Opportunities and Pitfalls in Drug Safety Studies after Marketing Approval

An Evaluation with a Focus on Older Patients

Dissertation

For the academic grade of

Doctor Public Health (Dr. PH)

Submitted by

Niklas Schmedt

University of Bremen

Faculty of Human and Health Sciences

Bremen, March 2016

First Reviewer: Prof. Dr. med. Hajo Zeeb

Second Reviewer: Prof. Dr. PH Falk Hoffmann

Date of Colloquium: July 8th, 2016

Acknowledgments

I would like to thank all those who contributed towards the completion of this thesis.

First, I am very grateful to my supervisor Prof. Dr. Edeltraut Garbe for her academic advice and her continuous and encouraging support of my scientific career over the last years.

I further thank the reviewers of this thesis, Prof. Dr. Falk Hoffmann and Prof. Dr. Hajo Zeeb, for their effort and advice.

I also extend my gratitude to my employer, the Leibniz Institute for Prevention Research and Epidemiology – BIPS, for giving me the opportunity to work with the German Pharmacoepidemiological Research Database (GePaRD).

In addition, I would like to take the opportunity to thank my BIPS colleagues and co-authors of the research articles which are part of this thesis. Special thanks go to Dirk Enders, Dr. Christoph Ohlmeier, Kathrin Jobski, Dr. Tania Schink and Dr. Heike Gerds for all the inspiring discussions and their helpful advice.

Finally, I would like to express my gratitude to my parents, family, friends, and my partner Natalie Skutzik for their support, moral and otherwise, over the last years and months.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADR	Averse drug reaction
AERS	U.S. Adverse Event Reporting System
BIPS	Leibniz Institute for Prevention Research and Epidemiology – BIPS
CPRD	Clinical Practice Research Datalink
CYP	Cytochrome P450
DDD	Defined Daily Dose
DRG	Diagnosis Related Group
EU	European Union
GePaRD	German Pharmacoepidemiological Research Database
HDPS	High-dimensional propensity score
NYHA	New York Heart Association
NSAID	Non-steroidal-anti-inflammatory drug
PhaSiNAg	Pharmacoepidemiological Safety Study of Neuroleptics and Antidepressants in the Area of Geriatric Psychiatrics
PML	Progressive multifocal leukoencephalopathy
PRR	Proportional reporting ratio
PS	Propensity score
PSS	Pharmacoepidemiological safety studies
RCT	Randomized controlled trial
ROR	Reporting odds ratio
SAFEGUARD	Safety Evaluation of Adverse Reactions in Diabetes
SHI	German statutory health insurance provider
START	Screening Tool to Alert to Right Treatment
STOPP	Screening Tool of Older Persons' Prescriptions
THIN	Health Improvement Network
VTE	Venous thromboembolism

Abstract

In the next decades, the number of older people will rise substantially in Germany. Likewise, drug treatment and polypharmacy will also augment due to the high prevalence of multimorbidity in this subpopulation. Since older people are often excluded from randomized controlled trials prior to drug approval, pharmacoepidemiological safety studies (PSS) based on spontaneous reporting systems and electronic healthcare databases often represent the only opportunity to investigate the safety of drugs in this population. However, these studies have specific methodological challenges related to the clinical characteristics of older patients and the nature of the data sources. Thus, the overall aim of this thesis is to (1) critically assess methodological challenges of PSS based on spontaneous reporting systems and electronic healthcare databases with a focus on older patients and (2) to define further areas of research to enhance their quality.

In this context, disproportionality analyses based on spontaneous reports, cohort studies, nested case-control studies, and case-only designs are introduced as study designs for PSS. Confounding (e.g., by frailty), outcome and exposure misclassification as well as time-related biases in PSS are illustrated as selected methodological challenges. These challenges are then discussed in the context of my research articles with a focus on older people, and opportunities to address these challenges are presented. More specifically, the role of spontaneous reporting systems in the detection of adverse drug reactions in older people is critically assessed. Afterwards, drug utilization studies as well as the application of high-dimensional propensity score methods and case-only designs are discussed as options to overcome the specific problem of confounding by indication and unmeasured confounding in PSS among older people. Further, a detailed review of the patient's profile is recommended to increase the specificity of the outcome case-algorithms in administrative claims databases. Moreover, it is highlighted that sensitivity analyses in drug utilization and safety studies are particularly important in the case of "as-needed" treatment among older patients and if information on the prescribed daily dose is missing. Finally, it is highlighted how time-related biases can be prevented in cohort and nested-case-control studies using a time-dependent analysis and risk-set sampling, respectively. In the conclusion, future research perspectives with regard to PSS in older patients are pointed out as, for instance, the use of semi-automated drug safety monitoring based on electronic healthcare databases, the availability of additional medical information in the context of the German "e-health" legislation or the need for external validation studies to study the impact of outcome and drug exposure misclassification.

1. Preface

This doctoral thesis was prepared based on several research articles attached in the appendix. The following sections are intended to integrate the content of these research articles into an overall public health context.

The following research articles are part of this doctoral thesis:

1. **Schmedt N**, Andersohn F, Garbe E. Signals of progressive multifocal leukoencephalopathy for immunosuppressants: a disproportionality analysis of spontaneous reports within the US Adverse Event Reporting System (AERS). Pharmacoepidemiology and Drug Safety 2012 Nov;21(11):1216-20.
   doi: 10.1002/pds.3320

In this article, we investigated safety signals of progressive multifocal leukoencephalopathy (PML) for different immunosuppressants based on spontaneous reports of possible adverse drug reactions in the U.S. Adverse Event Reporting System. Our analyses revealed signals of PML for several immunosuppressants including drugs previously not considered as potential risk factors for PML, e.g., azathioprine and cyclosporine.

2. **Schmedt N,** Garbe E. Antipsychotic drug use and the risk of venous thromboembolism in elderly patients with dementia. Journal of Clinical Psychopharmacology 2013;33:753-8.
   doi: 10.1097/JCP.0b013e3182a412d5

This article describes a nested case-control study based on the German Pharmacoepidemiological Research Database (GePaRD) in which we investigated whether the use of antipsychotics is associated with a higher risk of venous thromboembolism (VTE) in older patients with dementia. We found a higher risk of VTE for current users of antipsychotics in general and users of combination therapy of conventional and atypical antipsychotics compared to non-use. Among current users, the risk of VTE was only elevated for new users, i.e., during the first three months of treatment.

3. 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.
doi: 10.1186/s12889-015-1943-7

In this study, we compared mortality figures from GePaRD with external data from the Federal Statistical Office of Germany. We observed diverging mortality rates in several federal states that might result from the higher socioeconomic status of the study population in GePaRD compared to the overall population in Germany. In Bremen, where the socioeconomic representativeness is higher for GePaRD, the mortality rates were in good accordance with external data. In addition, the agreement of the percentage of hospital deaths in both data sources suggests completeness of outpatient mortality information in GePaRD.

4. **Schmedt N**, Jobski K, Kollhorst B, Krappweis J, Rüther E, Schink T, Garbe E. Treatment patterns and characteristics of older antipsychotic drug users in Germany. International Clinical Psychopharmacology 2016 May;31(3):159-69.
doi: 10.1097/YIC.0000000000000119

In this study, we explored the characteristics and treatment patterns of older antipsychotic users in Germany based on GePaRD. We found that most antipsychotics were frequently used for indications other than schizophrenia and bipolar disorders, e.g., dementia, depression, pain, vertigo or nausea. In addition, the baseline prevalence of co-morbidities and co-medications substantially differed for users of individual drugs. The observed low persistence in combination with a high number of treatment episodes suggests frequent “as-needed” treatment with antipsychotics, especially in patients with dementia. In general, this study can serve as a reference for drug utilization studies in other countries and provides important information for comparative safety studies in older antipsychotic users.

5. **Schmedt N**, Kollhorst B, Enders D, Jobski K, Krappweis J, Garbe E, Schink T. Comparative risk of death in older adults treated with antipsychotics: a population-based cohort study. European Neuropsychopharmacology 2016 Jul 27. doi: 10.1016/j.euroneuro.2016.07.006. [Epub ahead of print]

In this cohort study, we compared the risk of death in older patients initiating treatment with 15 different antipsychotics in Germany based on GePaRD. Since previous analyses indicated substantial differences for users of individual antipsychotics (Schmedt et al. 2016a), we also applied high-dimensional propensity score methods to further explore the potential for confounding. In summary, we found that the use of haloperidol, zuclopenthixol, and melperone is associated with an increased risk of death compared to risperidone. The same applies for the use of levomepromazine and chlorprothixene in patients aged 80 years and older and in those with dementia. While the slightly increased risk of death for melperone requires further investigation, our study suggests that haloperidol, zuclopenthixol, levomepromazine, and chlorprothixene should be avoided in older patients except in palliative care as long as other treatment options exist.

Further authored and co-authored research articles and “letters to the editor” dealing with the evaluation of drug safety after marketing approval are related to and cited in this thesis.

1. Amann U, **Schmedt N**, Garbe E. Prescribing of potentially inappropriate medications for the elderly: an analysis based on the PRISCUS list. Deutsches Arzteblatt international 2012;109:69-75.
   doi: 10.3238/arztebl.2012.0069

In this study, we explored the frequency of use of potentially inappropriate medications for older adults in Germany based on criteria of the PRISCUS list before its publication. The study highlights the need for drug safety studies in older drug users due to a frequent use of medications that might be associated with adverse drug reactions in this population.

2. Andersohn F, **Schmedt N**, Weinmann S, Willich SN, Garbe E. Priapism associated with antipsychotics: role of alpha1 adrenoceptor affinity. Journal of Clinical Psychopharmacology 2010 Feb;30(1):68-71
   doi: 10.1097/JCP.0b013e3181c8273d

The analyses described in this article were based on spontaneous reports of possible adverse drug reactions from the U.S. Food and Drug Administration. With this data, we conducted a disproportionality analysis and explored whether safety signals of priapism arose for antipsychotics depending on their alpha1 adrenoceptor affinity. As one article of this thesis (Schmedt et al. 2012) is based on the same data source, and disproportionality analyses of spontaneous reports are discussed as a tool to generate hypotheses of possible adverse drug reactions after marketing approval, this study is cited as another example of this method.

3. Mikolajczyk R, Horn J, **Schmedt N**, Langner I, Lindemann C, Garbe E. Injury prevention by medication among children with attention-deficit/hyperactivity disorder: a case-only study. JAMA Pediatrics 2015 Apr;169(4):391-5.
   doi:10.1001/jamapediatrics.2014.3275

In this article, we investigated whether treatment with methylphenidate and atomoxetine can prevent injuries among children with attention-deficit/hyperactivity disorder using the

self-controlled case series design. Since this case-only design is discussed as an option to prevent unmeasured confounding in drug safety studies, it is cited as an example of this methodological approach.

4. **Schmedt N**, Garbe E. Letter by Schmedt and Garbe regarding article, "statins and the risk of cancer after heart transplantation". Circulation 2013;127:e440. doi: 10.1161/CIRCULATIONAHA.112.135707

5. **Schmedt N**, Azoulay L, Hense S. Re.: "Reduced risk of lung cancer with metformin therapy in diabetic patients: a systematic review and meta-analysis". American Journal of Epidemiology 2014;180:1216-7. doi: 10.1093/aje/kwu311

In these “letters to the editor”, we pointed out that immortal time bias and time-window bias may have led to an overestimation of the beneficial effects of statins with regard to the occurrence of cancer after heart transplantation and of metformin regarding the risk of lung cancer in patients with diabetes. Since these forms of biases may also affect drug safety studies and can result in substantially biased study results, the “letters to the editor” are cited as examples.

In addition, some parts are based on unpublished analyses within the “Pharmacoepidemiological Safety Study of Neuroleptics and Antidepressants in the Area of Geriatric Psychiatrics (PhaSiNAg)” conducted at the Leibniz Institute for Prevention Research and Epidemiology – BIPS and funded by the Federal Institute for Drugs and Medical Devices. In this study, I was the project manager and epidemiologist.

2. Introduction

According to the 13th population projection of the Federal Statistical Office in Germany, the number of people aged 65 years and older will rise from 17 million in 2013 to 23 million in 2037 reflecting a relative increase of approximately 40%. The strongest relative elevation is expected for those aged 80 years and older with an 124% increase from 4 million in 2013 to 10 million in 2050 (Federal Statistical Office 2015). Since multimorbidity is prevalent in many older people and increases with rising age (Barnett et al. 2012, Marengoni et al. 2011), drug treatment and polypharmacy are likely to rise substantially in the next decades. In 2013, patients aged 65 years and older already received 55% of all "Defined Daily Doses" (DDDs) of drugs prescribed to members of German statutory health insurance providers (SHIs), although they represented only 22% of the total population (Schaufler and Telschow 2014). Compared to data from 2007, the total number of prescribed DDDs in this age group increased by 24% within 6 years (own calculations based on Coca and Nink 2008 and Schaufler and Telschow 2014).

Against this background, the evaluation of beneficial and especially of harmful drug effects in older patients is of high public health interest and will become even more important in the future. Although older patients may be particularly prone to serious adverse drug reactions (ADRs), they are often excluded from randomized controlled trials (RCTs) prior to drug approval due to restrictive inclusion criteria (e.g., Herland et al. 2005). Therefore, studies based on spontaneous reporting systems and electronic healthcare databases often represent the only opportunity to investigate the safety of drug use in this population after marketing approval (Garbe and Pigeot 2015, Garbe and Suissa 2014). Similar to pharmacoepidemiological studies in other subpopulations such as children or pregnant women, the evaluation of drug safety in older patients poses specific methodological challenges. Some of them refer to the clinical characteristics of older patients, others are related to the nature of the data sources.

Thus, the overall aim of this thesis is to (1) critically assess methodological challenges of pharmacoepidemiological safety studies (PSS) based on spontaneous reporting systems and electronic healthcare databases with a focus on older patients and (2) to define further areas of research to enhance their quality. Since four of five research articles included in this thesis are based on data from GePaRD, most discussions primarily focus on this data source.

The first part of this thesis provides important background information on the need for drug safety studies after marketing approval, data sources for drug safety studies and

specific aspects of drug use in older patients. Thereafter, the objectives are defined followed by a section briefly illustrating study designs available for drug safety studies and selected methodological challenges. Based on this, I discuss how these challenges can affect drug safety studies with a focus on older patients and how they can be addressed. In conclusion, future perspectives of research in this field are pointed out.

3. Background

3.1. Need for Drug Safety Studies after Marketing Approval

3.1.1. Legal Requirements for Drug Approval in Germany

From 1956 to 1961, thalidomide (Contergan®) was sold in Germany as a nonbarbiturate hypnotic sedative. At the time, the production and approval of drugs were not regulated for pharmaceutical companies in Germany and testing for possible harmful effects was not routinely undertaken in humans prior to drug marketing (Rehman et al. 2011). In 1961, Contergan® was withdrawn from the market by the manufacturer due to accumulated case reports of congenital malformations in newborns whose mothers had been exposed to the drug during pregnancy (Maio 2001). Later, it became clear that more than 10,000 malformed children were born worldwide due to exposure to Contergan® (Franks et al. 2004).

As a result of this tragedy, drug marketing and approval in Germany was increasingly regulated in the following years ending with the commencement of the second Pharmaceutical Drug Act in 1976, which is still in place in an updated version today. The key regulations of this law include requirements with regard to consistency, production and quality of the drug, drug approval, protection of humans in clinical trials, collection and analysis of drug-related risks, and liability in case of ADRs. Further, the quality, safety, and efficacy of a drug have to be demonstrated in appropriate studies before approval taking into account the respective directives of the European Union (Andersohn 2011).

According to the legal prerequisites, new drugs are investigated in preclinical animal and in vitro studies and in three phases of clinical trials in humans prior to marketing approval (Garbe and Suissa 2014). In phase I, a small number of individuals is exposed to the respective drug in order to obtain first information on tolerability as well as pharmacodynamic and pharmacokinetic aspects in humans. Phase II studies usually include several hundreds of patients and serve to define the appropriate dose range of the drug and to obtain first information on its efficacy and safety within the target population.

Last, phase III studies are mostly conducted as RCTs to prove the efficacy of the drug and to monitor possible ADRs compared to placebo or commonly used treatment alternatives in larger patient populations (approx. 1,000 to 3,000 persons) (Garbe and Suissa 2014).

3.1.2. Limitations of Randomized Controlled Trials

Pre-marketing RCTs provide important information on the efficacy and safety of drugs and efforts for drug approval are commonly terminated if ADRs are detected in these studies (Schuster et al. 2005). Nevertheless, many drugs are withdrawn after a long time on the market and after many patients have used them. Between 1980 and 2009, one out of seven newly approved drugs was later withdrawn by the U.S. Food and Drug Administration. Approximately 22% of withdrawals were related to safety concerns, i.e., 3.5% of all drugs approved during this period were withdrawn due to safety reasons (Qureshi et al. 2011). These withdrawals are the result of several limitations of RCTs prior to marketing approval.

First, the sample sizes of RCTs are usually not large enough to detect elevated risks of rare ADRs. For instance, a sample size of approximately 23,500 patients would be required per study arm to find a twofold increased risk of a drug compared to a reference group for an outcome with an incidence of 1 per 1,000 allowing for a beta error of 20% (Garbe and Suissa 2014). Since the assumed relative risk of an ADR in users of a specific drug compared to a reference category and the incidence of ADRs are usually much smaller, rare but severe ADRs can often not be identified in these studies.

Second, findings from RCTs usually cannot be extrapolated to routine care. Even phase III studies mostly have very restrictive inclusion criteria analyzing study populations who are not representative for subsequent drug users in clinical practice. Frequently, pregnant women, children, patients with severe co-morbidity, and especially older patients are not included or at least under-recruited (Garbe and Suissa 2014). For instance, Herland et al. (2005) showed that only 5.4% of a population with asthma from clinical practice and 17% with chronic obstructive pulmonary disease would fit into the restrictive criteria routinely used in RCTs.

Third, phase I to III trials are conducted under experimental conditions in specialized units by well-trained staff. Thus, study participants are closely monitored and treatment in clinical trials follows a standardized regimen for which adherence is tightly controlled. However, poor medication adherence and large treatment variations depending on individual patient characteristics and response to therapy are common in clinical practice (Garbe and Suissa 2014). This especially applies to older patients in whom medication

adherence may be lower in the case of polypharmacy (Pasina et al. 2014) and for whom many drugs are initiated at a lower dose and slowly up-titrated (Wehling and Burkhardt 2013).

Fourth, RCTs usually have a short follow-up which impedes the detection of so-called "type D ADRs" with a long induction period or occurring after cumulative drug exposure, e.g., vaginal adenocarcinoma after exposure to diethylstilbestrol during pregnancy or tardive dyskinesias after treatment with antipsychotics (Edwards and Aronson 2000).

For all these reasons, possible harmful effects of drugs have to be further investigated intensively after marketing approval.

3.2. Data Sources for Drug Safety Studies

3.2.1. Case Reports and Spontaneous Reporting Systems

Once a drug has been approved and is available on the market, several data sources can be used for post-marketing surveillance of its safety profile. The first is the assessment of case reports of possible ADRs published in the literature or sent to the responsible authorities by health professionals, consumers or manufacturers. Case reports of possible ADRs can be evaluated qualitatively based on established criteria for causality assessment (e.g., World Health Organization 2016a) or quantitatively using different statistical methods for so-called disproportionality analyses as soon as they have been entered into spontaneous reporting systems (Almenoff et al. 2007, Bate and Evans 2009). Information on case reports of possible ADRs usually include age, sex, suspected drugs and co-medication, the date of initiation and discontinuation of drug treatment, date and severity of the possible ADR, underlying indications for drug treatment, etc. In Germany, the Federal Institute for Drugs and Medical Devices and the Paul-Ehrlich Institute collect and analyze spontaneous reports of possible ADRs (Federal Institute for Drugs and Medical Devices 2016, Paul-Ehrlich Institute 2016) which are also combined for analyses on the supranational level by the European Union (European Medicines Agency 2016) and the World Health Organization (World Health Organization 2016b). The U.S. Food and Drug Administration gathers spontaneous reports from the United States and other countries, which are freely available on the internet for research purposes (U.S. Food and Drug Administration 2016). Although the assessment of case reports is particularly important for the detection of rare ADRs and is recognized as a useful tool to generate hypotheses about unknown ADRs at an early stage after marketing approval, the analysis of case reports is prone to various types of biases (see section 6.1). In addition, it is not

possible to determine risks of ADRs due to underreporting, stimulated reporting, and missing denominator information on individuals exposed to the drug, and the data quality of spontaneous reports is often poor (Garbe and Suissa 2014). Therefore, other data sources such as field studies or electronic healthcare databases have to be used to conduct PSS of higher quality to investigate the safety profile of newly approved drugs in the real world setting.

3.2.2. Primary vs. Secondary Data

Data sources for PSS after marketing approval can be subdivided into primary data gathered in field studies and secondary data mostly obtained by health insurance providers for reimbursement purposes or available as electronic medical records from general practitioners or hospitals (Strom 2012).

Field studies have significantly contributed to the knowledge of ADRs of drugs after marketing approval, e.g., with regard to the risk of VTE associated with the use of oral contraceptives (Lewis et al. 1997, Spitzer et al. 1996), and they are often conducted based on multi-purpose cohorts such as the Women's Health Initiative Study (Bavry et al. 2014, Desai et al. 2013). Other examples of PSS based on primary data are the Berlin Case-Control Surveillance Study (Andersohn et al. 2004) as well as studies from the pharmacovigilance center "Embryotox" (Weber-Schoendorfer and Schaefer 2016) or the German biologic registry (Strangfeld and Zink 2014). Even though field studies have to be considered a valuable tool for specific research questions, PSS based on secondary data have recently become more popular. Since their first use in the 1980s in North America, large electronic healthcare databases have increasingly been established worldwide (Strom 2012). This development did not occur by chance, but can be attributed to the advantages of using secondary data compared to primary data. An overview and a comparison of the characteristics of primary and secondary data for PSS are shown in Table 1.

Table 1 Comparison of primary and secondary data for pharmacoepidemiological studies (modified from Andersohn and Garbe 2008)

Characteristic	Primary data	Secondary data
Number of study participants	Lower	High
Time required for the study	High	Lower
Costs of the study	High	Lower
Calculation of the incidence of rare events possible	Mostly not	Yes
Study enrollment of patients with severe diseases possible	Mostly not	Yes
Exposure information of over-the-counter drugs available	Yes	No
Recall bias possible	Yes	No
Interview bias possible	Yes	No
Selection bias due to non-response possible	Yes	No
Information on life-style-related factors (e.g., alcohol consumption, smoking, physical activity) available	Yes	Mostly not (partly in medical record databases)
Laboratory test results and other clinical parameters available	Yes	Mostly not (partly in medical record databases)
Information on disease severity available	Yes	Mostly not

As described above, the outcomes of interest are usually rare events which could not be investigated adequately in pre-marketing RCTs, and PSS often require a very large sample size of 10,000 exposed persons or more combined with a respective control group of the same or larger size. Field studies of this size would be very time consuming and expensive. In contrast, studies based on electronic healthcare databases can be conducted much faster and are more cost-efficient, since the data have already been collected. Since safety issues usually need to be investigated rapidly to address regulatory and public health crises, electronic healthcare databases are often considered the preferred data source for PSS after marketing approval (Strom 2012). A further advantage of secondary data is that they include data of populations rarely investigated in RCTs or extremely difficult to acquire for field studies, e.g., older persons living in nursing homes or patients with distinct multimorbidity. For these patients, information on drug exposure is assumed to be complete for most drugs in electronic healthcare databases, whereas its collection is prone to recall or interviewer bias in field studies (Schneeweiss and Avorn 2005). On the other hand, electronic healthcare databases also have disadvantages compared to field studies, since they often lack information on possible important confounders in PSS (e.g., smoking, alcohol consumption or physical activity), and the validity of diagnoses used for identification of outcome events may be limited (Strom 2012). In field studies, confounder information can already be considered at the stage of data collection and the occurrence of the outcome can be assessed according to previously defined clinical definitions.

3.2.3. Administrative Claims vs. Medical Record Databases

As mentioned above, there is a large amount of electronic healthcare databases for PSS after marketing approval worldwide. These include different types of databases depending on the health care system in which they were generated. They are usually categorized into administrative claims databases and medical record databases (Strom 2012).

Administrative claims data typically appear when individuals use the health care system, e.g., when they receive drug dispensations, require hospital care or contact registered physicians in ambulatory care. In many health care systems, these services are paid for by health insurance providers or the state. For reimbursement purposes and quality control, specific information on these services is documented electronically. Commonly, data on pharmacy dispensations, hospital data, and ambulatory care data are reported to the patient's health insurance provider and can be used together with available sociodemographic information in PSS. In other cases, records from different areas of care

and registries (e.g., death or birth registries) are available and can be linked by an unique identifier (Strom 2012).

Medical record databases were developed more recently than administrative claims databases and are a consequence of the increasing computerization of documentation in medical practices and hospitals (Strom 2012). Two well-known examples are the Clinical Practice Research Datalink (CPRD) (Herrett et al. 2015) and the Health Improvement Network (THIN) (Hall 2009). Both databases from the United Kingdom primarily include medical records from ambulatory care; however, information on hospitalizations has been recently made available via linkage (Herrett et al. 2015).

Administrative claims databases have the advantage that they usually include data from all sites of care needed for PSS, i.e., pharmacy dispensations, hospitalization, and ambulatory care. Since the identity of the drug and the dispensed amount of the drug define the price for reimbursement, pharmacy claims are subject to high standards of quality control and can therefore be assumed to provide very accurate data (Schneeweiss and Avorn 2005, Strom 2012). The same applies for hospital data mostly reimbursed based on Diagnosis Related Groups (DRGs) and thus relying on appropriate coding of diagnoses. However, concerns have been raised regarding the practice of upcoding of diagnoses by hospitals to receive the largest possible payment (Lüngen and Lauterbach 2000). A major limitation of administrative claims data is often related to the validity of ambulatory diagnoses because their correct coding does not influence the physician's payment and consequently there is no incentive for correct coding in contrast to hospital data (Strom 2012). Compared to administrative claims data, medical record databases include very detailed information on diagnostic data from ambulatory care, but data from other physicians and other sites of care, especially hospital diagnosis data and dates of hospitalization, may be incomplete or missing if only data from the general practitioner are entered into the system (Andersohn and Garbe 2008).

3.2.4. Secondary Data in Germany

In Germany, secondary data from various sources are available for public health research purposes. The majority of Germans are insured with an SHI, whereas private health insurance providers are only accessible to a small proportion, i.e., to employees with an income higher than 56,250€[1] and specific occupational groups as for instance self-

[1] income threshold for compulsory health insurance in 2016 assessed at https://www.bundesregierung.de/Content/DE/Artikel/2015/10/2015-10-14-sozialversicherung.html

employed persons or civil servants. In 2014, 70.3 out of 81.2 million people in Germany (approx. 87%) were members of one of the 118 different SHIs, including 15.5 million members aged 65 years and older (Federal Ministry of Health 2016). Secondary data available from the SHIs for research purposes mainly comprise sociodemographic data, hospital data, ambulatory care data, and outpatient drug dispensing data, but also data on medical devices as well as dentist and nursing care data (Swart et al. 2014). However, the use of secondary data in Germany is strongly regulated by law. According to § 75 of the Social Act X, social data can be delivered for research purposes only if the public health interest of the project substantially outweighs data privacy concerns of the individuals. All research projects based on pseudonymized data from SHIs have to be approved by the respective SHI and the responsible authority (e.g., the German Federal Insurance Office for Germany-wide acting SHIs) (Garbe and Pigeot 2015).

Since four of five research articles included in this thesis were based on GePaRD, its characteristics and structure are introduced in more detail.
GePaRD was established by the Leibniz Institute for Prevention Research and Epidemiology – BIPS and currently consists of administrative claims data from four SHIs with information on about 20 million people. Thus, it represents nearly 25% of the German population from all federal states of Germany. The database comprises sociodemographic core data, hospital data, outpatient dispensation data, and outpatient care data of all individuals enrolled in one of the four SHIs since 2004. Sociodemographic core data contain information on sex, year of birth, insurance status, and reason for deregistration from the SHI. Hospital data comprise the date of admission and discharge, different types of diagnoses (admission, main discharge, secondary and ancillary diagnoses), diagnostic and therapeutic procedures, and the reason for hospital discharge (incl. death). Data from outpatient physician visits incorporate diagnoses on a quarterly basis including the diagnostic certainty (assured, suspected, excluded, and status post diagnosis), and types and dates of diagnostic and therapeutic procedures. All diagnoses are coded according to the German Modification of the 10th revision of the International Classification of Diseases. Dispensation data contain information on prescriptions dispensed in a pharmacy and reimbursed by the respective SHI. Drugs purchased over the counter and most of the medications administered in hospitals are not contained in GePaRD. Dispensation information also includes the dates of the prescription and dispensation, the number of prescribed packages, the specialty of the prescribing physician, and the central pharmaceutical number of the drug. Based on a pharmaceutical reference database,

information on the generic and brand name of the drug, packaging size, strength, DDD, and other pharmaceutical information can be linked to GePaRD (Garbe et al. 2011, Pigeot and Ahrens 2008).

3.3. Drug Use in Older Patients

Clinical pharmacology investigates the response to and the effect of drug treatment on the individual level while pharmacoepidemiology aims to measure the effects of drug treatment on the population level using epidemiological methods. In other words, pharmacoepidemiology is the application of epidemiological methods to clinical pharmacology (Hennessy 2006). From a pharmacological point of view, each application of a drug to a patient has to be regarded as a single experiment with an unknown outcome depending on individual characteristics of the patient and the correct application of the drug (Wehling and Burkhardt 2013). Several factors may influence the pharmacological response to a drug and may lead to substantial variation of drug effects in single patients. This has to be taken into account in the interpretation of PSS and especially applies to older patients. The most important factors in this context are reduced physiological resources or frailty leading to alterations in pharmacokinetics and pharmacodynamics as well as multimorbidity accompanied by polypharmacy and resulting drug interactions. In the following, the most important factors with a possible impact on individual treatment outcomes in older patients are described in more detail. In addition, the PRISCUS list and other criteria defining potentially inadequate medication in older patients are introduced.

3.3.1. Alterations in Pharmacodynamics and Pharmacokinetics

In clinical pharmacology, the study of drugs is usually subdivided into pharmacodynamics and pharmacokinetics. While pharmacodynamics concern the biochemical and physiological effects of the drug on the human body, pharmacokinetics focus on the process of absorption, distribution, and elimination of the drug in the human body (Hennessy 2006). In older patients, physiological changes typically occurring during the aging process or mediated by co-morbidity can cause alterations of pharmacodynamics and pharmacokinetics. The major concern in this context is usually an elevated plasma concentration causing possible drug toxicity and ADRs, especially for drugs with a narrow therapeutic window, e.g., cardiac glycosides or several anti-epileptics frequently used in older patients (Turnheim 2003, Turnheim 2004).

Pharmacokinetic alterations with increasing age are well understood and seem to have a greater impact on the variability of drug effects in older patients compared to pharmacodynamics. The most critical aspect is the accumulation of the drug due to reduced clearance from the body which might also lead to toxicity and ADRs. Impaired clearance is mainly caused by a reduced glomerular filtration rate for drugs being eliminated via the kidneys ("renal clearance") and a reduced hepatic blood flow or decreased capacity of enzymes for drugs being first metabolized in the liver ("hepatic clearance") (Wehling and Burkhardt 2013).

The glomerular filtration rate usually declines with increasing age due to an elevated failure of nephrons, but its reduction can also be triggered by co-morbidity such as atherosclerosis or type 2 diabetes mellitus. Renal clearance of a specific patient generally depends on age, weight, sex, and plasma creatinine concentration as a measure of renal function (Wehling and Burkhardt 2013). In contrast to hepatic clearance, it can be calculated easily in a "bed-side" approach by the physician using for instance the Cockcroft and Gault formula (Cockcroft and Gault 1976). One frequent mistake is that physicians only use the plasma creatinine concentration as a measure for renal clearance, although both formation and elimination of creatinine are reduced in older patients. This results in normal plasma creatinine concentration even in case of an impaired glomerular filtration rate (Turnheim 2003). It is estimated that this phenomenon accounts for approximately 25 percent of all preventable ADRs (Wehling and Burkhardt 2013).

While renal elimination mainly affects water soluble drugs, many other lipid soluble drugs cannot be eliminated via the kidneys and have to be transferred into water soluble products first. This metabolism is usually performed in the liver and catalyzed by the so-called cytochrome P450 (CYP) isoenzymes. Besides frequent genetic polymorphisms of these enzymes, hepatic disease, reduced hepatic blood flow due to other age-related diseases as, for instance, heart failure and a generally reduced activity of CYP isoenzymes can lead to drug toxicity due to diminished metabolism of the drug in older patients (Wehling and Burkhardt 2013). Since hepatic clearance may work despite severe liver disease, it is still an uncommon cause of impaired clearance compared to a reduced glomerular filtration rate (Hennessy 2006). Other factors that might lead to variations in drug responses in older patients involve the absorption and the distribution of the drug but are considered to have minor impact compared to renal and hepatic elimination and are therefore not further discussed here (Hennessy 2006, Turnheim 2003).

The effect of age on pharmacodynamics has not been well investigated and is more difficult to analyze compared to pharmacokinetics. Probably, physiological changes and

age-induced alterations in target-structures of the drug, e.g., with regard to the number of receptors, as well as impaired homeostatic regulation are responsible for different responses to drugs and a higher risk of ADRs in older patients (Hennessy 2006). One important example is the increased sensitivity of some older patients to benzodiazepines leading to strong sedative effects even at a very low dose which may partly be related to disturbed circadian rhythm and reduced nerve conduction velocity (Wehling and Burkhardt 2013).

In summary, impaired renal elimination in first instance and secondly reduced hepatic metabolism may lead to a higher risk of drug toxicity in older patients. In the interpretation of PSS, it is important to recognize that older patients may vary substantially with regard to physiological alterations meaning that responses to drugs and their beneficial and unintended effects will be even more heterogeneous than in younger populations.

3.3.2. Multimorbidity, Polypharmacy, and Drug Interactions

Increasing age is accompanied by a higher prevalence of age-related disease and multimorbidity (Barnett et al. 2012) constituting a major challenge for drug treatment, especially if more than one physician is involved in the care of the same patient. If other co-morbidity and co-medication is not adequately considered, several drugs are administered concomitantly which might lead to drug interactions and increased risks of ADRs (Wehling and Burkhardt 2013). Typical ADRs resulting from drug interactions are neuropsychological syndromes, hypotension with related falls as well as acute renal failure (Mallet et al. 2007). The occurrence of ADRs not recognized by physicians might lead to “prescribing cascades”, i.e., additional polypharmacy by the initiation of further drugs to treat unrecognized ADRs (Rochon and Gurwitz 1997).

Drug interactions can occur on the level of pharmacodynamics and pharmacokinetics. On the one hand, different drugs may compete for the same target structures, e.g., for the same receptor group which might lead to increased plasma concentrations of the drugs. On the other hand, drug interactions may emerge in the hepatic metabolism of the drugs if two or more substrates are metabolized by the same CYP isoenzyme or if drugs act as inducers or inhibitors of specific CYP isoenzymes. This can result in enhanced or diminished metabolism, reduced efficacy or possibly toxicity (Hennessy 2006, Turnheim 2003). An overview of possible interactions during hepatic metabolism via CYP isoenzymes can be derived from the Flockhart list (Flockhart 2011).

Although not all drug-drug interactions have clinical relevance and lead to ADRs, their prevalence is probably often underestimated and misinterpreted in clinical practice. This

especially applies to older patients who are more prone to ADRs due to frailty, reduced homeostasis, and multimorbidity and therefore constitutes a preventable cause of morbidity and mortality in the older population (Mallet et al. 2007). In PhaSiNAg, we investigated the concurrent use of potentially interacting drugs for selected antipsychotics according to the Flockhart list (Flockhart 2011). For instance, nearly two third of haloperidol users received concurrent treatment with a CYP2D6 inhibitor (data not shown in research article) leading to a possible diminished metabolism and increased plasma concentration of haloperidol. In a Swedish study by Johnell and Klarin (2007), the prevalence of potentially clinically relevant drug-drug interactions was 26% in the study population aged 75 years and older receiving at least two drugs concomitantly. The prevalence of potentially serious drug-drug interactions was five percent and strongly associated with the overall number of drugs used. Another study from Canada (Juurlink et al. 2003) investigated whether older patients admitted to the hospital with specific drug toxicities were more likely to have received an interacting drug in the week prior to admission. For instance, patients hospitalized due to digoxin toxicity were 12 times more likely to have been treated with the interacting antibiotic clarithromycin. Patients treated with ACE-inhibitors who were hospitalized due to hyperkalemia were even 20 times more likely to have been prescribed a potassium-sparing diuretic (Juurlink et al. 2003).

In summary, multimorbidity constitutes a major problem in drug therapy of older patients. In the cases of polypharmacy and of several involved prescribing physicians, ADRs due to drug-drug interactions are likely and a preventable cause of morbidity and mortality in the older population.

3.3.3. Potentially Inappropriate Medication in Older Patients

In the last two decades, many activities were undertaken to improve the quality of prescribing in older patients by defining lists of potentially inappropriate medications (Gallagher et al. 2008, Laroche et al. 2007, McLeod et al. 1997, Rognstad et al. 2009). The most famous list is the Beers Criteria for determining potentially inappropriate medication in older patients which was first developed for licensed drugs in the United States and has been updated regularly (Beers 1997, Fick et al. 2003, American Geriatrics Society Beers Criteria Update Expert Panel 2012, American Geriatrics Society Beers Criteria Update Expert Panel 2015). For Germany, potentially inappropriate medication was defined for the first time in the PRISCUS list in 2010 in which 83 drugs from 18 drug classes were classified as inadequate for the use in older patients based on expert consensus (Holt et al. 2010). This list includes drugs possibly associated with a higher risk

of ADRs due to their pharmacological properties as well as recommendations for possible treatment alternatives and actions to be taken if these drugs are still used in older patients (Holt et al. 2010). Just recently, the EU(7)-PIM list of potentially inappropriate medications was developed by consensus of experts from seven European countries (Renom-Guiteras et al. 2015).

We assessed the status of potentially inadequate prescribing in older patients in Germany according to the PRISCUS list before its publication to evaluate its relevance and to establish a reference for following studies. Approximately 25% of the study population received at least one prescription of a potentially inadequate medication, and the prevalence increased with age. Nearly nine percent received even four and more different potentially inadequate medications within the same year (Amann et al. 2012). An analysis by Linder et al. (2014) based on data from one large SHI showed that the percentage of older patients with at least one prescription of a potentially inappropriate medication decreased from 21.7% to 19.9% between 2008 and 2012.

Nevertheless, it has to be noted critically that the PRISCUS and most other lists defining potentially inappropriate medications remain superficial, since they do not account for important clinical characteristics such as concomitant use of other drugs, co-morbidity, and the disease severity of the indication. In this context, the Screening Tool of Older Persons' Prescriptions (STOPP) and the Screening Tool to Alert to Right Treatment (START) have been developed and updated by Gallagher and colleagues to take into account more clinical information (Gallagher et al. 2008, O'Mahony et al. 2015). In studies from Ireland and Sweden, inappropriate prescribing according to the STOPP criteria has been associated with an increased risk of avoidable adverse drug events in older patients compared to those without such drug prescriptions (Hamilton et al. 2011, Hedna et al. 2015). Unfortunately, not all clinical information for the application of the STOPP criteria, e.g., the severity of heart failure according to the "New York Heart Association (NYHA) Functional Classification," is available in detail in administrative claims databases like GePaRD. Therefore, their applicability in this setting is limited.

4. Objectives

Against this background, the evaluation of drug safety after marketing approval is of high public health interest. In this context, the subpopulation of older people receiving drug therapy is at particular risk of ADRs and will become more important in the future. Since

PSS in older patients based on spontaneous reporting systems and electronic healthcare databases face specific challenges, the objectives of this thesis are

1. To critically assess methodological challenges of PSS after marketing approval based on spontaneous reporting systems and electronic healthcare databases with a focus on older patients and
2. To define further areas of research to enhance the quality of PSS in older patients.

Basic pharmacoepidemiological study designs and some important methodological challenges of PSS will be introduced and discussed in the next sections. This discussion also includes a critical post-hoc evaluation of the research articles and outlines possible improvements of my research.

5. Methods for Drug Safety Studies

5.1. Study Designs

In the following, the basic study designs applied in the PSS described in this thesis are introduced (Schmedt et al. 2012, Schmedt and Garbe 2013a, Schmedt et al. 2016b). In general, pharmacoepidemiology applies classic epidemiological study designs, i.e., the (nested) case-control study and the cohort study with some specific adaptations mainly motivated by the nature of the exposure and electronic healthcare databases as primary data source (Schneeweiss and Avorn 2005). Although studies based on spontaneous reports cannot be interpreted in the same way as observational studies, the analysis strategy of a disproportionality analysis is similar to a case-control or cohort study. Besides classic epidemiological studies, case-only designs developed for specific drug-safety-specific research questions are described.

5.1.1. Disproportionality Analyses

Spontaneous reporting systems only include case reports of possible ADRs submitted to the responsible authority. Since not all suspected cases of ADRs are reported and no information on the overall number of drug users is available, it is not possible to calculate incidence rates or risks based on this data. However, it is possible to check whether the proportion of reports of a specific ADR is higher among all reports of a specific drug compared to the reports of all other drugs (disproportionality analysis). The underlying

assumption is that an unexpected high reporting of a specific drug-ADR combination might point to a possible risk of the ADR for this drug (Garbe and Suissa 2014).
As statistical measures of disproportionality, the proportional reporting ratio (PRR) and the reporting odds ratio (ROR) are frequently used (Bate and Evans 2009). These can easily be calculated based on a basic two-by-two table (Table 2).

Table 2 Calculation of the proportional reporting ratio (PRR) and the reporting odds ratio (ROR) in spontaneous reporting systems

	Event of interest	All other reported events
Drug of interest	a	b
All other reported drugs	c	d

Note: a,b,c and d are absolute frequencies

The PRR is calculated as a/(a+b) divided by c/(c+d) and the ROR is calculated as a*d/b*c. In most cases, both measures are computed in conjunction with the respective chi-square value. A higher PRR or ROR indicates that the drug-ADR combination was reported more frequently than expected based on all other reports, i.e., the higher the PRR/ROR, the stronger the disproportionate reporting. An advantage of the ROR compared to the PRR is that multivariable logistic regression with additional adjustment for possible confounders as, for instance, age and sex can be used for its calculation. Besides the PRR and the ROR, more complex methods based on Bayesian statistics such as the Empirical Bayes Geometric Mean and the Information Component have been developed. These approaches are particularly useful to decrease the number of false positive signals if many drug-ADR combinations with a low observed number of events are investigated (Bate and Evans 2009). As threshold for a signal, Evans et al. (2001) proposed a PRR =>2, a chi-square value of at least four, and a minimum of three exposed events of interest with the respective drug of interest.
Once a signal for a possible drug-ADR combination has emerged, it has to be further evaluated based on a detailed review of case reports or PSS of higher quality, e.g., in observational studies based on electronic healthcare databases (Garbe and Suissa 2014).

5.1.2. Cohort Studies

Many PSS based on electronic healthcare databases are conducted as retrospective cohort studies. In comparison to cohort studies in classic epidemiology, they differ predominantly with regard to the cohort entry which may be defined as date in time,

occurrence of specific a diagnosis or event or prescription of a drug (Garbe and Suissa 2014). While cohorts beginning at a specific calendar date are rare, cohort entry at diagnosis or drug prescription is more frequent. An important problem of pharmacoepidemiological cohort studies based on electronic healthcare databases is left truncation of the data which impedes obtaining the complete history of previous diseases and medications. For event-based cohorts, bias can be introduced if the duration or severity of the disease event leading to cohort entry and exposure are both associated with the outcome under study (Garbe and Suissa 2014). For prescription-based cohorts, bias can arise in PSS if only prevalent users of the drug are included and patients with ADRs at an early stage of treatment are not considered after discontinuation of treatment ("depletion of susceptibles") (Moride and Abenhaim 1994). Another problem of cohort studies is that the performance of the analysis and the interpretation may become very complex as soon as the exposure status changes during follow-up and exposure has to be assessed time-dependently. However, this may be necessary, since time-fixed analyses can introduce severe biases (immortal time bias, see section 5.2.4) or biased results through contamination of the exposure groups.

5.1.3. Nested Case-Control Studies

The nested case-control study also constitutes a frequently used design in pharmacoepidemiological research based on electronic healthcare databases and actually reflects an analysis strategy within a previously defined cohort. The entry of such a cohort can be defined as a date in time, a specific age, the date of a specific diagnosis or the prescription of a drug. The cohort exit is usually defined as the end of the study period, death, the occurrence of the outcome or loss-to-follow-up, e.g., due to ending insurance membership (Garbe and Suissa 2014).

Conducting a nested case-control analysis basically entails four steps. First, the time axis for the cohort has to be defined which can be calendar time or follow-up time. Second, all cases with the outcome event have to be identified and selected at the first occurrence during follow-up. Third, the risk set of possible controls has to be determined for each case. It consists of all individuals in the cohort at risk for the outcome at the date of the outcome occurrence of the respective case (index date). Last, a predefined number of controls has to be selected out of the risk set of each case (risk-set sampling). By definition, the control and the case have the same the index date (Essebag et al. 2003). In this context, the controls have to be considered as control person moments in time, i.e., every individual who served as a control can become a case later during follow-up. An

alternative approach to sampling controls from the risk-set at the index date is the random assignment of a person moment at risk for the outcome from the cohort as a control. Cases and controls are mostly matched by parameters such as age and sex, and statistical analysis is performed using conditional logistic regression. If the distribution of the exposure during person time of the controls is representative for the whole cohort, the obtained odds ratio can be interpreted as an efficient estimator of the rate ratio for both sampling approaches described above (incidence density sampling) (Garbe and Suissa 2014, Rothman et al. 2008a).

One major advantage of the nested case-control analysis based on electronic healthcare databases is that it can be performed more efficiently compared to the full cohort analysis with time-depending exposures. While the risk-set of each case consists of all controls at the index date using cox regression in the full cohort analysis, the nested case-control approach only requires a random sample of it (Garbe and Suissa 2014).

5.1.4. Case-Only Designs

Besides nested case-control and cohort studies, case-only designs are frequently used to overcome certain difficulties of pharmacoepidemiological studies. As described below, confounding by indication or unmeasured confounding might be a serious problem if both the underlying indication or its severity and the drug exposure are associated with the outcome or if important confounders cannot be assessed in electronic healthcare databases. Under these circumstances, case-only designs may serve as alternatives if the key assumptions of the respective study design are fulfilled (Garbe and Suissa 2014). In the following, the case-crossover and the self-controlled case series design as case-only designs predominantly used in pharmacoepidemiology are introduced in more detail.

5.1.4.1. Case-Crossover Design

The case-crossover design can be considered a case-control study in which cases serve as their own controls. For each case, exposure status within a predefined risk period prior to the event and within one or more control periods is assessed. The statistical analysis is performed as conditional logistic regression taking into account the matched nature of the data. In order to apply this design, three requirements have to be fulfilled. First, the outcome under study must be an acute ADR resulting from a transient drug effect. Possible ADRs with a long induction period and drugs used continuously over a long time period, e.g., antihypertensive or antidiabetic medication, are difficult to investigate based

on this design. Second, the time window of the drug effect has to be defined exactly as the period under risk of the ADR. Third, data for a sufficiently long time period have to be available for each case to assess information on the exposure prior to the event assuming that the probability of drug exposure has not changed over time (Maclure 1991). If the latter is not the case, the case-time-control design proposed by Suissa (1995) can be used as an extension of the case-crossover design to account for time-trend biases.

5.1.4.2. Self-Controlled Case Series Design

The self-controlled case series design was developed by Farrington (Farrington 1995). At the beginning, it was primarily used to study acute ADRs of vaccinations, but later it was also applied to investigate other drug exposures and non-acute outcomes, e.g., autism. Similar to the case-crossover design, it can be used to analyze possible ADRs of transient exposures considering only individuals who developed the outcome of interest. However, it rather conveys the idea of a cohort study and several events per individual can contribute to the analysis which is not possible in the case-crossover design (Garbe and Suissa 2014).

Basically, the self-controlled case series method is used to estimate the relative incidence of the possible ADR or outcome of interest during a predefined risk period after drug exposure (risk period) compared to the remaining time of the individual during follow-up as control period. At first, the study time window is defined which can be a specific age-range or calendar time period. Afterwards, all individuals who experienced the outcome of interest are selected and the follow-up time during the study time window is determined. Then, drug exposure is assessed for all cases and the risk period has to be defined depending on the pharmacological properties of the respective study drug and mechanisms leading to the outcome event. Thus, the whole observation period for each case can be subdivided into periods at risk and control periods and the date of the outcome events can be assigned to one of these periods. The relative incidence of the outcome between risk periods and control periods is then modeled using conditional Poisson regression (Garbe and Suissa 2014, Whitaker et al. 2006).

In general, three key assumptions have to be met in the basic approach. First, the occurrence of the outcome has to be the result of a non-homogeneous Poisson process. Second, the occurrence of an outcome must not affect the probability of drug exposure and, third, the outcome event must not censor or influence the observation period of an individual otherwise (Whitaker et al. 2009). For this reason, death and severe events leading to death shortly after their occurrence should not be investigated according to the

basic approach; however, extended methods have been developed to study such events (Whitaker et al. 2006).

5.2. Methodological Challenges in Drug Safety Studies

In this section, selected methodological challenges of PSS based on electronic healthcare databases are introduced. The selection of topics was based on aspects that had to be addressed in the research articles included in this thesis. The main limitations of disproportionality analyses refer to the quality and reporting of spontaneous reports and to a lesser extent to methodological aspects. Therefore, these limitations are not presented in this section as methodological aspects, but highlighted and discussed later in section 6.1. However, some challenges described hereafter can also affect studies based on spontaneous reporting systems, e.g., confounding as well as outcome and exposure misclassification.

5.2.1. Unmeasured Confounding, Confounding by Indication and Channeling Bias

As in all non-randomized observational research, one of the most important challenges in PSS after marketing approval is confounding bias (Schneeweiss and Avorn 2005). As described above, many administrative claims databases as GePaRD miss detailed information on important possible confounders, e.g., body weight, history of smoking, alcohol consumption, and other lifestyle-related factors such as physical activity (see section 3.2.2). If these factors are associated with both outcome and exposure status, unmeasured and residual confounding may lead to biased risk estimates and as a consequence to the misinterpretation of study results.

A general concern in pharmacoepidemiological studies after marketing approval is confounding by indication. If drug use is compared to non-use, it occurs if the indication for the drug itself is a risk factor for the outcome under study (Schneeweiss and Avorn 2005). In comparative effectiveness and safety studies, confounding by indication often appears if the drug under study and the comparator drug are indicated in patients at different stages of the treated disease (confounding by disease severity). If the disease severity is associated with the outcome, the drug used in those most severely affected by the illness will be disadvantaged by confounding (Garbe and Suissa 2014, Walker 1996). Of note, confounding by indication more often affects studies that investigate the intended effects of drugs because the indication for drug use is usually associated with the outcome by definition. In contrast, it is less common in studies investigating ADRs, since

these are usually unintended and not associated with the indication of the drug (Strom and Melmon 2012). For instance, confounding by indication will be a problem if the risk of myocardial infarction in statin users is compared to non-users, but not if the outcome under study is rhabdomyolysis. The reason behind this is that statins are indicated for hypercholesterolemia and for secondary prevention of cardiovascular events, both of which strong risk factors for myocardial infarction, but not for the unintended ADR rhabdomyolysis.

Even if the drugs compared with regard to a specific outcome are used for the same indication, drug prescribing by physicians in routine care may be influenced by patient characteristics leading to confounding by drug channeling to patients at high risk for the outcome (channeling bias) (Schneeweiss and Avorn 2005). For example, MacDonald et al. (2003) showed that meloxicam and cyclooxygenase-2 inhibitors as newer class of non-steroidal-anti-inflammatory drugs were predominantly prescribed to patients at very high risk for gastrointestinal bleeding in clinical practice because they were supposed to have gastro-protective effects compared to classic NSAIDs. This prescribing pattern probably led to the wrong conclusion in other studies that the use of these newer NSAIDs is not associated with a beneficial effect with regard to the risk of gastrointestinal bleeding (MacDonald et al. 2003). Another example is the higher risk of falls and hip fractures observed for newer benzodiazepines compared to older agents. In this case, newer agents were preferentially prescribed to frail older patients who are at higher risk for such ADRs resulting in unmeasured confounding and an overestimation of the risk of fall and hip fractures compared to older agents (Schneeweiss and Wang 2005).

In PSS among older patients, confounding by frailty is often a problem. Frailty is defined as higher vulnerability to stressors resulting from decreased physiological reserves in multiple systems characterized by nutritional decline, fatigue, decreased activity, and overall weakness leading to a higher risk of adverse health outcomes including mortality and a lower probability to receive preventive treatments (Fried et al. 2004). As the concept has to be distinguished from co-morbidity (Fried et al. 2004), frailty is very difficult to measure based on electronic healthcare databases (Glynn et al. 2001).

5.2.2. Outcome Misclassification

An important prerequisite to conduct high quality PSS is the accurate assessment of the study outcomes. For the evaluation of the accuracy of an outcome based on administrative claims data, the first important aspect is to know how the data were created and processed and for which purposes they were collected (Lanes et al. 2015). For

instance, hospital diagnoses are typically used for reimbursement based on DRGs. Therefore, hospitals have a high interest in correctly coding diseases, especially for severe cases (Schneeweiss and Avorn 2005). In addition, the quality of hospital diagnoses is often routinely checked by health insurance providers as they pay for hospital care of their insurance members. In contrast, ambulatory diagnoses typically do not affect reimbursement and are therefore subject to less quality checks (Lanes et al. 2015). If the diagnostic certainty is not adequately captured by physicians in the outpatient setting, bias can, for instance, be introduced by misclassification of suspected or ruled out diagnoses as ascertained diagnosis (Schneeweiss and Avorn 2005).

In general, the possibility of identifying an outcome based on administrative claims data depends on the acuity and onset of the respective event. For instance, stroke or myocardial infarction typically occur rapidly and require immediate treatment in the hospital setting, i.e., the date of onset can be determined exactly in most cases. Thus, these acute and severe outcomes are well captured based on administrative claims data as shown in validation studies from the Unites States (Brouwer et al. 2015, Lakshminarayan et al. 2014). In contrast, chronic disease outcomes with potentially long latency periods not necessarily requiring hospitalization, e.g., heart failure or type 2 diabetes, may be more difficult to investigate. Even if additional information on specific treatments (e.g., antidiabetic drugs) is used to define the onset date, the chronology of exposure and event may not be correctly assessed and protopathic bias might occur (Lanes et al. 2015).

In addition, outcome misclassification in PSS can especially bias the effect estimate if the specificity of the recorded diagnoses is low. In case of 100% specificity of the diagnosis, relative risk estimates will remain nearly unbiased in the case of non-differential misclassification even with a low sensitivity in most cases (Rothman et al. 2008b, Schneeweiss and Avorn 2005). A major concern is therefore differential misclassification, e.g., if an outcome is expected under treatment with a specific drug and these patients are more closely monitored compared to the unexposed group (detection bias).

5.2.3. Exposure Misclassification

Besides the adequate assessment of outcome events, it is crucial in PSS based on electronic healthcare databases to ascertain the exposure of drugs appropriately.

For instance, many ADRs occur during the first days of drug treatment. In this case, it is essential to have valid exposure information to prevent misclassification of the exposure status. This applies for both cohort studies and nested case-control studies in which drug

exposure has to be captured on a daily basis to define exposure risk windows (Garbe and Suissa 2014). Administrative claims data are often considered the gold standard for the assessment of drug exposure information, since prescriptions usually have to be filled in the pharmacy to get reimbursed by health insurance providers (Schneeweiss and Avorn 2005); however, some challenges to define the exact exposure status for patients remain. Misclassification may occur in administrative claims data because the exact duration and dose of drug treatment are typically not documented. While medical record databases often contain information on the prescribed daily dose or the amount of tablets etc. to be taken, pharmacy dispensing information usually does not provide this information, and the duration of therapy and dose have to be estimated based on assumptions on the regular drug intake of the patient, e.g., based on the prescribed amount of DDDs or by assigning a fixed supply for each prescription (Garbe and Suissa 2014, Schneeweiss and Avorn 2005). Under these assumptions, misclassification can mainly occur in two different ways. First, patients can be falsely considered as unexposed if they take the respective medication at a lower dosage or "as-needed" which might be a specific problem in older patients (see section 3.3.2). Second, patients might be misclassified as exposed if they filled the prescription, but decided to stop the drug for reasons like side effects or perceived missing efficacy (Schneeweiss and Avorn 2005). Furthermore, overlapping prescriptions may introduce exposure misclassification because it is not clear whether the patient used the drug at a higher dosage or the prescription was filled and the drug used later. Depending on the medication, it has to be decided if the supply of the previous prescriptions is concatenated or the supply of the following prescriptions is added to the previous one (stockpiling) (Greevy et al. 2011).

5.2.4. Time-Related Biases

The most important time-related biases in pharmacoepidemiological studies are immortal time and time-window bias. Immortal time bias typically occurs if a time-fixed analysis is used in cohort studies to emulate an intention-to-treat approach of an RCT (Suissa 2007). In this case, immortal time refers to a time period during follow-up, in which the outcome under study could not have occurred due to the definition of exposure assessment. In many studies, this time period refers to cohort time prior to initiating a specific drug treatment during which the patient had to stay event free until being classified as exposed. If this period is incorrectly considered as exposed (information bias) or alternatively disregarded in the analysis (selection bias), this leads to immortal time bias (Levesque et al. 2010, Suissa 2008b).

An example of immortal time bias is shown in Figure 1 adapted from a study by Lai et al. (2012) which found a 39% to 45% risk reduction for users of different antidiabetic drugs compared to non-users. In this study, metformin users were compared to non-users and cohort entry was defined as the first diagnosis of type 2 diabetes. Exposure status was defined as time-fixed based on the whole follow-up., i.e., a patient was considered as exposed if he or she had a prescription of metformin during follow-up. While exposure was defined correctly for all non-exposed patients and metformin users initiating treatment at the time of diagnosis, subjects starting metformin after the diagnosis (e.g., after six months of non-pharmacological treatment) were erroneously considered as exposed during the whole follow-up. For these patients, the time from diagnosis to first prescription of metformin is immortal time because the event lung cancer could not have occurred during this time period. And if it did, patients would have been considered unexposed. Using this study design, metformin users have a guaranteed event free follow-up time and, as a result of the exposure misclassification, the event rate of lung cancer is underestimated in the exposed group and additionally overestimated in the non-exposed group.

Depending on the magnitude of the misclassified immortal time, this bias may have an enormous impact on the rate ratio and interpretation of study results (Suissa 2008b). In previous studies, it often led to an overestimation with regard to the beneficial effect of drugs (Levesque et al. 2010, Schmedt et al. 2014, Schmedt and Garbe 2013b, Suissa 2007, Suissa 2008b); however, it can likewise affect cohort studies investigating the safety of drugs (Daniel et al. 2015) and hide possible drug risks.

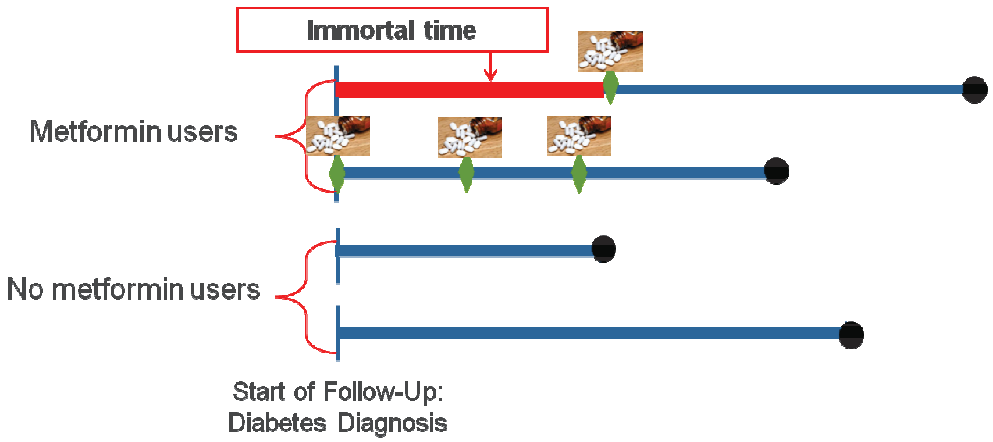

Figure 1 Illustration of Immortal Time Bias in a Study on Metformin and Lung Cancer (Lai et. al 2012)

Another time-related bias leading to substantial misinterpretation of pharmacoepidemiological case-control studies is the so-called time-window bias (Schmedt et al. 2014, Suissa and Azoulay 2012, Suissa et al. 2011). It typically occurs in (nested) case-control studies as a result of time-independent control selection and exposure assessment (Suissa et al. 2011). Suissa et al. (2011) illustrated the time-window bias based on a case-control study by Khurana et al. (2007) which found a 45% risk reduction for lung cancer among statin users compared to non-users. In this study with a maximal observation period of 67 months, exposure was defined as at least one prescription of a statin prior to lung cancer diagnosis for cases and during the whole observation period for controls. As illustrated in Figure 2, lung cancer cases had a lower probability of receiving statin treatment compared to controls due to a shorter observation period. This produced an overrepresentation of unexposed cases and again led to an overestimation of the beneficial effect of the drug under study (Suissa and Azoulay 2012, Suissa et al. 2011). In PSS, this phenomenon could result in an attenuation of risk estimates and conceal a possible risk of drug use.

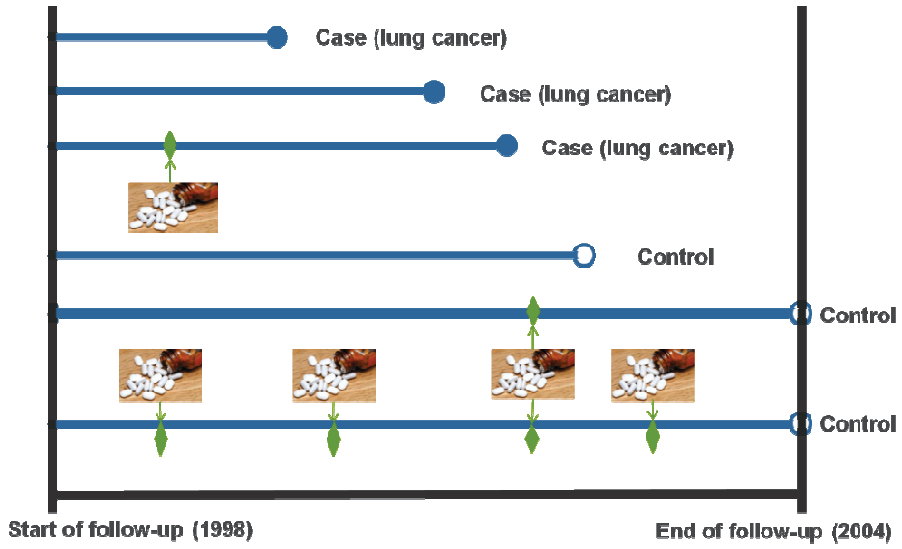

Figure 2 Illustration of Time-Window Bias in the Study by Khurana et al. (2007) (modified from Suissa et al. 2011)

6. Discussion

In this section, the methodological challenges in PSS described before are discussed against the background of my research with a focus on older patients. As far as possible, opportunities to address these challenges are presented based on examples in our

studies, and their advantages and disadvantages are pointed out. In addition, the first part of this section contains a discussion of the role of disproportionality analyses based on spontaneous reports for drug safety evaluations after marketing approval.

6.1. Disproportionality Analyses Based on Spontaneous Reporting Systems – A First Step

Disproportionality analyses based on spontaneous reporting systems have to be considered an important tool to evaluate drug safety after marketing approval. The main advantage of disproportionality analyses and spontaneous reporting systems is that case reports are immediately available for assessment upon submission. In many electronic healthcare databases, data delivery is often delayed, i.e., drug safety evaluations cannot be conducted until 1.5 to 2 years after marketing approval. However, at this time, many patients may already have experienced serious ADRs. In addition, disproportionality analyses are valuable to study rare ADRs mostly related to drug use. For instance, we studied safety signals of PML in users of selected immunosuppressants (Schmedt et al. 2012). PML was suitable for a disproportionality analysis because it is very rare and most cases - except in patients with HIV/AIDS - are related to immunosuppressive drug treatment (Weber 2008). Besides, as PML is uncommon it would probably not be possible to conduct a PSS with sufficient statistical power based on large electronic healthcare databases such as GePaRD.

On the other hand, disproportionality analyses based on spontaneous reporting systems have inherent limitations. Therefore, they can only serve as a first step to generate hypotheses about possible ADRs and cannot replace PSS on more accurate data (Garbe and Suissa 2014).

The first limitation of disproportionality analyses refers to the data quality of case reports in spontaneous reporting systems. Based on our studies (Andersohn et al. 2010, Schmedt et al. 2012), important information such as on age and sex is often missing. A recent study showed that case reports submitted to the U.S. Food and Drug Administration by drug manufacturers in 2014 contained complete information on age, sex, event date, and the medical term of the ADR in only 41%. Age was missing in 38% (Moore et al. 2016). In this case, it is difficult to investigate age effects and disproportionate reporting in specific age groups such as older patients. In addition, follow-up reports of the same suspected ADRs are often not adequately recorded in the spontaneous reporting system. In our study investigating the association between immunosuppressants and PML (Schmedt et al.

2012), we excluded 84 duplicates out of 719 reports of PML from the analysis after manual review. For several case reports, more than five duplicates were identified. If these cases had not been excluded, this would have led to substantial bias in the calculation of the ROR in our study.

The second important limitation of disproportionality analyses is selective reporting of specific drug-ADR combinations (reporting bias) (Moore et al. 2003). The correct interpretation of the ROR and PRR actually requires that the underreporting of ADRs is similar for the drug of interest and all other drugs in the comparison group, and that it is stable over time. However, it has been shown that increased medical and media attention for a specific drug-ADR combination can lead to selective reporting (Griffin 1986) and that public safety warnings can stimulate the reporting of ADRs for a specific drug (notoriety bias) (Pariente et al. 2007). In general, reporting rates typically change over time. The Weber effect was first described in the 1980s and replicated in later studies (Hartnell and Wilson 2004). It depicts the phenomenon that ADR reporting continuously rises within the first 18 to 24 months after marketing approval and then steadily decreases despite a growing market share (Weber 1984). Although other analyses did not reveal the typical pattern for the Weber effect, reporting of ADRs in those studies was also not stable over time for most drugs (Hoffman et al. 2014, Wallenstein and Fife 2001). As a result, selective reporting can lead to false-positive signals of specific ADR-drug combinations, but it can also mask signals of the same ADR for other drugs (competition bias) (Pariente et al. 2012a). For example, in our study on PML (Schmedt et al. 2012) it is likely that the ROR was overestimated for some drugs that had gained particular media attention due to a possibly higher risk, e.g., natalizumab which had been withdrawn for 16 months due to case reports of PML. On the other hand, higher reporting of PML cases associated with natalizumab could have obscured the detection of signals for other immunosuppressants

The third limitation of disproportionality analyses is the fact that they are less suitable for specific safety outcomes; this particularly applies to older patients. ADRs occurring a long time after drug treatment (“type D reactions”) may often not be reported, since they are not considered related to a drug intake from many years before. The same applies to ADRs with a high background incidence in the population. For example, possible life-threatening ADRs such as ischemic stroke or myocardial infarction are difficult to analyze because they frequently occur among older patients and may thus not be considered possible ADRs. In this context, it has to be assumed that many serious ADRs are underreported in the subgroup of older patients.

6.2. Confounding in Drug Safety Studies – Big Problems, Promising Approaches

Confounding bias can be addressed in several ways in PSS. For measured confounders, restriction or matching can be applied at the stage of the study design and standardization, stratification or adjustment in multivariable regression can be used in the analysis. For unmeasured confounding, case-only designs and advanced methods as the high-dimensional propensity score (HDPS) can be used. The application and usefulness of these methods are discussed in the following section based on my research with a focus on older patients. Further approaches to address unmeasured confounding in PSS are, for instance, the application of instrumental variables or two-stage sampling (Schneeweiss 2006); however these methods are not discussed here.

6.2.1. Drug Utilization Studies to Detect Possible Confounding and Channeling

The first step to mitigate confounding in PSS is to carefully define and assess possible confounders available in the electronic healthcare database, i.e., all diseases, drug treatments etc. that may be independently associated with the study endpoint.

Before we compared the risk of death between different antipsychotics and risperidone in older patients (Schmedt et al. 2016b), we conducted a detailed drug utilization study to assess the baseline prevalence of important possible confounders and indications (Schmedt et al. 2016a). This was particularly important in the analysis of older patients, in whom antipsychotics are used for various indications, e.g., dementia, sleeping disorders, depression, etc. For instance, we found that users of haloperidol more frequently had a diagnosis of metastatic solid tumors than patients treated with other antipsychotics (Schmedt et al. 2016a), since it is also used as an anti-emetic and sedative in palliative care (McLean et al. 2013). Consequently, a strong association with haloperidol and the outcome death was expected. This would have resulted in strong confounding and an overestimation of the increased risk of death for haloperidol compared to other antipsychotics. To avoid this bias, we excluded patients with any cancer diagnosis, antineoplastic treatment or other records indicative of palliative care prior to cohort entry from the safety analysis (Schmedt et al. 2016b).

Another finding to highlight the importance of detailed drug utilization studies as a prerequisite for a comparative safety study is the example of the long acting antipsychotic fluspirilene. It is currently only indicated for the treatment of schizophrenia, but older

studies suggested that fluspirilene may also be used as an alternative to or add-on for benzodiazepine tranquilizers in older patients (Schmidt 1989). Our drug utilization study revealed that patients treated with fluspirilene had less diagnosed somatic and psychiatric co-morbidity and that the drug was more often used in mild psychiatric conditions compared to other antipsychotics (Schmedt et al. 2016a). We finally refrained from analyzing fluspirilene in the safety analysis because of suspected biased results through strong unmeasured confounding (Schmedt et al. 2016b).

6.2.2. High-Dimensional Propensity Score

The HDPS has been proposed by Schneeweiss et al. (2009) as an extension of the traditional propensity score (PS) method. The PS is calculated with logistic regression modeling the probability of receiving a drug compared to another drug or no treatment depending on a pre-defined set of confounding variables (Sturmer et al. 2014). In contrast, the HDPS is additionally calculated incorporating a set of many empirically selected confounding variables (often up to 500). The calculation of the HDPS basically includes the following five steps: (i) identification of the data dimensions, e.g., inpatient and outpatient diagnoses, inpatient operations and procedures, outpatient diagnostic and therapeutic procedures, and outpatient dispensations, (ii) selection of candidate codes based on their prevalence, (iii) construction of covariates based on the recurrence of the respective codes, (iv) prioritization of the covariates based on the estimated strength of confounding according to the Bross formula within each data dimension, and (v) selection of the covariates with the strongest confounding and calculation of the HDPS (Schneeweiss et al. 2009). As with traditional PS analyses, a comprehensive diagnostic of the HDPS analysis must be performed which may involve the use of HDPS in several ways, e.g., stratification, adjustment, inverse probability of treatment weighting, and matching in combination with or without different trimming approaches (Sturmer et al. 2014). The correct implementation and the appropriateness of the HDPS analysis always depend on the research question and the population under study.

In our comparative safety study on antipsychotics and death in older patients, we used the HDPS as matching parameter and adjusted for HDPS-deciles in the Cox-proportional hazards model to account for possible unmeasured confounding in the conventional multivariable regression analysis. Although no substantial differences were observed, the hazard ratio for some comparisons in the HDPS analysis tended towards a null effect in the HDPS adjusted and matched analysis which may indicate remaining unmeasured or residual confounding (Schmedt et al. 2016b). The HDPS has to be considered particularly

useful in our study among the heterogeneous group of older patients in whom the drugs under study were used for various indications and therefore strong confounding must be assumed. In this case, the HDPS approach might facilitate the empirical selection of proxies for unmeasured confounders, e.g., frailty, for which there is no established score or index based on administrative claims data (Kim and Schneeweiss 2014). In future studies, the performance of the HDPS in older patients would benefit if data on nursing care and remedies could be used as additional data dimension. Probably, these include important additional proxies on possible unmeasured confounders. Unfortunately, these data were not available in our study (Schmedt et al. 2016b).

6.2.3. Application of case-only designs

In general, case-only designs can be considered valuable tools to investigate possible acute ADRs for transient drug exposures because of their ability to inherently control for time-invariant confounders that cannot be assessed in electronic healthcare databases. This is especially advantageous in PSS among older patients whose individual constitution and response to drugs may differ substantially (see section 3.3). In addition, many drugs are not used over longer periods in older patients. For instance, antipsychotics are often used “as-needed” in patients with dementia and for a short duration (Schmedt et al. 2016a). For acute ADRs assumed to occur under current treatment of antipsychotics, e.g., ischemic stroke, myocardial infarction, ventricular arrhythmia (Trifiro et al. 2014), a case-only design can therefore be considered an equal or even preferable study design compared to a classic cohort or nested-case control study. Consequently, the risk of myocardial infarction and stroke associated with antipsychotic use was recently evaluated based on case-only designs in the whole population (Brauer et al. 2015, Lin et al. 2014) and more specifically among older patients (Pariente et al. 2012b, Shin et al. 2013). Retrospectively, a case-only design would also have been an interesting additional approach to study the risk of VTE under antipsychotic treatment in patients with dementia complementing our nested-case control study (Schmedt and Garbe 2013a).

On the other hand, case-only designs are often not suitable for relevant drug safety questions. For instance, it is nearly impossible to study long-term safety outcomes such as cancer, and it is also not possible to evaluate safety concerns of drugs continuously used over long time periods. In addition, it is only possible to control for time-invariant unmeasured confounders; however, many possible confounders vary with time in the same individual in clinical practice (Hallas and Pottegard 2014). For instance, in our self-

controlled case series analysis on stimulant use and the risk of injuries in children with attention deficit/hyperactivity disorder, a major concern was that the severity of symptoms may vary over time and trigger prescriptions of methylphenidate or atomoxetine (Mikolajczyk et al. 2015). We did not consider this as a problem because such confounding by disease severity would have diluted a possible beneficial effect that was still observed; however, in a PSS this could lead to an overestimation of drug risks. Another problem that may particularly limit the applicability of case-only studies is exposure misclassification in older patients. Although this also applies to all other study designs, the correct assessment of the exposure risk window is particularly important in case-only studies in which usually transient drug effects are measured (Hallas and Pottegard 2014). Unfortunately, the assessment of transient drug exposure is extremely challenging if the duration of use has to be estimated, and exposure misclassification is more likely in older patients due to larger variations of doses, lower adherence, up-titration, and tapering etc. For a more detailed discussion of exposure misclassification, see section 6.4.

6.3. Avoiding Outcome Misclassification – We Can Do Better!

Outcome misclassification is often neglected in PSS based on administrative claims data, although it can be addressed in several ways.

To assess the accuracy of a case-algorithm, it should ideally be validated against a well-suited gold standard, e.g., medical records or physician questionnaires. If this is not possible, alternative strategies for validation may be used as, for instance, the comparison of event rates against reliable data from other sources (Lanes et al. 2015, Schneeweiss and Avorn 2005). As an example, we compared the age- and sex-standardized mortality rates in GePaRD to those from the German Federal Statistical Office. Although we found discrepancies that may be related to differences with regard to the socioeconomic status of the study populations, the results led to the conclusion that death is reliably recorded in GePaRD (Ohlmeier et al. 2015). In a similar approach, we compared the event rates of several study outcomes, e.g., stroke, acute myocardial infarction, pancreatic cancer, etc., in the EU-funded multi-database study “Safety Evaluation of Adverse Reactions in Diabetes” (SAFEGUARD) across different data sources and with event rates reported in the literature (Schmedt et al. 2013). However, there are two major limitations of this approach: (i) even if the event rates are similar across data sources, it is possible that false-positive events mask false-negatives, and (ii) the results of this approach cannot be transferred to a specific study population, e.g., a cohort of older patients with dementia.

If external validation cannot be performed at all, other strategies to enhance the specificity of a case-algorithm should be considered. In GePaRD, the main hospital discharge diagnosis is often used for acute and severe events and diagnoses from the outpatient setting are usually not taken into account. We also used this approach in our nested-case control study on the risk of VTE in older users of antipsychotics (Schmedt and Garbe 2013a). In most cases, the specificity of the case-algorithm may be further increased by using additional information on drug treatment of the outcome or obligatory diagnostic tests (The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance 2015). For instance, a case-algorithm for VTE could have included additional information on the use of anticoagulants or possible records for an ultrasound. Although this is rarely done, a detailed review of the patient's profile based on the available information in administrative claims data can also be performed. Two experienced reviewers can evaluate all possible cases, and implausible events can be excluded to increase the specificity of the case algorithm. A manual review might be considered particularly important in older patients because there are usually more differential diagnoses, and less diagnostic tests may be performed which can decrease the specificity of using diagnoses alone. In the case of many outcomes for manual review, diagnostic tests or information on drug treatment can also be implemented in automated case-algorithms.

In order to detect possible differential misclassification of the outcome case-algorithm (detection bias), it can be useful to count the number of performed diagnostic tests for the outcome for all exposure groups. For instance, the number of liver tests can be counted if the outcome under study is acute liver injury. This approach can be particularly important if the occurrence of a specific ADR is already suspected for a specific drug, and treated patients are more closely monitored.

6.4. Assessing Drug Exposure in Administrative Claims Data – A "Double-Edged Sword"

On the one hand, the assessment of drug exposure in older patients based on administrative claims data has many advantages and is often considered the gold standard in PSS as recall bias can be ruled out and prescribing information is also available for subpopulations that are difficult to recruit for field studies, e.g., older patients suffering from severe disease such as dementia (Schneeweiss and Avorn 2005) (see also section 3.2.2). For instance, in GePaRD, more detailed information is available via linkage

to a central pharmaceutical reference database, e.g., on the route of administration and the strength of the drug, which can provide a comprehensive picture of the respective drug exposure in older patients.

On the other hand, exposure misclassification can occur in several ways because the duration of drug treatment has to be estimated because the exact prescribed dose is not available. For chronic treatment such as antidiabetic drugs or antihypertensives, the problem of misclassification is of less relevance, since continuous refills at least indicate regular treatment and gaps between prescriptions may be combined to continuous treatment episodes by adding pre-defined fractions to the estimated drug supply (e.g., seven days or 50% of the prescribed DDD) (Gardarsdottir et al. 2010, Schneeweiss and Avorn 2005). However, the correct assignment of the exposure status for drugs assumed to be rather used “as-needed” or for very short time periods is more challenging (Schneeweiss and Avorn 2005) because the actual duration of use between two prescriptions or for a single prescription remains unknown. In older patients, “as-needed” treatment frequently occurs, e.g., in treatment with nonsteroidal anti-inflammatory drugs or with antipsychotics in dementia. In our studies on antipsychotics (Schmedt et al. 2016a, Schmedt et al. 2016b), we added 150% of the prescribed DDD for each antipsychotic dispensation and then created episodes of continuous treatment as displayed in Figure 3.

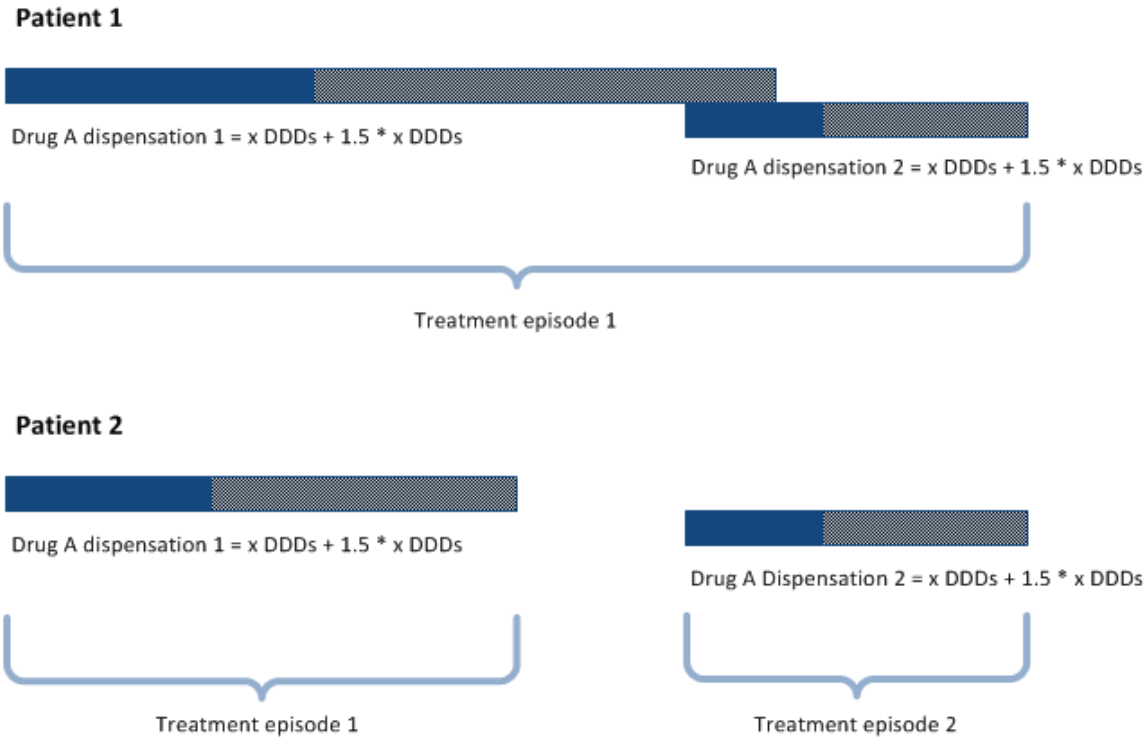

Figure 3 Construction of antipsychotic treatment episodes from Schmedt et al. 2016a

In the case of longer time periods between two prescriptions for the same patient (e.g., patient 2), it is unclear if the patient continuously used the drug at a low dose or sporadically at a higher dose. In this case, administrative claims data are still useful; however, an exposure misclassification cannot be ruled out and sensitivity analyses are crucial to assess the consistency of the results under different assumptions. This particularly applies to studies in which the timing of exposure has to be determined exactly, e.g., in case-only studies (see section 5.1.4). In our study, we performed sensitivity analyses adding 0%, 50%, 100%, 200%, and 300% of the prescribed DDD to assess the consistency of the treatment duration under different assumptions (Schmedt et al. 2016a).

Besides, exposure misclassification may occur in administrative claims data if drug exposure in the hospital setting is not documented as is the case in GePaRD. If this misclassification is differential between exposure groups, this may lead to biased risk estimates, especially if the outcome under study is mortality ("immeasurable time bias") (Suissa 2008a). Of course, this form of exposure misclassification is most relevant in older patients with severe co-morbidity who are hospitalized more often and for longer time periods, e.g., patients with dementia. Although this bias cannot be prevented, its possible impact and direction can be estimated by comparing the amount of hospitalized person time between the exposure groups. We did this in our comparative safety study on antipsychotics and death and found no substantial differences between the exposure groups (Schmedt et al. 2016b). In this context, it also needs to be emphasized that nested case-control studies, in which exposure status is often defined based on the last drug prescription before the index date, are more prone to this form of exposure misclassification. This was one of the reasons why we chose a cohort study with a modified intention-to-treat design in our study on the comparative risk of death in older antipsychotic users (Schmedt et al. 2016b).

Finally, a general problem in PSS with regard to the exposure based on administrative claims data is the missing information on the prescribed dose. Again, for chronic treatments, the dose can be estimated based on the time between two prescriptions. In contrast, it cannot be determined if "as-needed" treatment is frequent. This was the case in our drug utilization study of antipsychotics in older patients in which 44% of new users had only one treatment episode, and the median treatment duration was only 20 days (data not shown in research articles). Therefore, we were not able to analyze dose effects in the following safety study (Schmedt et al. 2016b), although previous studies had

reported substantial dose effects regarding the risk of death under antipsychotic medication (Gerhard et al. 2014, Huybrechts et al. 2012).

6.5. Time-Related Biases – Unnecessary Problems!

Immortal time bias and time-window bias have been introduced as potential pitfalls that can affect cohort studies and case-control studies, respectively (see section 5.2.4). This is true for PSS among all age groups and is not restricted to older patients. Instead, it is important to emphasize that these forms of biases are always introduced by the choice of the study design and can easily be prevented (Suissa 2008b).

As described above, immortal time bias is usually introduced through an inappropriate time-fixed analysis in which unexposed person time is wrongly considered as exposed person time (information bias) or unexposed person time before drug exposure is deleted entirely from the analysis (selection bias) (see section 5.2.4). This problem can be solved in most circumstances by using a time-dependent analysis (e.g., in the Cox proportional hazard model) (Suissa 2008b). In addition, it is important that treatment groups are not assigned hierarchically because exposure periods of drugs at a lower hierarchy level would thus be disregarded in the analysis. In the same manner, patients should not be excluded based on drug exposure assessed during follow-up (Levesque et al. 2010).

Time-window bias is introduced in nested-case control studies if the length of the exposure time window is systematically different for cases and controls. This form of bias can be avoided by using incidence density sampling (Suissa et al. 2011). For instance, in our study on VTE in older users of antipsychotics with dementia, we used risk-set sampling to select controls and defined the exposure status based on antipsychotic prescriptions in the 365 days before the index date of the case. Since all persons in the cohort had a minimum look-back period of 365 days before cohort entry, the exposure time window of all cases and controls had the same length (Schmedt and Garbe 2013a). In studies without a fixed exposure time window (e.g., if exposure is defined as “ever use” of a drug), additional matching by duration of follow-up or the date of cohort entry between cases and controls can be performed to prevent time-window bias.

7. Conclusions and Future Perspectives

The post-marketing evaluation of drug safety in older patients including the prevention of ADRs is of enormous public health interest. In light of the ageing society and the corresponding increase of drug use, its importance will increase even more in the future.

In this context, this thesis demonstrated that spontaneous reporting systems and PSS based on electronic healthcare databases are essential to detect ADRs after marketing approval and to overcome limitations of RCTs. However, it also emphasized important problems with regard to the quality of spontaneous reports and electronic healthcare databases as well as methodological pitfalls of PSS. Since some of these are particularly relevant for PSS in older patients, these challenges have to be addressed in future research activities.

More specifically, this thesis demonstrated that disproportionality analyses based on spontaneous reports are useful to analyze rare and specific ADRs such as PML; however, the method is generally hampered by the poor quality of the data. In addition, disproportionality analyses based on spontaneous reports are less suitable for the detection of ADRs in older patients if the background incidence of the ADR in the population is high. In this context, it has recently been shown that data mining approaches based on electronic healthcare databases have a greater potential to detect signals of more common ADRs (Patadia et al. 2015). As another example, sequential matched cohort studies have been applied as semi-automated active safety monitoring of newly marketed drugs based on electronic healthcare databases in the United States (Gagne et al. 2012). In addition, different approaches for signal detection have also been applied based on GePaRD (Suling and Pigeot 2012). Given the higher quality of data compared to spontaneous reports and the availability of established data sources like GePaRD, active semi-automated safety monitoring of new drugs based on electronic healthcare databases should be more often considered and further developed in Germany in the future. Of course, as a prerequisite, the data from SHIs would have to be delivered in a timelier manner to detect ADRs soon after marketing approval.

Moreover, this thesis revealed that confounding bias is one of the most important potential pitfalls in non-randomized drug safety research. Although PSS based on electronic healthcare databases can be considered the preferred alternative to study the safety of drugs in older people, confounding was identified as a particular concern in this very heterogeneous and vulnerable subpopulation (e.g., confounding by frailty). In this context, this thesis showed that drug utilization studies are crucial to detect possible drug channeling and to explore the potential for confounding by indication prior to the safety study itself. Furthermore, the HDPS method was applied in a study on the comparative risk of death in older antipsychotic users and discussed as a new promising approach to address unmeasured confounding based on empirical selection of potential proxies for confounders not captured in electronic healthcare databases (e.g., frailty). In addition,

case-only designs have been introduced as an alternative study design to overcome unmeasured confounding in the case of transient drug treatments, e.g., with antipsychotics in older patients. Beyond these approaches, the increasing digitalization may offer additional opportunities in the future if confounder information available in electronic healthcare databases can be enriched through linkage with other data sources. These could, for instance, include data from large field studies or information on physical activity from electronic tracking devices. In particular for Germany, the recent passage of the “e-health” legislation and the introduction of the electronic health card could imply further alternatives since more detailed clinical information such as laboratory test results, more detailed information on drug treatment and images from x-rays, computed tomography or magnetic resonance imaging will become available as digital content. Of course, data protection concerns will be a tremendous challenge to be solved in this context. If issues related to confounding impede PSS based on electronic healthcare databases alone, large simple trials (Lesko and Mitchell 2012) and randomized database studies (van Staa et al. 2014, van Staa et al. 2012) constitute an auspicious opportunity in combining randomization and facilitated follow-up through electronic healthcare databases.

In addition, this thesis focused on the correct assessment of safety outcomes in administrative claims data which is often neglected. It was highlighted that a high specificity of the case-algorithm of the outcome is essential in comparative drug safety studies. If a validation study with an external gold standard is not possible, a comparison of event rates to external data sources has been introduced as an opportunity to check the plausibility of a case-algorithm. Further, a detailed review of the available information (diagnoses, procedures, medications, etc.) was proposed as a useful approach to enhance the specificity of the case-algorithm based on administrative claims data.

Furthermore, this thesis showed that the correct assessment of exposure based on pharmacy dispensing data in electronic healthcare databases can be critical in specific situations. This is particularly true for studies in the heterogeneous population of older patients in whom the prescribed dose may vary substantially and sporadic “as-needed” treatment is frequent. Thus, drug utilization should be investigated carefully before a safety study is planned and conducted. If exposure misclassification is likely, it should be evaluated whether the study can be performed and whether sensitivity analyses have to be conducted to test the consistency of the results under different assumptions of drug intake (i.e., frequency and dose). In the future, it will be crucial to perform external validation studies, for instance, with medical records to estimate the extent of exposure

misclassification and gain more detailed information on the prescribed dose if this information is not available in administrative claims databases.

Another important topic in PSS among older patients is the effect of potentially interacting drugs. This thesis showed that polypharmacy and drug-drug interactions are frequent in older patients and may lead to a higher risk of ADRs. Although outcomes of drug-drug interactions itself are increasingly being studied, possible effect modification by potentially interacting drugs is rarely investigated. In future PSS in older patients, this should routinely be taken into account even though these drugs are often not independently associated with the outcome.

Finally, time-related biases have been pointed out as an important threat to the credibility of drug safety studies based on electronic healthcare databases. As outlined in this thesis, these forms of biases can be avoided by using an appropriate study design.

In the near future, the need for PSS after marketing approval will substantially grow, since post-authorization safety studies are commonly required in risk management plans which have to be submitted at the time of application for a marketing authorization. Because individual electronic healthcare databases are often too small, multidatabase studies such as in the SAFEGUARD project and large research networks, e.g., the Sentinel Initiative of the U.S. Food and Drug Administration (U.S. Food and Drug Administration 2010) or the Canadian Network for Observational Studies of Drug Effects (Suissa et al. 2012) will become more important to answer research questions with regard to the safety of newly marketed drugs. Many of these studies will analyze the safety of drugs in older patients, whose proportion in the general population will continuously increase and who will use the majority of drugs in the following decades. In this context, this thesis illustrated opportunities to assess the safety of drugs after marketing approval essentially contributing to the understanding of potential pitfalls to be addressed to enhance the quality of future PSS.

8. References

Almenoff, J. S., Pattishall, E. N., Gibbs, T. G., DuMouchel, W., Evans, S. J. and Yuen, N. (2007) Novel statistical tools for monitoring the safety of marketed drugs. *Clinical Pharmacology and Therapeutics, 82*(2), pp. 157-66.

Amann, U., Schmedt, N. and Garbe, E. (2012) Prescribing of potentially inappropriate medications for the elderly: an analysis based on the PRISCUS list. *Deutsches Arzteblatt International, 109*(5), pp. 69-75.

American Geriatrics Society Beers Criteria Update Expert Panel (2012) American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *Journal of the American Geriatrics Society, 60*(4), pp. 616-31.

American Geriatrics Society Beers Criteria Update Expert Panel (2015) American Geriatrics Society 2015 Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults. *Journal of the American Geriatrics Society, 63*(11), pp. 2227-46.

Andersohn, F. (2011) Identifizierung unerwünschter Arzneimittelwirkungen nach Marktzulassung. Postdoctoral thesis [Habilitationsschrift], Freie Universität Berlin. Available at: http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000022410?lang=en. Accessed: 11.03.2016

Andersohn, F., Bronder, E., Klimpel, A. and Garbe, E. (2004) Proportion of drug-related serious rare blood dyscrasias: estimates from the Berlin Case-Control Surveillance Study. *American Journal of Hematology, 77*(3), pp. 316-8.

Andersohn, F. and Garbe, E. (2008) Pharmacoepidemiological research with large health databases. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 51*(10), pp. 1135-1144.

Andersohn, F., Schmedt, N., Weinmann, S., Willich, S. N. and Garbe, E. (2010) Priapism associated with antipsychotics: role of alpha1 adrenoceptor affinity. *Journal of Clinical Psychopharmacology, 30*(1), pp. 68-71.

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S. and Guthrie, B. (2012) Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet, 380*(9836), pp. 37-43.

Bate, A. and Evans, S. J. (2009) Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiology and Drug Safety, 18*(6), pp. 427-36.

Bavry, A. A., Thomas, F., Allison, M., Johnson, K. C., Howard, B. V., Hlatky, M., Manson, J. E. and Limacher, M. C. (2014) Nonsteroidal anti-inflammatory drugs and cardiovascular outcomes in women: results from the women's health initiative. *Circulation: Cardiovascular Quality and Outcomes, 7*(4), pp. 603-10.

Beers, M. H. (1997) Explicit criteria for determining potentially inappropriate medication use by the elderly: an update. *Archives of Internal Medicine, 157*(14), pp. 1531-1536.

Brauer, R., Smeeth, L., Anaya-Izquierdo, K., Timmis, A., Denaxas, S. C., Farrington, C. P., Whitaker, H., Hemingway, H. and Douglas, I. (2015) Antipsychotic drugs and risks of myocardial infarction: a self-controlled case series study. *European Heart Journal, 36*(16), pp. 984-92.

Brouwer, E. S., Napravnik, S., Eron, J. J., Jr., Simpson, R. J., Jr., Brookhart, M. A., Stalzer, B., Vinikoor, M., Floris-Moore, M. and Sturmer, T. (2015) Validation of Medicaid Claims-based Diagnosis of Myocardial Infarction Using an HIV Clinical Cohort. *Medical Care, 53*(6), pp. e41-8.

Coca, V. and Nink, K. (2008) Arzneimittelverordnungen nach Alter und Geschlecht. In Schwabe, U. and Paffrath, D., (eds.) *Arzneiverordnungsreport 2008*, Heidelberg: Springer Medizin Verlag. pp. 909-923.

Cockcroft, D. W. and Gault, M. H. (1976) Prediction of creatinine clearance from serum creatinine. *Nephron, 16*(1), pp. 31-41.

Daniel, S., Koren, G., Lunenfeld, E. and Levy, A. (2015) Immortal time bias in drug safety cohort studies: spontaneous abortion following nonsteroidal antiinflammatory drug exposure. *American Journal of Obstetrics and Gynecology, 212*(3), pp. 307 e1-6.

Desai, P., Chlebowski, R., Cauley, J. A., Manson, J. E., Wu, C., Martin, L. W., Jay, A., Bock, C., Cote, M., Petrucelli, N., Rosenberg, C. A., Peters, U., Agalliu, I., Budrys, N., Abdul-Hussein, M., Lane, D., Luo, J., Park, H. L., Thomas, F., Wactawski-Wende, J. and Simon, M. S. (2013) Prospective analysis of association between statin use and breast cancer risk in the women's health initiative. *Cancer Epidemiology, Biomarkers and Prevention, 22*(10), pp. 1868-76.

Edwards, I. R. and Aronson, J. K. (2000) Adverse drug reactions: definitions, diagnosis, and management. *The Lancet, 356*(9237), pp. 1255-1259.

Essebag, V., Genest, J., Jr., Suissa, S. and Pilote, L. (2003) The nested case-control study in cardiology. *American Heart Journal, 146*(4), pp. 581-90.

European Medicines Agency (2016) EudraVigilance. Available at: https://eudravigilance.ema.europa.eu/human/index.asp. Accessed: 11.03.2016

Evans, S. J., Waller, P. C. and Davis, S. (2001) Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiology and Drug Safety, 10*(6), pp. 483-6.

Farrington, C. P. (1995) Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics, 51*(1), pp. 228-35.

Federal Institute for Drugs and Medical Devices (2016) UAW-Publikationsplattform des Bundesinstituts für Arzneimittel und Medizinprodukte.
Available at: http://www2.bfarm.de/uaw/. Accessed: 11.03.2016

Federal Ministry of Health (2016) Zahlen und Fakten zur Krankenversicherung.

Available at: http://www.bmg.bund.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung.html. Accessed: 16.03.2016

Federal Statistical Office (2015) Bevölkerung Deutschlands bis 2060 - 13. koordinierte Bevölkerungsvorausberechnung. Wiesbaden. Available at: https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/VorausberechnungBevoelkerung/BevoelkerungDeutschland2060Presse5124204159004.pdf?__blob=publicationFile. Accessed: 11.03.2016

Fick, D. M., Cooper, J. W., Wade, W. E., Waller, J. L., Maclean, J. R. and Beers, M. H. (2003) Updating the Beers criteria for potentially inappropriate medication use in older adults: results of a US consensus panel of experts. *Archives of Internal Medicine, 163*(22), pp. 2716-2724.

Flockhart, D. (2011) Drug Interactions: Cytochrome P450 Drug Interaction Table. Indiana University School of Medicine. Available at: http://medicine.iupui.edu/clinpharm/ddis/table.aspx. Accessed: 11.03.2016

Franks, M. E., Macpherson, G. R. and Figg, W. D. (2004) Thalidomide. *The Lancet, 363*(9423), pp. 1802-11.

Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D. and Anderson, G. (2004) Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *Journals of Gerontology. Series A: Biological Sciences and Medical Sciences, 59*(3), pp. 255-63.

Gagne, J. J., Glynn, R. J., Rassen, J. A., Walker, A. M., Daniel, G. W., Sridhar, G. and Schneeweiss, S. (2012) Active safety monitoring of newly marketed medications in a distributed data network: application of a semi-automated monitoring system. *Clinical Pharmacology and Therapeutics, 92*(1), pp. 80-6.

Gallagher, P., Ryan, C., Byrne, S., Kennedy, J. and O'Mahony, D. (2008) STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment). Consensus validation. *International Journal of Clinical Pharmacology and Therapeutics, 46*(2), pp. 72-83.

Garbe, E. and Pigeot, I. (2015) Benefits of large healthcare databases for drug risk research. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 58*(8), pp. 829-37.

Garbe, E. and Suissa, S. (2014) Pharmacoepidemiology. In Ahrens, W. and Pigeot, I., (eds.) *Handbook of Epidemiology*, New York: Springer Science+Business Media. pp. 1875-1915.

Garbe, E., Suling, M., Kloss, S., Lindemann, C. and Schmid, U. (2011) Linkage of mother-baby pairs in the German Pharmacoepidemiological Research Database. *Pharmacoepidemiology and Drug Safety, 20*(3), pp. 258-264.

Gardarsdottir, H., Souverein, P. C., Egberts, T. C. and Heerdink, E. R. (2010) Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *Journal of Clinical Epidemiology, 63*(4), pp. 422-427.

Gerhard, T., Huybrechts, K., Olfson, M., Schneeweiss, S., Bobo, W. V., Doraiswamy, P. M., Devanand, D. P., Lucas, J. A., Huang, C., Malka, E. S., Levin, R. and Crystal, S. (2014) Comparative mortality risks of antipsychotic medications in community-dwelling older adults. *British Journal of Psychiatry, 205*(1), pp. 44-51.

Glynn, R. J., Knight, E. L., Levin, R. and Avorn, J. (2001) Paradoxical relations of drug treatment with mortality in older persons. *Epidemiology, 12*(6), pp. 682-9.

Greevy, R. A., Jr., Huizinga, M. M., Roumie, C. L., Grijalva, C. G., Murff, H., Liu, X. and Griffin, M. R. (2011) Comparisons of persistence and durability among three oral antidiabetic therapies using electronic prescription-fill data: the impact of adherence requirements and stockpiling. *Clinical Pharmacology and Therapeutics, 90*(6), pp. 813-9.

Griffin, J. P. (1986) Survey of the spontaneous adverse drug reaction reporting schemes in fifteen countries. *British Journal of Clinical Pharmacology, 22 Suppl 1*, pp. 83S-100S.

Hall, G. C. (2009) Validation of death and suicide recording on the THIN UK primary care database. *Pharmacoepidemiology and Drug Safety, 18*(2), pp. 120-131.

Hallas, J. and Pottegard, A. (2014) Use of self-controlled designs in pharmacoepidemiology. *Journal of Internal Medicine, 275*(6), pp. 581-9.

Hamilton, H., Gallagher, P., Ryan, C., Byrne, S. and O'Mahony, D. (2011) Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients. *Archives of Internal Medicine, 171*(11), pp. 1013-9.

Hartnell, N. R. and Wilson, J. P. (2004) Replication of the Weber effect using postmarketing adverse event reports voluntarily submitted to the United States Food and Drug Administration. *Pharmacotherapy, 24*(6), pp. 743-9.

Hedna, K., Hakkarainen, K. M., Gyllensten, H., Jonsson, A. K., Petzold, M. and Hagg, S. (2015) Potentially inappropriate prescribing and adverse drug reactions in the elderly: a population-based study. *European Journal of Clinical Pharmacology, 71*(12), pp. 1525-33.

Hennessy, S. (2006) Basic principles of clinical pharmacology relevant to pharmacoepidemiological studies. In Strom, B. L. and Kimmel, S. E., (eds.) *Textbook of Pharmacoepidemiology*, Chichester: Wiley-Blackwell. pp. 35-54.

Herland, K., Akselsen, J.-P., Skjønsberg, O. H. and Bjermer, L. (2005) How representative are clinical study patients with asthma or COPD for a larger “real life” population of patients with obstructive lung disease? *Respiratory Medicine, 99*(1), pp. 11-19.

Herrett, E., Gallagher, A. M., Bhaskaran, K., Forbes, H., Mathur, R., van Staa, T. and Smeeth, L. (2015) Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology, 44*(3), pp. 827-36.

Hoffman, K. B., Dimbil, M., Erdman, C. B., Tatonetti, N. P. and Overstreet, B. M. (2014) The Weber effect and the United States Food and Drug Administration's Adverse

Event Reporting System (FAERS): analysis of sixty-two drugs approved from 2006 to 2010. *Drug Safety, 37*(4), pp. 283-94.

Holt, S., Schmiedl, S. and Thurmann, P. A. (2010) Potentially inappropriate medications in the elderly: the PRISCUS list. *Deutsches Arzteblatt International, 107*(31-32), pp. 543-551.

Huybrechts, K. F., Gerhard, T., Crystal, S., Olfson, M., Avorn, J., Levin, R., Lucas, J. A. and Schneeweiss, S. (2012) Differential risk of death in older residents in nursing homes prescribed specific antipsychotic drugs: population based cohort study. *BMJ, 344*, pp. e977.

Johnell, K. and Klarin, I. (2007) The relationship between number of drugs and potential drug-drug interactions in the elderly. *Drug Safety, 30*(10), pp. 911-918.

Juurlink, D. N., Mamdani, M., Kopp, A., Laupacis, A. and Redelmeier, D. A. (2003) Drug-drug interactions among elderly patients hospitalized for drug toxicity. *JAMA, 289*(13), pp. 1652-1658.

Khurana, V., Bejjanki, H. R., Caldito, G. and Owens, M. W. (2007) Statins reduce the risk of lung cancer in humans: a large case-control study of US veterans. *Chest, 131*(5), pp. 1282-8.

Kim, D. H. and Schneeweiss, S. (2014) Measuring frailty using claims data for pharmacoepidemiologic studies of mortality in older adults: evidence and recommendations. *Pharmacoepidemiology and Drug Safety, 23*(9), pp. 891-901.

Lai, S. W., Liao, K. F., Chen, P. C., Tsai, P. Y., Hsieh, D. P. and Chen, C. C. (2012) Antidiabetes drugs correlate with decreased risk of lung cancer: a population-based observation in Taiwan. *Clinical Lung Cancer, 13*(2), pp. 143-8.

Lakshminarayan, K., Larson, J. C., Virnig, B., Fuller, C., Allen, N. B., Limacher, M., Winkelmayer, W. C., Safford, M. M. and Burwen, D. R. (2014) Comparison of Medicare claims versus physician adjudication for identifying stroke outcomes in the Women's Health Initiative. *Stroke, 45*(3), pp. 815-21.

Lanes, S., Brown, J. S., Haynes, K., Pollack, M. F. and Walker, A. M. (2015) Identifying health outcomes in healthcare databases. *Pharmacoepidemiology and Drug Safety, 24*(10), pp. 1009-16.

Laroche, M.-L., Charmes, J.-P. and Merle, L. (2007) Potentially inappropriate medications in the elderly: a French consensus panel list. *European Journal of Clinical Pharmacology, 63*(8), pp. 725-731.

Lesko, S. M. and Mitchell, A. A. (2012) The use of randomized controlled trials for pharmacoepidemiological studies. In Strom, B. L., Kimmel, S. E. and Hennessy, S., (eds.) *Pharmacoepidemiology*, Chichester: Wiley-Blackwell. pp. 640-654.

Levesque, L. E., Hanley, J. A., Kezouh, A. and Suissa, S. (2010) Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ, 340*, pp. b5087.

Lewis, M. A., Heinemann, L. A., Spitzer, W. O., MacRae, K. D. and Bruppacher, R. (1997) The use of oral contraceptives and the occurrence of acute myocardial infarction in young women. Results from the Transnational Study on Oral Contraceptives and the Health of Young Women. *Contraception, 56*(3), pp. 129-40.

Lin, S. T., Chen, C. C., Tsang, H. Y., Lee, C. S., Yang, P., Cheng, K. D., Li, D. J., Wang, C. J., Hsieh, Y. C. and Yang, W. C. (2014) Association between antipsychotic use and risk of acute myocardial infarction: a nationwide case-crossover study. *Circulation, 130*(3), pp. 235-43.

Linder, R., Schneider, U., Kothemann, M. and Verheyen, F. (2014) Physicians' prescription behavior of potentially inappropriate medications for elderly people: an analysis using the PRISCUS list based on TK routine data. *Deutsche Medizinische Wochenschrift, 139*(19), pp. 983-9.

Lüngen, M. and Lauterbach, K. W. (2000) Upcoding--a risk for the use of diagnosis-related groups. *Deutsche Medizinische Wochenschrift, 125*(28-29), pp. 852-856.

MacDonald, T. M., Pettitt, D., Lee, F. H. and Schwartz, J. S. (2003) Channelling of patients taking NSAIDs or cyclooxygenase-2-specific inhibitors and its effect on interpretation of outcomes. *Rheumatology, 42 Suppl 3*, pp. iii3-10.

Maclure, M. (1991) The case-crossover design: a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology, 133*(2), pp. 144-53.

Maio, G. (2001) On the history of the Contergan (thalidomide) catastrophe in the light of drug legislation. *Deutsche Medizinische Wochenschrift, 126*(42), pp. 1183-6.

Mallet, L., Spinewine, A. and Huang, A. (2007) The challenge of managing drug interactions in elderly people. *The Lancet, 370*(9582), pp. 185-191.

Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., Meinow, B. and Fratiglioni, L. (2011) Aging with multimorbidity: a systematic review of the literature. *Ageing Research Reviews, 10*(4), pp. 430-9.

McLean, S. L., Blenkinsopp, A. and Bennett, M. I. (2013) Using haloperidol as an antiemetic in palliative care: informing practice through evidence from cancer treatment and postoperative contexts. *Journal of Pain and Palliative Care Pharmacotherapy, 27*(2), pp. 132-135.

McLeod, P. J., Huang, A. R., Tamblyn, R. M. and Gayton, D. C. (1997) Defining inappropriate practices in prescribing for elderly people: a national consensus panel. *Canadian Medical Association Journal, 156*(3), pp. 385-391.

Mikolajczyk, R., Horn, J., Schmedt, N., Langner, I., Lindemann, C. and Garbe, E. (2015) Injury prevention by medication among children with attention-deficit/hyperactivity disorder: a case-only study. *JAMA Pediatrics, 169*(4), pp. 391-5

Moore, N., Hall, G., Sturkenboom, M., Mann, R., Lagnaoui, R. and Begaud, B. (2003) Biases affecting the proportional reporting ratio (PPR) in spontaneous reports pharmacovigilance databases: the example of sertindole. *Pharmacoepidemiology and Drug Safety, 12*(4), pp. 271-81.

Moore, T. J., Furberg, C. D., Mattison, D. R. and Cohen, M. R. (2016) Completeness of serious adverse drug event reports received by the US Food and Drug Administration in 2014. *Pharmacoepidemiology and Drug Safety*. [Epub ahead of print].

Moride, Y. and Abenhaim, L. (1994) Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. *Journal of Clinical Epidemiology, 47*(7), pp. 731-7.

O'Mahony, D., O'Sullivan, D., Byrne, S., O'Connor, M. N., Ryan, C. and Gallagher, P. (2015) STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. *Age and Ageing, 44*(2), pp. 213-8.

Ohlmeier, C., Langner, I., Hillebrand, K., Schmedt, N., Mikolajczyk, R., Riedel, O. and Garbe, E. (2015) Mortality in the German Pharmacoepidemiological Research Database (GePaRD) compared to national data in Germany: results from a validation study. *BMC Public Health, 15*, pp. 570.

Pariente, A., Avillach, P., Salvo, F., Thiessard, F., Miremont-Salame, G., Fourrier-Reglat, A., Haramburu, F., Begaud, B., Moore, N. and Association Francaise des Centres Regionaux de, P. (2012a) Effect of competition bias in safety signal generation: analysis of a research database of spontaneous reports in France. *Drug Safety, 35*(10), pp. 855-64.

Pariente, A., Fourrier-Reglat, A., Ducruet, T., Farrington, P., Beland, S. G., Dartigues, J. F., Moore, N. and Moride, Y. (2012b) Antipsychotic use and myocardial infarction in older patients with treated dementia. *Archives of Internal Medicine, 172*(8), pp. 648-53; discussion 654-5.

Pariente, A., Gregoire, F., Fourrier-Reglat, A., Haramburu, F. and Moore, N. (2007) Impact of safety alerts on measures of disproportionality in spontaneous reporting databases: the notoriety bias. *Drug Safety, 30*(10), pp. 891-8.

Pasina, L., Brucato, A. L., Falcone, C., Cucchi, E., Bresciani, A., Sottocorno, M., Taddei, G. C., Casati, M., Franchi, C., Djade, C. D. and Nobili, A. (2014) Medication non-

adherence among elderly patients newly discharged and receiving polypharmacy. *Drugs and Aging, 31*(4), pp. 283-9.

Patadia, V. K., Coloma, P., Schuemie, M. J., Herings, R., Gini, R., Mazzaglia, G., Picelli, G., Fornari, C., Pedersen, L., van der Lei, J., Sturkenboom, M., Trifiro, G. and consortium, E.-A. (2015) Using real-world healthcare data for pharmacovigilance signal detection - the experience of the EU-ADR project. *Expert Review of Clinical Pharmacology, 8*(1), pp. 95-102.

Paul-Ehrlich Institute (2016) Datenbank mit Verdachtsfällen von Impfkomplikationen und unerwünschten Arzneimittelwirkungen. Available at: http://52625146fm.pei.de/fmi/iwp/cgi?-db=UAWDB&-loadframes. Accessed: 11.03.2016

Pigeot, I. and Ahrens, W. (2008) Establishment of a pharmacoepidemiological database in Germany: methodological potential, scientific value and practical limitations. *Pharmacoepidemiology and Drug Safety, 17*(3), pp. 215-223.

Qureshi, Z. P., Seoane-Vazquez, E., Rodriguez-Monguio, R., Stevenson, K. B. and Szeinbach, S. L. (2011) Market withdrawal of new molecular entities approved in the United States from 1980 to 2009. *Pharmacoepidemiology and Drug Safety, 20*(7), pp. 772-7.

Rehman, W., Arfons, L. M. and Lazarus, H. M. (2011) The rise, fall and subsequent triumph of thalidomide: lessons learned in drug development. *Therapeutic Advances in Hematology, 2*(5), pp. 291-308.

Renom-Guiteras, A., Meyer, G. and Thurmann, P. A. (2015) The EU(7)-PIM list: a list of potentially inappropriate medications for older people consented by experts from seven European countries. *European Journal of Clinical Pharmacology, 71*(7), pp. 861-75.

Rochon, P. A. and Gurwitz, J. H. (1997) Optimising drug treatment for elderly people: the prescribing cascade. *BMJ, 315*(7115), pp. 1096.

Rognstad, S., Brekke, M., Fetveit, A., Spigset, O., Wyller, T. B. and Straand, J. (2009) The Norwegian General Practice (NORGEP) criteria for assessing potentially inappropriate prescriptions to elderly patients. A modified Delphi study. *Scandinavian Journal of Primary Health Care, 27*(3), pp. 153-9.

Rothman, K. J., Greenland, S. and Lash, T. L. (2008a) Case-control studies. In Rothman, K. J., Greenland, S. and Lash, T. L., (eds.) *Modern Epidemiology*, Philadelphia: Lippincott Williams & Wilkins. pp. 128-147.

Rothman, K. J., Greenland, S. and Lash, T. L. (2008b) Validity in epidemiologic studies. In Rothman, K. J., Greenland, S. and Lash, T. L., (eds.) *Modern Epidemiology*, Philadelphia: Lippincott Williams & Wilkins. pp. 128-147.

Schaufler, J. and Telschow, C. (2014) Arzneimittelverordnungen nach Alter und Geschlecht. In Schwabe, U. and Paffrath, D., (eds.) *Arzneiverordnungsreport 2014*, Berlin, Heidelberg: Springer-Verlag. pp. 1077-1094.

Schmedt, N., Andersohn, F. and Garbe, E. (2012) Signals of progressive multifocal leukoencephalopathy for immunosuppressants: a disproportionality analysis of spontaneous reports within the US Adverse Event Reporting System (AERS). *Pharmacoepidemiology and Drug Safety, 21*(11), pp. 1216-20.

Schmedt, N., Azoulay, L. and Hense, S. (2014) Re.: "Reduced risk of lung cancer with metformin therapy in diabetic patients: a systematic review and meta-analysis". *American Journal of Epidemiology, 180*(12), pp. 1216-7.

Schmedt, N., Bezemer, I., De Berardis, G., de Vries, C., Gil, M., Hense, S., Leal, I., Martin, E., Masclee, G., Trifiro, G., Requena, G., Rijnbeek, P., Romio, S., Sammon, C., Scotti, L., Seeger, J., Smits, M., Schink, T., Varas-Lorenzo, C., Sturkenboom, M. and Garbe, E. (2013) Benchmarking of definitions and algorithms to extract outcome events in the EU-funded SAFEGUARD project. 8th annual meeting of the German Society of Epidemiology (DGEpi), Leipzig.
Abstract available at:
http://dgepi.de/fileadmin/pdf/tagungen_workshops/Abstracts_23_08_2013__komplett_.pdf. Accessed: 11.03.2016

Schmedt, N. and Garbe, E. (2013a) Antipsychotic drug use and the risk of venous thromboembolism in elderly patients with dementia. *Journal of Clinical Psychopharmacology, 33*(6), pp. 753-8.

Schmedt, N. and Garbe, E. (2013b) Letter by Schmedt and Garbe regarding article, "statins and the risk of cancer after heart transplantation". *Circulation, 127*(5), pp. e440.

Schmedt, N., Jobski, K., Kollhorst, B., Krappweis, J., Ruther, E., Schink, T. and Garbe, E. (2016a) Treatment patterns and characteristics of older antipsychotic users in Germany. *International Clinical Psychopharmacology*. [Epub ahead of print].

Schmedt, N., Kollhorst, B., Enders, D., Jobski, K., Krappweis, J., Garbe, E. and Schink, T. (2016b) Comparative Risk of Death in Older Adults Treated with Antipsychotics: A Population-based Cohort Study. [Manuscript under review at Journal of Internal Medicine].

Schmidt, L. G. (1989) Utilization and safety of fluspirilene in nonpsychotic outpatients. *Pharmacopsychiatry, 22*(5), pp. 188-91.

Schneeweiss, S. (2006) Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and Drug Safety, 15*(5), pp. 291-303.

Schneeweiss, S. and Avorn, J. (2005) A review of uses of health care utilization databases for epidemiologic research on therapeutics. *Journal of Clinical Epidemiology, 58*(4), pp. 323-337.

Schneeweiss, S., Rassen, J. A., Glynn, R. J., Avorn, J., Mogun, H. and Brookhart, M. A. (2009) High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology, 20*(4), pp. 512-522.

Schneeweiss, S. and Wang, P. S. (2005) Claims data studies of sedative-hypnotics and hip fractures in older people: exploring residual confounding using survey information. *Journal of the American Geriatrics Society, 53*(6), pp. 948-54.

Schuster, D., Laggner, C. and Langer, T. (2005) Why drugs fail--a study on side effects in new chemical entities. *Current Pharmaceutical Design, 11*(27), pp. 3545-59.

Shin, J. Y., Choi, N. K., Jung, S. Y., Lee, J., Kwon, J. S. and Park, B. J. (2013) Risk of ischemic stroke with the use of risperidone, quetiapine and olanzapine in elderly patients: a population-based, case-crossover study. *Journal of Psychopharmacology, 27*(7), pp. 638-44.

Spitzer, W. O., Lewis, M. A., Heinemann, L. A., Thorogood, M. and MacRae, K. D. (1996) Third generation oral contraceptives and risk of venous thromboembolic disorders: an international case-control study. Transnational Research Group on Oral Contraceptives and the Health of Young Women. *BMJ, 312*(7023), pp. 83-8.

Strangfeld, A. and Zink, A. (2014) Safety of biologic therapy - results from the German biologics register RABBIT. *Deutsche Medizinische Wochenschrift, 139*(37), pp. 1817-20.

Strom, B. L. (2012) Overview of automated databases in pharmacoepidemiology. In Strom, B. L., Kimmel, S. E. and Hennessy, S., (eds.) *Pharmacoepidemiology*, Chichester: Wiley-Blackwell. pp. 158-162.

Strom, B. L. and Melmon, K. L. (2012) The use of pharmacoepidemiology to study beneficial drug effects. In Strom, B. L., Kimmel, S. E. and Hennessy, S., (eds.) *Pharmacoepidemiology*, Chichester: Wiley-Blackwell. pp. 655-677.

Sturmer, T., Wyss, R., Glynn, R. J. and Brookhart, M. A. (2014) Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *Journal of Internal Medicine, 275*(6), pp. 570-80.

Suissa, S. (1995) The case-time-control design. *Epidemiology, 6*(3), pp. 248-53.

Suissa, S. (2007) Immortal time bias in observational studies of drug effects. *Pharmacoepidemiology and Drug Safety, 16*(3), pp. 241-9.

Suissa, S. (2008a) Immeasurable time bias in observational studies of drug effects on mortality. *American Journal of Epidemiology, 168*(3), pp. 329-35.

Suissa, S. (2008b) Immortal time bias in pharmaco-epidemiology. *American Journal of Epidemiology, 167*(4), pp. 492-9.

Suissa, S. and Azoulay, L. (2012) Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care, 35*(12), pp. 2665-73.

Suissa, S., Dell'aniello, S., Vahey, S. and Renoux, C. (2011) Time-window bias in case-control studies: statins and lung cancer. *Epidemiology, 22*(2), pp. 228-31.

Suissa, S., Henry, D., Caetano, P., Dormuth, C. R., Ernst, P., Hemmelgarn, B., Lelorier, J., Levy, A., Martens, P. J., Paterson, J. M., Platt, R. W., Sketris, I., Teare, G. and Canadian Network for Observational Drug Effect, S. (2012) CNODES: the Canadian Network for Observational Drug Effect Studies. *Open Medicine, 6*(4), pp. e134-40.

Suling, M. and Pigeot, I. (2012) Signal detection and monitoring based on longitudinal healthcare data. *Pharmaceutics, 4*(4), pp. 607-40.

Swart, E., Ihle, P., Gothe, H. and Matusiewicz, D. (2014) *Routinedaten im Gesundheitswesen - Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven,* 2 ed.*,* Bern: Verlag Hans Huber.

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (2015) Guide on Methodological Standards in Pharmacoepidemiology (Revision 4). Available at: http://www.encepp.eu/standards_and_guidances/documents/ENCePPGuideofMethStandardsinPE_Rev4.pdf. Accessed: 16.03.2016

Trifiro, G., Sultana, J. and Spina, E. (2014) Are the safety profiles of antipsychotic drugs used in dementia the same? An updated review of observational studies. *Drug Safety, 37*(7), pp. 501-20.

Turnheim, K. (2003) When drug therapy gets old: pharmacokinetics and pharmacodynamics in the elderly. *Experimental Gerontology, 38*(8), pp. 843-853.

Turnheim, K. (2004) Drug therapy in the elderly. *Experimental Gerontology, 39*(11-12), pp. 1731-8.

U.S. Food and Drug Administration (2010) The Sentinel Initiative. Available at: http://www.fda.gov/downloads/Safety/FDAsSentinelInitiative/UCM233360.pdf. Accessed: 16.03.2016

U.S. Food and Drug Administration (2016) FDA Adverse Event Reporting System (FAERS). Available at: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/ucm082193. Accessed: 16.03.2016

van Staa, T. P., Dyson, L., McCann, G., Padmanabhan, S., Belatri, R., Goldacre, B., Cassell, J., Pirmohamed, M., Torgerson, D., Ronaldson, S., Adamson, J., Taweel, A., Delaney, B., Mahmood, S., Baracaia, S., Round, T., Fox, R., Hunter, T., Gulliford, M. and Smeeth, L. (2014) The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. *Health Technology Assessment, 18*(43), pp. 1-146.

van Staa, T. P., Goldacre, B., Gulliford, M., Cassell, J., Pirmohamed, M., Taweel, A., Delaney, B. and Smeeth, L. (2012) Pragmatic randomised trials using routine electronic health records: putting them to the test. *BMJ, 344*, pp. e55.

Walker, A. M. (1996) Confounding by indication. *Epidemiology, 7*(4), pp. 335-6.

Wallenstein, E. J. and Fife, D. (2001) Temporal patterns of NSAID spontaneous adverse event reports: the Weber effect revisited. *Drug Safety, 24*(3), pp. 233-7.

Weber-Schoendorfer, C. and Schaefer, C. (2016) Pregnancy outcome after tocilizumab therapy in early pregnancy-a case series from the German Embryotox Pharmacovigilance Center. *Reproductive Toxicology, 60*, pp. 29-32.

Weber, J. (1984) Epidemiology of adverse reactions to nonsteroidal antiinflammatory drugs. In Rainsford, K. D. and Velo, G. P., (eds.) *Advances in Inflammation Research*, New York: Raven Press. pp. 1-7.

Weber, T. (2008) Progressive multifocal leukoencephalopathy. *Neurologic Clinics, 26*(3), pp. 833-54, x-xi.

Wehling, M. and Burkhardt, H. (2013) *Arzneitherapie für Ältere,* 3 ed.*,* Berlin, Heidelberg: Springer Verlag.

Whitaker, H. J., Farrington, C. P., Spiessens, B. and Musonda, P. (2006) Tutorial in biostatistics: the self-controlled case series method. *Statistics in Medicine, 25*(10), pp. 1768-97.

Whitaker, H. J., Hocine, M. N. and Farrington, C. P. (2009) The methodology of self-controlled case series studies. *Statistical Methods in Medical Research, 18*(1), pp. 7-26.

World Health Organization (2016a) The use of the WHO-UMC system for standardised case causality assessment. Available at: http://who-umc.org/Graphics/24734.pdf. Accessed: 11.03.2016

World Health Organization (2016b) VigiBase®. Available at: http://www.who-umc.org/DynPage.aspx?id=98082&mn1=7347&mn2=7252&mn3=7322&mn4=7326. Accessed: 11.03.2016

Appendix A: Declaration

Hiermit versichere ich, dass ich die 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.

Niklas Schmedt Bremen, 21. März 2016