Aus dem
Helmholtz Zentrum München –
Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Institut für Gesundheitsökonomie und Management im Gesundheitswesen

Institutsleiter Professor Dr. Reiner Leidl

Evidence generation in health services research - illustrative examples and conceptual thoughts on claims data-based applications

Kumulative Habilitationsschrift
zur Befähigung für das Fachgebiet
Gesundheitsökonomie und Versorgungsforschung

vorgelegt von

Dr. LARISSA CHRISTINA MARIA SCHWARZKOPF

aus

NÖRDLINGEN

Jahr

2018

Danksagung

Dankbarkeit gehört zu den Schulden, die jeder Mensch hat,
aber nur die wenigsten tragen sie ab.

Volksmund

An dieser Stelle möchte ich allen Menschen, die wissentlich oder unwissentlich zum Gelingen meiner Habilitation beigetragen haben, aufrichtig danken.

Zuvorderst stehen die Mitglieder meines Fachmentorats, Professor Dr. Dennis Nowak, Professor Dr. Reiner Leidl und Professor Dr. Ulrich Mansmann. Sie haben mich über den gesamten Habilitationsprozess hinweg in fachlichen und organisatorischen Fragen kompetent unterstützt und mit einer Mischung aus Lebenserfahrung und Pragmatismus meine teilweise grenzwertige Detailverliebtheit in geordnete Bahnen gelenkt.

Meinen Koautoren Jürgen Bernhard-Adler, Julia Ertl, Christian Günster, Jana Hapfelmeier, Rolf Holle, Rudolf Maria Huber, Michael Kreuter, Reiner Leidl, Katharina Larisch, Markus Polke, Michaela Schunk, Amanda Tufman, Margarethe Wacker, Julia Wälscher, Julia Walter und Sabine Witt gebührt dafür Dank, dass sie ihre eigenen Perspektiven konstruktiv kritisch in unsere gemeinsamen Manuskripte eingebracht haben. Dieser Dialog hat meinen Horizont erweitert und maßgeblich zur wissenschaftlichen Qualität der in diese Habilitation eingehenden Artikel beigetragen. Bei Alina Brandes, Matthias Hunger und Doris Schopohl bedanke ich mich dafür, mir die Chance gegeben zu haben, mich als Koautorin in ihre eigenen Projekte einzubringen.

Meiner Arbeitsgruppe und allen (ehemaligen) Kollegen am IGM danke ich für ihre fortwährende Bereitschaft über einer Tasse Kaffee auch die exotischsten Projektideen zu diskutieren. Aufgrund dieses wertschätzenden und offenen Klimas war die Habilitation auch mit einem deutlichen ‚Spaßfaktor' verbunden.

Das größte ‚Dankeschön' aber gilt meiner Familie, die gemeinsam mit mir Phasen der Euphorie und der Frustration durchlaufen hat: Danke Georg, für die vorbehaltlose Unterstützung unseres unkonventionellen Familienmodells. Danke Johannes, Andreas und Lidwina, für die klaglose Akzeptanz wissenschaftlicher Texte als gute Nachtlektüre. Danke Euch allen insbesondere dafür, dass ihr mich immer wieder daran erinnert, dass akademische Titel für das ‚wahre Leben' von nachrangiger Bedeutung sind.

Alle den namentlich nicht genannten sei gewiss – Ihr seid unerwähnt doch unvergessen.

Table of contents

List of abbreviations

COPD	Chronic Obstructive Pulmonary Disease
DRG	Diagnosis Related Group
GEE	Generalized Estimation Equations
GPS	Gute Praxis Sekundärdatenanalyse
HSR	Health Services Research
ICD	International Classification of Diseases and related Health Problems
IHD	Ischaemic Heart Disease
ILD	Insterstitial Lung Disease
MEA	Managed Entry Agreement
SGB	Sozialgesetzbuch
SHI	Statutory Health Insurance
STROBE	STrengthening the Reporting of OBservational studies in Epidemiology
STROSA	STandardisierte Berichtsroutine für Sekundärdatenanalysen

Part I: Executive summary

1. Background information

a. Conceptual approach of health services research

In Germany, just about 20 years ago, 'Versorgungsforschung' (Health Services Research/HSR) was acknowledged as an independent field of research rooted in the traditions of medical sociology and health care systems research with close connections to clinical research and public health (Pfaff and Kaiser 2006).

Whilst clinical research is focused on generating either evidence on efficacy in a clinical setting or in providing early evidence on the 'added value' of a technology in routine care settings, HSR targets assessment, analysis, prognosis, appraisal, development and evaluation of routine care per se. In contrast to public health, which has a strongly preventive orientation and often targets non-medical population-based intervention strategies, HSR focuses on the management of distinct indications respectively the management of distinct patient populations by predominantly medical means. However, a clear distinction between these related research areas does not exist and the border is fluent (Raspe, Pfaff et al. 2010).

HSR is multidisciplinary in nature and broad in scope: it comprises all kinds of evidence generation on the 'use, costs, quality, accessibility, delivery, organization, financing, and outcomes of health care services' with the purpose of creating in-depth understanding of the structures, processes and effects of health care (Lohr and Steinwachs 2002). In the German context, HSR has been characterized by its population focus, the high level of patient orientation, the context-related research perspective, the pursuit of cross-sectional solutions, the overarching aim of improving the current status quo and the self-commitment to provide application-oriented, evidence-based support to policy decision makers (Pfaff and Schrappe 2011).

In consequence, the spectrum of HSR studies is manifold. It incorporates descriptive analyses of the current health care system, the evaluation of determinants of health care utilization, the conceptual development of health care interventions, the identification of key success criteria for innovative care strategies within real-world settings, the elaboration of best practice examples as well as theoretical groundwork on the operationaliza-

tion of deciding factors (e.g. ‘equity’, ‘added value’, etc.)(Nellesen-Martens, Edmund et al. 2014).

Corresponding research can address the structural framework of health care provision (input), the interactive process of health care service delivery (throughput), the intermediate results of health care service utilization (output) as well as long-term effects (outcome). According to the recent concept paper by Schrappe and Pfaff, the outcome is the most important of these four related but to some extent independent elements of a health care system. Thus, their current definition of HSR is as follows (Schrappe and Pfaff 2016):

> ‘HSR is a cross-professional field of research which – surmising from a patient and population perspective and considering complex context conditions – investigates health care structures and processes, describes outcomes on the level of daily routine care and evaluates complex interventions to improve health care.’

Owing to the long-lasting perspective and the mission to evolve health care in daily routine, randomized trials are sparse in HSR (Greiner, Witte et al. 2014), and observational studies are the main source of information (Pfaff, Glaeske et al. 2009). These observational data predominantly stem from primary data sources such as patient and population cohorts or disease-specific registries. More recently, there has been growing interest in using secondary data from claims to the Statutory Health Insurance (SHI) funds for HSR (Schubert, Köster et al. 2008, Müller and Rothgang 2015).

b. Status quo of claims data analysis in Germany

SHI claims data are the major contributor to secondary data, which are defined as any data that are routinely collected in the health care sector without a primary interest in their scientific evaluation (Swart and Ihle 2005). SHI claims data are mainly collected for the purpose of reimbursement and refer to insurants, service providers and in some parts to employers. Given this primarily documentation-oriented routine collection, claims data-based HSR can be interpreted as any kind of analyses of SHI data beyond the bare documentation purpose itself.

The content and scope of claims data are regulated by Volume V of the German Social Insurance Code (SGB V) and include in- and outpatient hospital care, outpatient physician care (general practitioner and medical specialist), drug prescriptions, rehabilitation, non-physician services (e.g. physiotherapy, occupational therapy, speech therapy) and medical aids. In addition to SHI itself, information on compulsory long-term care insurance (SGB XI) is also often augmented in the analyses.

Health insurance is mandatory in Germany, with about 80% of the resident population being insured within the SHI system. Thus, the German HSR movement soon recognized the potential of claims data for epidemiological and health economic research (Müller and Rothgang 2015). Over the last two decades, claims data analyses have become a valuable source of information, with a considerable increase in significance after the introduction of the morbidity-oriented risk structures scheme in 2009. Most contributions to the field are empirical, but a growing number of methodological applications are also found (Kreis, Neubauer et al. 2016).

To ensure and improve the methodological quality of claims data-based research, national guidelines for their conduct have been agreed (Good practice secondary data analysis/GPS) (Swart, Gothe et al. 2015). Subsequently, a uniform reporting standard, the STROSA (STandardisierte BerichtsROutine für SekundärdatenAnalysen) oriented at the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) concept for epidemiological studies, has been proposed (Swart, Bitzer et al. 2016). Moreover, ambiguous terms regarding the definition of 'health care utilization' have been identified – inclusive of suggestions on which one to apply (Schwarzkopf, Menn et al. 2012). This can be seen as an initial step towards a uniform terminology in claims data analyses. Finally, the AGENS (Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten) has established itself as a national platform for claims data analysts for methodological discussions on strategies for dealing with claim data-specific challenges.

c. Contribution of this 'Habilitationsschrift' to the scientific discussion

This 'Habilitationsschrift' aims to further contribute to the ongoing discussion on the perspectives of claims data analyses as a cornerstone of HSR in Germany in a two-fold

manner. First, it introduces illustrative examples of empirical applications to shed light on the manifold potentials of claims data-based HSR. Second, based on conceptual work, it critically appraises the usefulness of claims data-based studies for decision support in distinct aspects of HSR.

The 'Habilitationsschrift' is structured as follows: Chapter 2 portrays current standard procedures and innovative methodological advances that deal with prominent challenges in the design of claims data-based studies. With the support of illustrative examples, Chapter 3 elucidates promising fields of application for claims-data based HSR. Subsequently, Chapter 4 elaborates some conceptual thoughts on the prospects of claims data as a source for evidence generation in HSR. Each chapter culminates in a brief appraisal of the presented findings. As a complement, the final Chapter 5 provides concluding remarks on future perspectives for claims data-based HSR in Germany.

2. Conceptual challenges in the design of claims data-based studies

a. Inference in non-randomized-large scale data sets

Owing to their observational nature, the well-known issues of interpreting the results of non-randomized studies (Rovithis 2013, Martin 2014) apply to claims data as well. In consequence, established techniques for the analysis of observational data have been adopted from epidemiology and public health. These methods include multivariate regression analysis (for an application, see Walter et al. 2018 who performed a stratified comparison of different closing techniques in femoral cardiac catheterization interventions (Walter, Brandes et al. 2017), propensity score techniques ((Caliendo and Kopeinig 2008, Austin 2011); for an application, see Witt et al. who addressed the survival effects of a disease management programme via propensity score matching in open cohorts (Witt, Leidl et al. 2014)) and entropy balancing ((Hainmueller 2012); for an application, see Hofer et al. who applied a difference-in-difference approach to evaluate a tele-monitoring intervention among chronic obstructive pulmonary disease (COPD) patients (Hofer, Achelrod et al. 2016)).

Apart from general methodological developments in these fields, just recently, claims data-specific applications have been proposed for propensity scores. Braun et al. designed an algorithm that accounts for the inaccuracy of billing codes (Braun, Gorfine et al. 2017), Schneeweis et al. suggested an automated technique for propensity weighting (Schneeweiss, Rassen et al. 2009, Wyss, Fireman et al. 2018) and Karim et al. combined high dimensional propensity scores with machine learning (Karim, Pang et al. 2018).

Usually, each single study decides on just one of these techniques to achieve 'ex post' randomization. Thus, the quantitative effect of the chosen method on resulting effect estimates is not well understood. At least for the cost side, we could partially close this knowledge gap. A comparison of multivariate regression, matched pairs (age and gender) and 1:2 Greedy propensity score matching revealed substantially differing estimates for mean annual per capita SHI expenditures in COPD patients with and without ischaemic heart disease (IHD)[1] (Schwarzkopf, Wacker et al. 2016). However, when it comes to judging treatment effects in groups with non-random treatment assignment, a comprehensive comparison of the various strategies at hand is still pending. Thus, it remains a case-by-case decision for the researcher as to how to approximate the 'true' intervention effect.

In addition to the question of how to best possibly achieve quasi-randomization among the distinct groups of comparisons, the interpretation of effect sizes is not necessarily straightforward. Owing to large sample sizes, even small differences become statistically significant, even though their clinical relevance might be questionable (Lin, Jr et al. 2013). Our 2014 comparison of comorbidity burden in individuals with and without dementia revealed a significantly increased likelihood of comorbid cardiac arrhythmia in dementia patients compared with elderly control subjects without dementia, but the prevalence differed by only 0.7% (28.8% vs. 28.1%) (Bauer, Schwarzkopf et al. 2014). Regarding the costs of care, a recent master thesis at our institute identified comorbid COPD as a significant cost-driving factor in interstitial lung disease (ILD) but, given baseline costs of €4523, the resulting additional €78 (+1.7%) might not be of economic relevance (Frank 2018).

[1] Propensity score not included in final paper.

Against this background, relying on p-values alone seems fallacious. Instead, reporting of relative effects (e.g. percentages), mirroring against – if available – minimum important clinical differences, and a tentative interpretation is paramount to avoid the over-interpretation of statistically significant results.

b. Internal validation of diagnostic information

In general, claims data-based studies aim to shed light on health care service provision in patient populations that are characterized by a distinct disease. To select this study population, disease incidence/prevalence is operationalized via specific diagnostic ICD10 and ICD9 codes.

The accuracy of these codes is open to speculation, as there is always a risk of coding errors (O'Malley, Cook et al. 2005). Moreover, the current coding system is sometimes not detailed enough to allow a precise distinction for different subtypes of a disease. Regarding ILDs, for example, the different subtypes of fibrosing ILDs cannot be disentangled (Schwarzkopf, Witt et al. 2018) and, for lung cancer, even the basic distinction between small cell lung cancer and non-small cell lung cancer is not feasible (Schwarzkopf, Wacker et al. 2015). Finally, there is the non-quantifiable issue of strategic coding, which introduces a bias towards more frequent coding of diagnoses that have a positive impact on remuneration (Reinhold, Thierfelder et al. 2009).

To enhance the internal validity of documented diagnoses, several strategies have been recommended that mitigate the risk of 'false positives' (Hoffmann, Andersohn et al. 2008, Schubert, Ihle et al. 2010, Hartmann, Weidmann et al. 2016). All these techniques do not rely on one single diagnosis but require distinct patterns of diagnoses to classify a condition as present. The most prominent example of these approaches is the so-called M2Q criterion, which is used for the Hierarchical Morbidity Groups of Germany's morbidity-oriented risk structures scheme. Here, a condition (e.g. COPD) is only considered prevalent if an insurant is diagnosed in at least two different quarters of a year or if s/he received at least one inpatient diagnosis (German Federal Insurance Office 2008).

Additionally, cross-validation concepts that aim to validate the presence of a disease via a combined look at (in- and outpatient) diagnoses, drug prescriptions and the conduct of distinct medical examinations were recommended. E.g. for type II diabetes, the 'Hauner

criteria' combine diabetes diagnoses with prescriptions for anti-diabetic drugs and blood glucose measurement (Hauner, Köster et al. 2007). Moreover, a ranking of diagnoses followed by prospective validation – defined as a distinct repetition pattern of diagnoses in a pre-defined timeframe after the first diagnosis – can be applied.

As example might serve the selection of patients with ILD in the study by Schwarzkopf et al. (Schwarzkopf, Witt et al. 2018). Starting with 454,254 diagnosed individuals[2], first those without an inpatient or without an outpatient diagnosis by a relevant medical specialist (pulmonologist, internal specialist, rheumatologist) were excluded. Then, insurants without a relevant diagnostic procedure and those with implausible diagnostic patterns (e.g. 'exclusion of' diagnosis after 'confirmed' diagnosis) were dropped. In doing so, 73,167 ILD patients remained. After further omission of individuals with incomplete demographic information or with interrupted enrolment, the remainder were split into 21,543 prevalent and 36,821 incident cases. The latter were finally prospectively validated by requiring further ILD diagnoses across all quarters of the patients' individual follow-up period, allowing only one diagnosis-free quarter. Thus, 21,581 remained.

All these strategies have the disadvantage of increasing the share of 'false-negative' classifications because they might be too restrictive to address less severe cases. E.g. not every individual with diabetes requires medical treatment (Schwarzkopf, Holle et al. 2017). Thus, Schubert et al. suggested a patient grouping into 'confirmed', 'probable' and 'questionable' cases, depending on the extent and stringency of the inclusion criteria fulfilled (Schubert, Ihle et al. 2010). However, even this idea of creating an upper and a lower threshold is of limited use when acute conditions with short disease duration are investigated.

c. Quantification of economic burden

Health economic theory defines cost as a monetary valuation of resource utilization, which is not necessarily linked to cash flow, and distinguishes direct medical (resource consumption within the health care sector), direct non-medical (resource consumption

[2] ICD10 codes applied were J84.1 for 'Idiopathic Interstitial Pneumonia', J84.0, J84.8, J84.9, D48.1 for 'Other Fibrosing ILDs', D86.0–D86.9 for 'Sarcoidosis', J70.2–J70.4 for 'Drug-associated ILD', J62.0–J62.8 and J63.0–J63.8 for 'Pneumoconiosis', J70.1 for 'Radiation-associated Pneumonitis', J82 for 'Eosinophilic Pneumonia', J67.9 for 'Hypersensitivity Pneumonitis' and J99.1 for 'Connective Tissue-associated ILD'.

outside the health care sector) and indirect (productivity loss) costs of health care (Brouwer, Rutten et al. 2001).

Claims data are collected for the purpose of reimbursement within the SHI system and thus reflect just a - nevertheless meaningful - section of direct costs. In contrast, they disregard crucial cost components borne by the patients or other parts of society as well as indirect costs owing to working days lost. This narrow payer perspective has its own reasons for being, but health economic standards recommend the more comprehensive 'societal perspective' to support policy decision making (von der Schulenburg, Greiner et al. 2008, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2017).

Apart from a time lag between resource consumption and billing (Reinhold, Andersohn et al. 2011), two crucial challenges are inherent in claims data. First, services are generally reimbursed on an aggregated level that reflects a lump sum for a bundle of distinct components. In the inpatient sector, a diagnosis related group (DRG) accounts for all procedures during a patient's stay. Outpatient physician care is billed by quarter-specific account cases that summarize the entire service provision given to a patient during a quarter of the year, irrespective of the number of patient–physician contacts. HSR is often interested in a more granular view on distinct components of service provision (e.g. diagnostic costs only), which cannot be precisely disentangled. All the more challenging is the fact that services and diagnoses are not linked with each other. This is a substantial issue for cost of illness analyses per se because claims data reflect expenditures per patient rather than the disease-specific cost of care.

To best possibly disentangle related and unrelated expenditures, many claims data studies rely on excess cost approaches that compare a population with a distinct condition against a population without (Akobundu, Ju et al. 2006). The spending differences observed - after adjustment for crucial covariates - are assumed to be disease related. Referring to the example of COPD patients with and without IHD, we identfied both cohorts via M2Q criterion for the year 2011 and then assessed 2012 SHI expenditures (Schwarzkopf, Wacker et al. 2016). Here, the matched pair analyses revealed an excess of ca. €1500 in mean annual per capita costs for the cohort with comorbid IHD. This additional spending was mainly triggered by hospitalizations.

However, in absence of a control group, the only means of reflecting cost of illness is a summation of disease-specific spending (Akobundu, Ju et al. 2006). Here, the

classification of services as disease related opens the scope for discretionary decisions. For example, we developed two alternative strategies to reflect lung cancer-specific cost of care (Schwarzkopf, Wacker et al. 2015). The narrow definition ('main analysis') accounted for spending on traditional chemotherapeutics, monoclonal antibodies, inpatient treatment with a principal diagnosis of lung cancer, rehabilitation stays with lung cancer as the reason for admission and outpatient physician contacts with at least one 'confirmed' diagnosis of lung cancer. Our 'broad' definition (sensitivity analysis) amended this spectrum by hospitalizations with a secondary diagnosis of lung cancer and by chemo- and radiotherapy-relevant co-medication (antiemetic preparations, antibiotics, bisphosphonates, anti-anaemic preparations and mesna). Given the lack of a link between diagnoses and related services, both operationalizations to some extent imprecise.

Moreover, the sum disease-specific approach might not fully capture costs of illness. The presence of a distinct condition most probably impacts on the costs of care in not per se disease-associated situations, because morbidity affects treatment options. This hypothesis was supported by an add-on analysis in the COPD example, which quantified expenditures for IHD-related drugs and hospitalizations owing to IHD. For both sectors, this disease-specific summation remained substantially below the sector-specific excess costs (inpatient: €819 vs. €65, drugs €297 vs. €120) (Schwarzkopf, Wacker et al. 2016). Therefore, the excess cost method, which does not intend to explicitly model disease-related expenditures, seems to be the method of choice and should be implemented whenever applicable.

3. Fields of application for claims data-based health services research

a. Health impact of comorbidity burden

The presence of distinct comorbid conditions influences treatment options, clinical prognosis and management requirements. Thus, detailed information on comorbidity burden is required to judge the health care service needs of a patient population comprehensively.

In this regard, the first descriptive step is a detailed assessment of comorbidity burden in the population of interest. Given that ICD10 codes enable an almost unlimited variety of diagnoses, setting a focus on distinct conditions seems paramount. To do so, several claims data-based algorithms have been suggested (Charlson, Pompei et al. 1987, Incalzi, Capparella et al. 1997, Elixhauser, Steiner et al. 1998, Gagne, Glynn et al. 2011). These generic indices by trend disregard certain conditions that are either highly prevalent or of particular clinical interest in distinct diseases (Bauer, Schwarzkopf et al. 2014, Schwarzkopf, Witt et al. 2018). Therefore, an indication-oriented case-by-case adaptation is more likely to reflect the care-relevant burden of disease.

Using the example of ILD (Schwarzkopf, Witt et al. 2018), the comorbidities included in the Elixhauser comorbidity index include some ILD subtypes within the domain 'chronic pulmonary disease', but disregard highly prevalent conditions such as IHD and conditions of clinical importance such as 'gastro-oesophageal reflux disease'. After respective corrections of the index, we were able to comprehensively portray the comorbidity burden associated with ILDs. Moreover, this piece of research elucidated which comorbid conditions are especially sensitive to concepts of multiple coding (M2Q criterion) and prospective validation.

Being aware of epidemiological burden, associations between comorbid conditions and outcomes of interest such as e.g. survival are of relevance for HSR to identify vital starting points for comprehensive patient management. In this regard, a simplified look at comorbidity prevalence per se is not sufficient to address associations concisely. It is even more important to raise awareness as to what extent the (pharmaceutical) treatment of comorbidity influences prognosis. In doing so, some kind of hierarchy which conditions might be prioritized in case of multi-morbidity becomes feasible. Proceeding with the ILD example, it could be shown that IHD only had a detrimental association with survival if untreated, whereas treated IHD had no association with all-cause mortality (Schwarzkopf, Witt et al. 2018).

Subsequently, the cost impact of distinct comorbid conditions should be quantified to reflect the economic burden on the individual, but also at the population level. Regarding ILDs, the master thesis of Frank demonstrated that particularly comorbidities of low prevalence were highly cost driving (e.g. lung cancer), whereas highly prevalent comorbid conditions often had a limited cost impact (e.g. COPD) (Frank 2018).

As a further outlook, combining information on comorbidity prevalence with information on cost impact opens the box for comprehensive budget impact analyses, which might also take age and gender aspects (potential for age- and gender-specific comorbidity and cost profiles) into account. Corresponding figures would be of utmost importance to support evidence-based priority setting in comorbidity management.

b. Structures of health care spending

Claims data cover various aspects of health care service provision with inpatient care, outpatient care and drug prescriptions as the most relevant elements. In consequence, not only the expenditure level per se, but also the underlying structures can be reflected in cross-sectional and longitudinal analyses. Given that claims data often follow up an insurant over a period of several years, they have particular potential for incidence-based cost of illness approaches that accumulate health care expenditures over the lifetime course of a disease (Tarricone 2006) or specific periods as exemplified in Schwarzkopf et al. 2015 (Schwarzkopf, Wacker et al. 2015).

A necessary precondition in this regard is the realignment of costs of care with patient/individual follow-up periods because, other than for prevalence-based approaches (for an application see the COPD with(out) IHD study of Schwarzkopf et al. (Schwarzkopf, Wacker et al. 2016)), relying on accounting years is not feasible. Outpatient diagnoses are reported on a quarterly basis and, therefore, defining 'incidence' not by date but by quarter yields the highest degree of precision. This operationalization was chosen for an exemplary study of ours that reports all-cause and disease-specific costs of care among lung cancer patients from the quarter of diagnosis over a patient-individual 3-year follow-up period (Schwarzkopf, Wacker et al. 2015).

The corresponding analyses revealed costs per case of approximately €38,700 over the entire observation period. Just over half this amount was related to lung cancer treatment. Referring to disease-related costs, inpatient care was the main contributor followed by lung cancer-related drugs. This structure could be replicated within different treatment strata (surgery, chemo-/radiotherapy, no tumour-directed therapy), even though proportions and absolute amounts varied to some extent.

In addition to reporting costs per case, the longitudinal nature of claims data enables the detection of changes in the cost structure over time. Proceeding with the lung cancer example, domain-specific expenditures were also reported quarterwise over the entire patient-individual follow-up period.

These analyses revealed that the initial 6 months post diagnosis accounted for more than two-fifths of all-cause costs per case, and for just half of lung cancer-related costs per case. After this initial peak, all-cause spending, lung cancer-related spending and the share of lung cancer-related spending in all cause-spending decreased over time. This effect was mainly driven by an abrupt decline in spending in the inpatient sector right after the quarter of diagnosis. The most probable explanation for this observation is that cost-intense surgical procedures were in general performed witin 6 months after diagnosis. Moreover, it became obvious that spending on lung cancer-related drugs increased in the initial phase post diagnosis but remained quite stable from the third quarter onwards. This might be related to the fact that, after the onset of chemotherapeutic interventions, corresponding treatment is provided continuously over an extended period of time. In the course of the disease, lung cancer-related drugs become the crucial component of disease-related costs because they exceed inpatient spending from the second year post diagnosis onwards (Schwarzkopf, Wacker et al. 2015).

These trends and shifts would have been masked if only aggregated costs per case had been evaluated. Indeed, this disaggregated perspective is a valuable contribution towards a better judgement on ‘common’ care pathways in the course of the disease. To further enhance the methodological quality of time trend analysis, and to allow investigations on the effect of potential impact factors over time (e.g. age, gender, morbidity), a standard application of techniques that account for intra-subject correlation (Fitzmaurice, Laird et al. 2011), such as generalized estimation equations (GEEs) (for an eapplication, see (Schwarzkopf, Hao et al. 2014)) and mixed models (for an application, see (Schwarzkopf, Holle et al. 2017)), is strongly recommended.

c. Subpopulation-specific display of health care service utilization pattern

As a key component of HSR, claims data not only include diagnoses and reimbursement information but incorporate detailed information on service provision within the SHI system. This opens the scope for in-depth hypothesis-driven comparisons of health care service provision between different patient populations. Here, aspects of equity in access as well as on quality of care are of particular interest. Despite claims data not including direct information on output and outcome, these aspects of care quality can be approximated with some creative efforts.

An example in this regard is putting diagnostic information on complications (outcome) into context with data on provision of distinct medical services (input), as done within our study on diabetes care in community-living and institutionalized individuals with dementia (Schwarzkopf, Holle et al. 2017). Here, we contrasted the setting-specific conduct of guideline-recommended diabetes-relevant medical examinations as well as the occurrence of diabetes-related complications in community-dwellers and nursing home residents. The care level-stratified analyses unveiled a consistent trend towards less frequent controls of blood parameters and ocular background in the institutional setting. In parallel, we observed a rather heterogeneous picture regarding the occurrence of diabetes-related complications (Schwarzkopf, Schunk et al. 2014). Altogether, this piece of evidence reflects an initial example of how to (partially) appraise the effectiveness of health care provision within claims data. Moreover, the findings at the population level point to subpopulations that are in particular need of extended service provision. This modus operandi matches with the improvement-oriented character of HSR.

As soon as subgroup differences in health care service utilization patterns are unveiled, subsequent explanatory analyses to identify the key impact factors on these differences are useful. Assuming that less aggressive end-of-life care (process) corresponds to patient preferences and thus indirectly enhances quality of life (outcome), we reflected aggressiveness of end-of-life care in lung cancer patients based on the receipt of distinct services (Walter, Tufman et al. 2018). To factor in potential inequity in service access between the rural and the urban setting, patient-individual zip codes were linked with official statistics from the ‘Bundesamt für Bauwesen und Raumordnung’ (i.e. ‘German Federal Institute for Research on Building, Urban Affairs, and Spatial Development’) to

classify the distinct residential areas correspondingly. The analyses substantiated evidence on equity in service access for rural and urban populations with supportive care being in general capable of expansion. Moreover, the presence of distinct comorbid conditions and previous tumour-directed therapy were identified as the main triggering factors for differences regarding the structure of end-of-life care.

To forge the bridge from population-level evidence on health care utilization to patient-level care trajectories, longitudinal analyses that take advantage of the techniques of data mining and machine learning seem to be particular promising. A first descriptive step in this regard is a recent paper by Vogt and colleagues (Vogt, Scholz et al. 2018), who identified typical outpatient careers in individuals with heart failure via sequence clustering. Subsequently, these careers might be compared regarding their effectiveness (e.g. time to hospitalization) via the techniques outlined in Chapter 2a to derive best practice examples for advisable patient management strategies.

4. Informative value of claims data to generate 'real-world' evidence

a. Relationship between claims data and other data sources

To judge the informative value of claims data for evidence generation soundly, a discussion against other sources of evidence is of utmost importance. In HSR, these are particularly disease-specific registries and (in part) survey-based cohort studies.

Referring to the example of haemophilia (Schopohl, Bidlingmaier et al. 2018), we could demonstrate that claims data outperform registries regarding economic research and comprehensive addressing of morbidity-related confounders. Moreover, they were classified as comparably well suited to reflect aspects of pharmacovigilance with in parallel substantial drawbacks in outcomes research (e.g. quality of life, short- and long-term clinical results). These findings seem by and large to be transferable to other indication areas where disease-specific registries exist (e.g. ILDs). This is also true for the conclusion that some of the claims data-related limitations could be mitigated by further ad-

vancements in the ICD coding system. Here, shifting the coding requirements to the recently published ICD11 might be a step in the right direction.

However, knowledge of different spectra of information compared with other sources of evidence is a necessary but not sufficient condition. All the more, therefore, further insights on systematic differences between secondary data-based information and self-reports on health care service utilization are paramount. Our comparison of survey-reported health care utilization from the KORA cohort with external data from claims data and official statistics indicated a substantial amount of selection and recall bias (Hunger, Schwarzkopf et al. 2013). Particularly service utilization related to increased frailty of the user such as nursing care or inpatient treatment was found to be heavily underreported in self-reports. This piece of evidence can be considered as starting point for deriving age- and gender-specific correction factors for survey-based reports on health care service utilization.

b. Claims data analyses as a tool for policy decision support

Qualitative and quantitative details on comprehensiveness and representativeness of information derived from claims data create a vital basis for evaluating their prospects for evidence generation in support of reimbursement decisions.

Particularly in the Anglo-American setting, it is common to not fully reimburse innovative technologies right from market entry because there is a substantial amount of uncertainty regarding health outcome, service utilization and long-term costs of care. Instead, further evidence on the technology of interest is collected in the real-world setting, and provider payment is linked to the performance of the technology. To do so, service providers and manufacturer conclude so-called managed entry agreement (MEA) contracts, where both parties agree on the conditions for full service coverage at a later point in time. The final decision is tied to the evaluation of the added value of the innovative technology in the routine care setting with corresponding evidence being collected during the introduction phase (Hutton, Trueman et al. 2007, Stafinski, McCabe et al. 2010).

Regarding the German SHI, MEAs are still of subordinate relevance. Since the Health Care Provision Act (GKV-Versorgungsstrukturgesetz) in 2011, the interest in such schemes has been growing because regulatory bodies opened the scope for selective

contracts with evidence development (§137e SGBV). For this evolving field, we elaborated a generic value-based typology of MEAs (Brandes, Schwarzkopf et al. 2016). Subsequently, the contribution of claims data to the respective forms was appraised based on the criteria data availability, completeness, timeliness, confidentiality, reliability and validity. These qualitative analyses revealed limited usefulness of claims data to target the safety aspects of a new technology, while acknowledging their substantial benefits regarding a reduction in uncertainty about the utilization and costs associated with a new technology. This typology might be a valuable supporting tool for SHIs that are not well versed yet with key success criteria for selective contracting.

5. Concluding summary

Having depicted the challenges of claims data in the context of HSR, this 'Habilitationsschrift' has outlined their prospects for descriptive, analytical, and some extent quasi-experimental HSR with illustrative examples. To advocate the role of SHI claims data as a cornerstone of HSR in Germany, additional conceptual thoughts on their contributions towards evidence generation for policy decision making were elaborated.

Chapter 2 expounded prominent challenges for the design of claims data-based HSR studies and unveiled state of the art strategies to deal with these issues. In summary, traditional methods from epidemiological and public health statistics can be transferred to claims data analyses in many instances. To enhance precise patient classification, the application of internal validation techniques, e.g. the M2Q criterion, combined with prospective validation was recommended.

Based on illustrative examples, Chapter 3 introduced 'morbidity burden', 'budget impact' and 'patient pathways' as particularly promising fields of application for claims data-based HSR. As take home messages, this section emphasized the necessity to move to disease-specific adaptations of established measures of comorbidity burden, encouraged detailed disentangling of cost structures in the longitudinal view and provided suggestions for an operationalization of aspects related to quality of care.

Subsequently, Chapter 4 classified claims data as a valuable source for (early) evidence generation in economically oriented HSR and HSR in especially vulnerable populations.

In summary, the work presented has emphasized that the use of claims data for questions of HSR is a vital challenge, because data are initially collected for the purpose of reimbursement. In consequence, the conduct of claims data-based HSR requires profound knowledge of health care system-related framework conditions (and their changes over time) and clinical disease-specific expertise to ensure a scientifically sound interpretation of the observed documentation patterns.

In this regard, intense communication between data owners and external analysts in combination with early cooperation with clinical specialists enhances the quality of research. Such multi-disciplinary cooperation promotes the case-by-case provision of data sets tailored to the research question of interest. Nevertheless, this cannot fully overcome the generic issues of claims data-based research: first, the clients of a distinct SHI fund might not be representative of the resident German population. Second, the lack of clinical data (e.g. disease severity) and patient-reported outcomes (e.g. health-related quality of life) limit the scope for HSR. Referring to the input–throughput–output model, the strengths of claims data are 'input', but 'process' and 'long-term result' can only be addressed with great creative effort. However, these two components are crucial for sound support of policy decision makers.

In addition, sound policy decision support requires scientific consensus on how to conduct and present claims data analyses. In this regard, national standards such as GPS and STROSA are a necessary initial step in the right direction but further harmonization beyond seems deeply required. There is great need for a common understanding on how to validate diagnoses, how to define disease incidence and how to operationalize disease-related costs of care. Moreover, the installation of some kind of permanent team that focuses on the prospects and strategies for the operationalization of patient-reported outcomes within claims data will mean significant progress. Finally, a careful adaptation of current data protection laws that alleviates the linkage of information from different data sources (claims data, survey, registries), such as is possible in some Nordic countries, is indispensable in the long run to take full advantage of mutual synergies to foster patient-centred HSR.

6. References

Akobundu, E., J. Ju, L. Blatt and C. D. Mullins (2006). "Cost-of-illness studies : a review of current methods." Pharmacoeconomics **24**(9): 869-890.

Austin, P. C. (2011). "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." Multivariate Behav Res **46**(3): 399–424.

Bauer, K., L. Schwarzkopf, E. Graessel and R. Holle (2014). "A claims data-based comparison of comorbidity in individuals with and without dementia." BMC Geriatr **14**(1): 10.

Brandes, A., L. Schwarzkopf and W. H. Rogowski (2016). "Using claims data for evidence generation in managed entry agreements." Int J Technol Assess Health Care: 1-9.

Braun, D., M. Gorfine, G. Parmigiani, N. D. Arvold, F. Dominici and C. Zigler (2017). "Propensity scores with misclassified treatment assignment: a likelihood-based adjustment." Biostatistics **18**(4): 695-710.

Brouwer, W., F. Rutten and M. Koopmanschap (2001). Costing in economic evaluations. Economic evaluation in health care. Merging theory with practice. M. Drummond and A. McGuire. Oxford, Oxford University Press**:** 68-93.

Caliendo, M. and S. Kopeinig (2008). "Some practical guidance for the implementation of propensity score matching." J Econ Surveys **22**(1): 31-72.

Charlson, M. E., P. Pompei, K. L. Ales and C. R. MacKenzie (1987). "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." J Chronic Dis **40**(5): 373-383.

Elixhauser, A., C. Steiner, D. R. Harris and R. M. Coffey (1998). "Comorbidity measures for use with administrative data." Med Care **36**(1): 8-27.

Fitzmaurice, G., N. Laird and J. Ware (2011). Applied Longitudinal Analysis. 2nd edition,. New York, Wiley.

Frank, A. L. (2018). Economic burden of incident interstitial lung disease (ILD) and the impact of comorbidity on costs of care. Master Thesis, MSc Public Health, Ludwig-Maximilians-Universität München.

Gagne, J. J., R. J. Glynn, J. Avorn, R. Levin and S. Schneeweiss (2011). "A combined comorbidity score predicted mortality in elderly patients better than existing scores." J Clin Epidemiol **64**(7): 749-759.

German Federal Insurance Office. (2008). "The new risk structure equalisation.", from http://www.bundesversicherungsamt.de/risikostrukturausgleich.html.

Greiner, W., J. Witte, M. Steffens, R. Böttger and C. Burgardt (2014). "Methodische und institutionelle Hürden der Versorgungsforschung in Deutschland." Gesundh ökon Qual manag **19**(04): 184-193.

Hainmueller, J. (2012). "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." Political Analysis **20**(1): 25-46.

Hartmann, J., C. Weidmann and R. Biehle (2016). "[Validation of SHI Claims Data Exemplified by Gender-specific Diagnoses]." Gesundheitswesen **78**(10): e53-e58.

Hauner, H., I. Köster and I. Schubert (2007). "Prevalence of Diabetes mellitus and Quality of Care in Hesse, Germany, 1998–2004." Dtsch Arztebl Int **104**(41): 2799-2805.

Hofer, F., D. Achelrod and T. Stargardt (2016). "Cost-Utility Analysis of Telemonitoring Interventions for Patients with Chronic Obstructive Pulmonary Disease (COPD) in Germany." Appl Health Econ Health Policy **14**(6): 691-701.

Hoffmann, F., F. Andersohn, K. Giersiepen, E. Scharnetzky and E. Garbe (2008). "[Validation of secondary data. Strengths and limitations]." Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz **51**(10): 1118-1126.

Hunger, M., L. Schwarzkopf, M. Heier, A. Peters, R. Holle and K. S. Group (2013). "Official statistics and claims data records indicate non-response and recall bias within survey-based estimates of health care utilization in the older population." BMC Health Serv Res **13**: 1.

Hutton, J., P. Trueman and C. Henshall (2007). "Coverage with evidence development: an examination of conceptual and policy issues." Int J Technol Assess Health Care **23**(4): 425-432.

Incalzi, R. A., O. Capparella, A. Gemma, F. Landi, E. Bruno, F. Di Meo and P. Carbonin (1997). "The interaction between age and comorbidity contributes to predicting the mortality of geriatric patients in the acute-care hospital." J Intern Med **242**(4): 291-298.

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2017). Allgemeine Methoden : Version 5.0 / Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Version 5.0 Köln.

Karim, M. E., M. Pang and R. W. Platt (2018). "Can We Train Machine Learning Methods to Outperform the High-dimensional Propensity Score Algorithm?" Epidemiology **29**(2): 191-198.

Kreis, K., S. Neubauer, M. Klora, A. Lange and J. Zeidler (2016). "Status and perspectives of claims data analyses in Germany-A systematic review." Health Policy **120**(2): 213-226.

Lin, M., H. C. L. Jr and G. Shmueli (2013). "Research Commentary—Too Big to Fail: Large Samples and the p-Value Problem." Information Syst Res **24**(4): 906-917.

Lohr, K. N. and D. M. Steinwachs (2002). "Health Services Research: An Evolving Definition of the Field." Health Serv Res **37**(1): 15-17.

Martin, W. (2014). "Making valid causal inferences from observational data." Prev Vet Med **113**(3): 281-297.

Müller, R. and H. Rothgang (2015). Möglichkeiten und Grenzen der Versorgungsforschung mit Kassendaten. Versorgungsforschung - zwischen Routinedaten, Qualitätssicherung und Patientenorientierung. C. Trittin. Sieburg**:** 15-26.

Nellesen-Martens, G., A. Edmund and M. Neugebauer (2014). "Qualität, Nutzen und Wirksamkeit der Versorgungsforschung - Ableitung von Konsequenzen für den Innovationsfond; Stellungnahme des Deutschen Netzwerks Versorgungsforschung (DNVF) e.V.".

O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle and C. M. Ashton (2005). "Measuring diagnoses: ICD code accuracy." Health Serv Res **40**(5 Pt 2): 1620-1639.

Pfaff, H., G. Glaeske, E. A. M. Neugebauer and M. Schrappe (2009). "Memorandum III: Methoden für die Versorgungsforschung (Teil I)." Gesundheitswesen **71**(08/09): 505-510.

Pfaff, H. and C. Kaiser (2006). "Aufgabenverständnis und Entwicklungsstand der Versorgungsforschung." Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz **49**(2): 111-119.

Pfaff, H. and M. Schrappe (2011). Einführung in die Versorgungsforschung. Lehrbuch Versorgungsforschung. H. Pfaff, G. Glaeske, E. Neugebauer and M. Schrappe. Stuttgart, Schattauer Verlag**:** 2-57.

Raspe, H., H. Pfaff, M. Härter, D. Hart, U. Koch-Gromus, F. Schwartz, J. Siegrist, H. Wittchen and F. Wissing (2010). Versorgungsforschung in Deutschland: Stand - Perspektive - Förderung: Stellungnahme, Standpunkte. Weinheim, Wiley-VCH.

Reinhold, T., F. Andersohn, F. Hessel, B. Brüggenjürgen and S. N. Willich (2011). "Die Nutzung von Routinedaten der gesetzlichen Krankenkassen (GKV) zur Beantwortung gesundheitsökonomischer Fragestellungen – eine Potenzialanalyse." Gesundh ökon Qual manag **16**(03): 153-159.

Reinhold, T., K. Thierfelder, F. Muller-Riemenschneider and S. N. Willich (2009). "[Health economic effects after DRG-implementation--a systematic overview]." Gesundheitswesen **71**(5): 306-312.

Rovithis, D. (2013). "Do health economic evaluations using observational data provide reliable assessment of treatment effects?" Health Econ Rev **3**(1): 21.

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

Schopohl, D., C. Bidlingmaier, D. Herzig, R. Klamroth, K. Kurnik, D. Rublee, W. Schramm, L. Schwarzkopf and K. Berger (2018). "Prospects for research in

haemophilia with real-world data-An analysis of German registry and secondary data." Haemophilia. **24**(4):584-94.

Schrappe, M. and H. Pfaff (2016). "Versorgungsforschung vor neuen Herausforderungen: Konsequenzen für Definition und Konzept." Gesundheitswesen **78**(11): 689-694.

Schubert, I., P. Ihle and I. Köster (2010). "[Internal confirmation of diagnoses in routine statutory health insurance data: concept with examples and case definitions]." Gesundheitswesen **72**(6): 316-322.

Schubert, I., I. Köster, J. Küpper-Nybelen and P. Ihle (2008). "Versorgungsforschung mit GKV-Routinedaten." Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz **51**(10): 1095-1105.

Schwarzkopf, L., Y. Hao, R. Holle and E. Graessel (2014). "Health Care Service Utilization of Dementia Patients before and after Institutionalization: A Claims Data Analysis." Dementia and Geriatric Cognitive Disorders Extra **4**(2): 195-208.

Schwarzkopf, L., R. Holle and M. Schunk (2017). "Effects of Nursing Home Residency on Diabetes Care in Individuals with Dementia: An Explorative Analysis Based on German Claims Data " Dementia and Geriatric Cognitive Disorders Extra **7**(1): 41-51.

Schwarzkopf, L., P. Menn and R. Holle (2012). "[Discretionary decisions in claims data analyses and their effects - an explorative comparison using the example of a cost of illness study on dementia]." Gesundheitswesen **74**(8-9): e76-83.

Schwarzkopf, L., M. Schunk and R. Holle (2014). Abschlussbericht - Analyse der Versorgungssituation der Versorgungssituation von Menschen mit Demenz und Diabetes. Forschungsförderung 2012 der Deutschen Alzheimer Gesellschaft e.V. https://www.deutsche-alzheimer.de/fileadmin/alz/forschung/2012_Abschlussbericht_Kassendatenbasierte_Analyse_der_Versorgungssituation_von_Menschen_mit_Demenz_und_Diabetes.pdf.

Schwarzkopf, L., M. Wacker, J. Ertl, J. Hapfelmeier, K. Larisch and R. Leidl (2016). "Impact of chronic ischemic heart disease on the health care costs of COPD patients – An analysis of German claims data." Resp Med **118**: 112-118.

Schwarzkopf, L., M. Wacker, R. Holle, R. Leidl, C. Günster, J.-B. Adler and R. M. Huber (2015). "Cost-components of lung cancer care within the first three years after initial diagnosis in context of different treatment regimens." Lung Cancer **90**(2): 274-280.

Schwarzkopf, L., S. Witt, J. Waelscher, M. Polke and M. Kreuter (2018). "Associations between comorbidities, their treatment and survival in patients with interstitial lung diseases – a claims data analysis." Resp Res **19**(1): 73.

Stafinski, T., C. J. McCabe and D. Menon (2010). "Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems." Pharmacoeconomics **28**(2): 113-142.

Swart, E., E. M. Bitzer, H. Gothe, M. Harling, F. Hoffmann, D. Horenkamp-Sonntag, B. Maier, S. March, T. Petzold, R. Röhrig, A. Rommel, T. Schink, C. Wagner, S. Wobbe and J. Schmitt (2016). "A Consensus German Reporting Standard for Secondary Data Analyses, Version 2 (STROSA-STandardisierte BerichtsROutine für SekundärdatenAnalysen)." Gesundheitswesen **78**(S 01): e145-e160.

Swart, E., H. Gothe and S. Geyer (2015). "Gute Praxis Sekundärdatenanalyse (GPS) - Leitlinien und Empfehlungen." Das Gesundheitswesen **77**(02): 120-126.

Swart, E. and P. Ihle (2005). "Gute Praxis Sekundärdatenanalyse (GPS)". Das Gesundheitswesen **67**(07): P38.

Tarricone, R. (2006). "Cost-of-illness analysis: What room in health economics?" Health Policy **77**(1): 51-63.

Vogt, V., S. M. Scholz and L. Sundmacher (2018). "Applying sequence clustering techniques to explore practice-based ambulatory care pathways in insurance claims data." Eur J Public Health **28**(2): 214-219.

von der Schulenburg, J.-M., W. Greiner, F. Jost, N. Klusen, M. Kubin, R. Leidl, T. Mittendorf, H. Rebscher, O. Schoeffski, C. Vauth, T. Volmer, S. Wahler, J. Wasem and C. Weber (2008). "German recommendations on health economic evaluation: third and updated version of the Hanover Consensus." Value Health **11**(4): 539-544.

Walter, J., A. Brandes, M. Sinner, W. Rogowski and L. Schwarzkopf (2017). "A Comparison of Costs and Complications of Manual Compression and Vascular Closing Devices in Femoral Cardiac Catheterization Interventions." Interventional Card J **3**(2).

Walter, J., A. Tufman, R. Leidl, R. Holle and L. Schwarzkopf (2018). "Rural versus urban differences in end-of-life care for lung cancer patients in Germany." Support Care Cancer **26**(7): 2275-2283.

Witt, S., R. Leidl, C. Becker, R. Holle, M. Block, J. Brachmann, S. Silber and B. Stollenwerk (2014). "The effectiveness of the cardiovascular disease prevention programme 'KardioPro' initiated by a German sickness fund: a time-to-event analysis of routine data." PLoS One **9**(12): e114720.

Wyss, R., B. Fireman, J. A. Rassen and S. Schneeweiss (2018). "Erratum: High-dimensional propensity score adjustment in studies of treatment effects using health care claims data." Epidemiology.

Part II: Publications

1. Empirical Applications

Schwarzkopf L, Witt S, Waelscher J, Polke M, Kreuter M. *Associations between comorbidities, their treatment and survival in patients with interstitial lung diseases – a claims data analysis*. Respiratory Research. 2018;19(1):73.

doi: 10.1186/s12931-018-0769-0.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5918773/

Schwarzkopf L, Wacker M, Holle R, Leidl R, Günster C, Adler J-B, Huber RM. *Cost-components of lung cancer care within the first three years after initial diagnosis in context of different treatment regimens*. Lung Cancer. 2015;90(2):274-80.

doi: 10.1016/j.lungcan.2015.09.005.

https://www.sciencedirect.com/science/article/pii/S0169500215300490?via%3Dihub

Schwarzkopf L, Wacker M, Ertl J, Hapfelmeier J, Larisch K, Leidl R. *Impact of chronic ischemic heart disease on the health care costs of COPD patients – An analysis of German claims data*. Respiratory Medicine. 2016;118:112-8.

doi: 10.1016/j.rmed.2016.08.001

https://www.sciencedirect.com/science/article/pii/S0954611116301743?via%3Dihub

Schwarzkopf L, Holle R, Schunk M. *Effects of Nursing Home Residency on Diabetes Care in Individuals with Dementia: An Explorative Analysis Based on German Claims Data* Dementia and Geriatric Cognitive Disorders Extra. 2017;7(1):41-51.

doi: 10.1159/000455071

https://www.karger.com/Article/FullText/455071

Walter J, Tufman A, Leidl R, Holle R, **Schwarzkopf L**. *Rural versus urban differences in end-of-life care for lung cancer patients in Germany*. Supportive care in cancer: 2018;26(7):2275-83.

doi: 10.1007/s00520-018-4063-y

https://link.springer.com/article/10.1007%2Fs00520-018-4063-y

2. Conceptual thoughts

Schopohl D, Bidlingmaier C, Herzig D, Klamroth R, Kurnik K, Rublee D, Schramm W, **Schwarzkopf L**, Berger K. *Prospects for research in haemophilia with real-world data-An analysis of German registry and secondary data*. Haemophilia: 2018;24(4):584-94.

doi: 10.1111/hae.13443.

https://onlinelibrary.wiley.com/doi/full/10.1111/hae.13443

Hunger M, **Schwarzkopf L**, Heier M, Peters A, Holle R. *Official statistics and claims data records indicate non-response and recall bias within survey-based estimates of health care utilization in the older population*. BMC Health Serv Res. 2013;13:1.

doi: 10.1186/1472-6963-13-1.

https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-13-1

Brandes A, **Schwarzkopf L**, Rogowski WH. *Using claims data for evidence generation in managed entry agreements.* International journal of technology assessment in health care. 2016:1-9.

doi: 10.1017/S0266462316000131

https://www.cambridge.org/core/journals/international-journal-of-technology-assessment-in-health-care/article/using-claims-data-for-evidence-generation-in-managed-entry-agreements/D0A8C77B75F7D080E86D69324E0BEB7B

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Habilitationsschrift mit dem Titel

Evidence generation in health services research
- illustrative examples and conceptual thoughts
on claims data-based applications

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Fahrenzhausen, den 8. Oktober 2018 _____

Larissa Schwarzkopf