© The Author(s) 2020

F. Peters[1] · T. Kreutzburg[1] · J. Kuchenbecker[1] · U. Marschall[2] · M. Remmel[3] ·
M. Dankhoff[4] · H.-H. Trute[5] · T. Repgen[5] · E. S. Debus[1] · C.-A. Behrendt[1]

[1] Forschungsgruppe GermanVasc, Klinik und Poliklinik für Gefäßmedizin, Universitäres Herz- und Gefäßzentrum UKE Hamburg, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

[2] BARMER Hauptverwaltung, Wuppertal, Germany

[3] Klinik und Poliklinik für Kardiologie, Universitäres Herz- und Gefäßzentrum UKE Hamburg, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

[4] DAK-Gesundheit, Hamburg, Germany

[5] Fakultät für Rechtswissenschaft, Universität Hamburg, Hamburg, Germany

Quality of care in surgical/interventional vascular medicine: what can routinely collected data from the insurance companies achieve?

Introduction

Countless laws and directives oblige stakeholders in the German healthcare system to continuously assure and improve the quality of patient care. Further motivators arise from reasons of competition, regulatory influence (e.g. pay for performance, EU Medical Device Regulation) or scientific interest; however, assuring and improving quality requires valid measurement of the quality of medical care.

Although modern funding structures were developed for the healthcare system as early as the nineteenth century, a broad public debate on performance-related remuneration only began in the USA in the 1960s [17]. It was also during this time that work started at Yale University on the development of a uniform classification system for diseases and treatments. This classification was introduced some 20 years later as the diagnosis-related groups (DRG) within the Centers for Medicare & Medicaid Services (CMS) system [5]. Even back

The German version of this article can be found under https://doi.org/10.1007/s00772-020-00664-x.

then, the treatment and its quality were to already be linked to the remuneration structures, whereby this raised countless questions and led to partial interests. At its core, the discussion has always been about the objective measurement of the quality of medical care. A construct that differentiates between the quality of the structure and process and the quality of the outcome is still generally accepted today [17]. To date, this has required the collection of valid primary data in clinical registries [8]; however, more than half a century after the broad public debate, many questions have still not been conclusively answered.

Surgical and interventional vascular medicine is a subject that is continuously evolving and counts among the great challenges in modern medicine, both in clinical and scientific terms. Here, common diseases such as peripheral arterial occlusive disease (PAOD) occur at the same time as rare diseases, such as the genetic aortic syndrome. Established low-risk procedures are accompanied by extremely complex endovascular procedures involving very high risks. The discipline is also characterised by rapid innovation cycles in medical devices. Besides the procedure-related parameters and di-

rect technical outcomes, numerous predictors play a vital role in the long-term survival of patients. Finally, patients with vascular diseases are nowadays treated by several specialist disciplines using complementary conservative pharmacological and endovascular treatment as well as open surgery procedures. These factors significantly impact the measurement of all levels of quality and their validity and complicate the comprehensive collection of primary data in clinical registries for quality assurance [9].

Legally binding registries of external cross-sectoral quality assurance in Germany (§ 137a of Book V of the German Social Code [SGB V]) harbor advantages due to the potentially high external validity. They comprehensively cover almost all procedures performed, although only the respective inpatient stay. A few registries sometimes provide a longer follow-up and include subsequent hospital stays. As a rule, however, a hospital's participation in the registries of professional societies or research groups is voluntary, and the selection of cases and endpoints is not homogeneous, which reduces the internal and external validity [58]. This aspect may be less significant in highly centralised Scandinavian
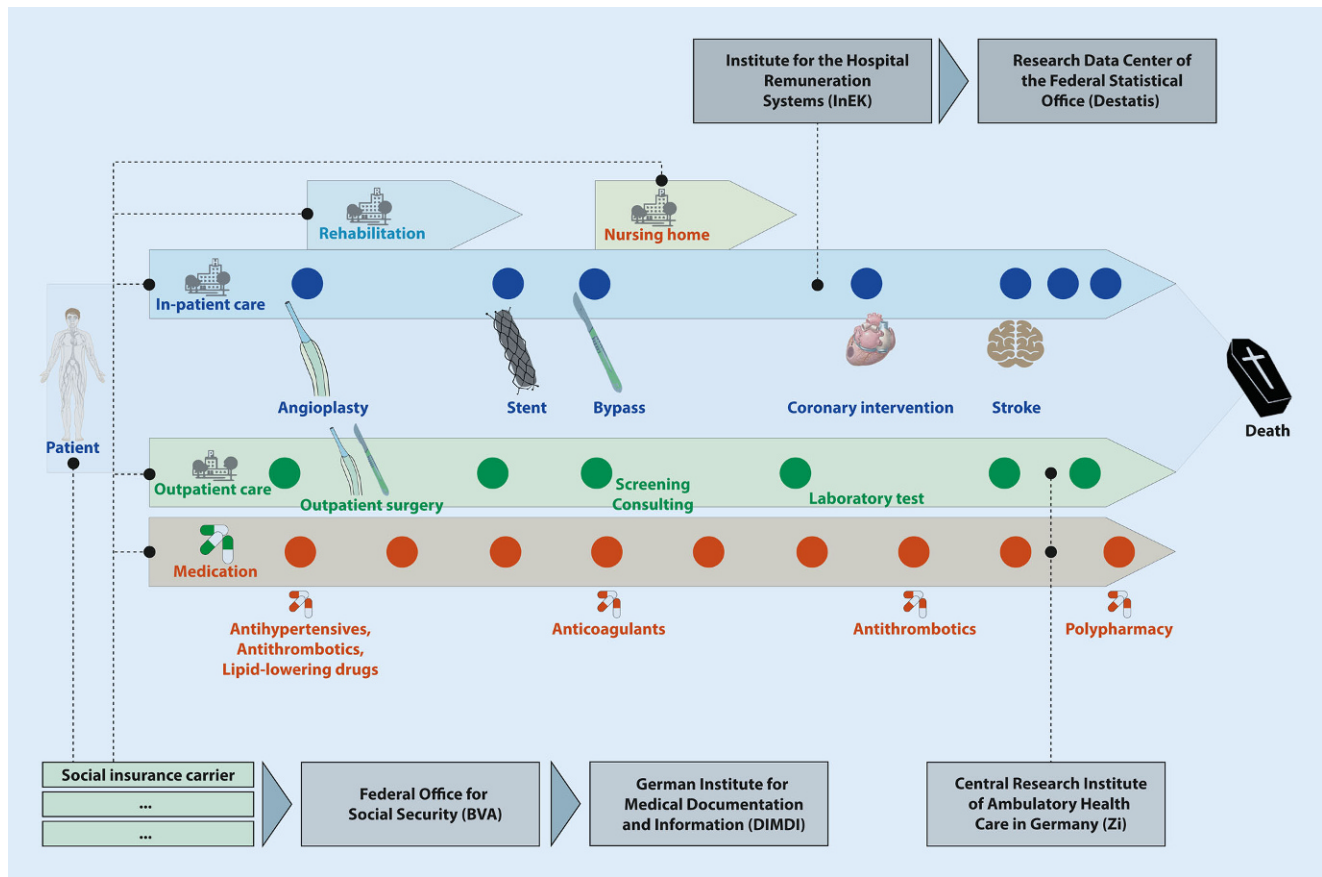
Fig. 1 ▲ Examples of data generation and data flow in cross-sectoral care for vascular diseases. A sensible link between all sectors and the long-term course is only possible through the social data of the social insurance carrier

countries than in Germany, where of the around 2000 hospitals up to 650 centers with different specialist departments provide vascular treatment. It is also possible to achieve an almost national survey of inpatient care using the data obtained pursuant to § 21 of the German Hospital Reimbursement Act (*Krankenhausentgeltgesetz* [KHEntgG]) from the Institute for the Hospital Remuneration System (Institut für das Entgeltsystem im Krankenhaus [InEK]), which ultimately also forms the basis for the diagnostic hospital statistics of the German Federal Statistical Office (Deutsches Statistisches Bundesamt [Destatis]). Here too, however, the observation period is restricted to the hospital stay. This limits the validity and clinical relevance of the findings and key conclusions. Furthermore, the lack of longitudinal linking between the cases prevents a patient-oriented approach. Depending on the reason for examination or index disease, this may ultimately make quality assurance impos-

sible. Although the interdisciplinary outpatient care of vascular patients also plays an important role in the long-term success of treatment, this aspect has not yet been adequately represented in all available data sources on inpatient care. Hence differences in outpatient secondary prevention are an important influencing factor in the assessment of the quality of revascularization.

Against this backdrop, suitably prepared routinely collected data from social insurance institutions provide a useful complement for quality improvement projects. Motivators inherent to the system (revenue generation) enable the timely and comprehensive generation of these data by the service providers without any limitation to individual specialist disciplines. Through a person-related longitudinal linkage of all available datasets, long-term observation of individual patients over periods of up to 15 years is theoretically possible, provided that the appropriate

anonymization using all of the currently established measures has been carried out beforehand (◘ **Fig. 1**; [11, 54]). Care in accordance with the guidelines with respect to inpatient treatment but also outpatient aftercare can therefore be adequately evaluated using routinely collected data from the social insurance institutions [37].

This review article examines the advantages and challenges of using routinely collected data for quality improvement projects in interdisciplinary vascular medicine. After reading this article, readers will be able to assess the different sources of routinely collected data and their potential usefulness for improving the quality of vascular treatment.

Routinely collected data in Germany

Routinely collected data is commonly associated with other terms, although on closer inspection the distinction is not

always clear cut. Thus, all data generated during routine care and their metadata can be considered routinely collected data [9, 50]. A number of research consortiums are currently working intensively on the harmonization and scientific use of these data [8, 22, 46]. The term secondary data derives from secondary purposes (e.g. quality assurance, research) beyond the primary purpose (mainly administration and accounting) [51]. Today, the debate on the evaluation and management of the healthcare system primarily centers on the data collected by social insurance institutions (◘ **Fig. 1**). Of particular importance and central to this article are the data from insured persons available here (§ 288 SGB V), along with the data on inpatient and outpatient care (§ 301, § 115 f. SGB V), contractual medical care (§ 295 SGB V), pharmaceutical billing (§ 300 SGB V), prevention and rehabilitation (§ 301 SGB V), incapacity to work (§ 295 SGB V), provision of therapeutic products (§ 302 SGB V), care for the chronically ill (§ 137f SGB V) and nursing care (§§ 36–38, § 41 SGB XI; § 37, § 43 SGB V). The data are also available at various institutions in the corresponding anonymized format for scientific use by authorized institutions. It is then no longer considered social data in the legal sense (◘ **Table 1**).

The scientific use of routinely collected data on inpatient care is most common, whereby the International Classification of Diseases (ICD) of the World Health Organization (WHO) valid in the reporting year is used to analyze primary and secondary diagnoses as well as operations and procedures. Deviating from this, the German uniform assessment standard (*Einheitlicher Bewertungsmaßstab* [EBM]) and the drug regulations via the anatomical therapeutic chemical (ATC) classification system are for example available for the outpatient sector in addition to the less comprehensive diagnosis coding. Despite the differences in diagnostic quality between sectors, a lack of better alternatives means that routinely collected data is frequently used.

The relevance of routinely collected data in the observation of population-related morbidity structures is thus demonstrated directly through the use of outpa-

F. Peters · T. Kreutzburg · J. Kuchenbecker · U. Marschall · M. Remmel · M. Dankhoff · H.-H. Trute · T. Repgen · E. S. Debus · C.-A. Behrendt

Quality of care in surgical/interventional vascular medicine: what can routinely collected data from the insurance companies achieve?

Abstract

The complexity and diversity of surgical/interventional vascular medicine necessitate innovative and pragmatic solutions for the valid measurement of the quality of care in the long term. The secondary utilization of routinely collected data from social insurance institutions has increasingly become the focus of interdisciplinary medicine over the years. Owing to their longitudinal linkage and pan-sector generation, routinely collected data make it possible to answer important questions and can complement quality development projects with primary registry data. Various guidelines exist for their usage, linkage, and reporting.

Studies have shown good validity, especially for endpoints with major clinical relevance. The numerous advantages of routinely collected data face several challenges that require thorough plausibility and validity procedures and distinctive methodological expertise. This review presents a discussion of these advantages and challenges and provides recommendations for starting to use this increasingly important source of data.

Keywords

Routinely collected data · Health research · Validity · Quality indicators · Administrative data

Behandlungsqualität in der operativ-interventionellen Gefäßmedizin – was können Routinedaten der Krankenkassen leisten? Englische Version

Zusammenfassung

Die Komplexität und Diversität der operativ-interventionellen Gefäßmedizin macht innovative und pragmatische Lösungsansätze zur validen Messung der langfristigen Behandlungsqualität erforderlich. Die sekundäre Nutzung von Routinedaten der Sozialversicherungsträger gerät dabei seit Jahren zunehmend in den Fokus der interdisziplinären Fachwelt. Routinedaten ermöglichen durch ihre longitudinale Verknüpfung und sektorenübergreifende Generierung die Beantwortung wichtiger Fragestellungen und können Qualitätsentwicklungsprojekte mit Primärdaten komplementär ergänzen. Es stehen verschiedene Leitlinien zu deren Nutzung, Verknüpfung und Berichterstattung zur Verfügung. Insbesondere bei Endpunkten

mit großer klinischer Relevanz wurde in Studien eine gute Validität nachgewiesen. Den vielen Vorteilen von Routinedaten stehen spezifische Herausforderungen gegenüber, die umfassende Plausibilitäts- und Validierungsverfahren und eine ausgeprägte Methodenkompetenz erfordern. Diese Übersichtsarbeit beschäftigt sich kritisch mit diesen Vorteilen und Herausforderungen und bietet Empfehlungen für den Einstieg in die Nutzung dieser zunehmend wichtigen Datenquelle.

Schlüsselwörter

Routinedaten · Versorgungsforschung · Validität · Qualitätsindikatoren · Administrative Daten

tient and inpatient information in Germany (◘ **Table 2**; [27]).

Consensus recommendations and guidelines on routinely collected data

In the international context, the validity and (international) comparability of routinely collected data has been the subject of controversial debate since it first began

to be used for scientific purposes [6, 20]. One main criticism is the limited comparability of the findings of routinely collected data studies due to the large number of classification systems used, with different versions and revisions as well as the project-specific selection of suitable inclusion criteria. With this in mind, uniform classifications of comorbidities and risk scores have been developed over the past 20 years to predict the severity

Institution and data owner	Characteristics
Scientific institutes of health insurance funds	E.g. BARMER science data warehouse (W-DWH), DAK-Gesundheit, AOK Research Institute (WIdO); cross-sectoral longitudinal and cross-sectional studies; limited to insured population
Federal Institute for Drugs and Medical Devices (DIMDI)	Partially limited datasets (e.g. no place of residence, admission date or quarterly pharmaceutical prescriptions)
Central Research Institute of Ambulatory Health Care in Germany (ZI)	Outpatient standard care and supply of pharmaceuticals; no data linkage to inpatient care
Institute for the Hospital Remuneration System (InEK)	Inpatient care data in accordance with § 21 KHEntgG; no data linkage
Federal Statistical Office (Destatis)	Microdata by the Research Date Center; hospital statistics

KHEntgG German Hospital Reimbursement Act

of diseases and in-hospital mortality [18, 41, 59].

Numerous guidelines and recommendations are currently also available from the German Society for Epidemiology (DGEpi) and the German Society for Social Medicine and Prevention (DGMSP) on the subject of routinely collected data. The Working Group for the Survey and Utilization of Secondary Data (AGENS) already published guidelines on secondary data analysis back in 2005; the publication is now in its third edition [50]. Against this backdrop, further guidelines developed since 2016 address the increasingly important subject of ensuring compliance with data protection regulations when linking datasets [34]. Finally, the standardized reporting of secondary data analyses (STROSA) reporting standard is available for Germany and also in the second edition [49].

Validation studies and transferability of findings

In addition to random sampling and additional needs-based reviews by the medical service of German statutory health insurance providers (Medizinischer Dienst der Krankenversicherung [MDK]), scientifically initiated validation studies with routinely collected data sources are available in Germany, for example, on mortality in the German pharmacoepidemiological research database (GePaRD) with high sensitivity (95.9%) and specificity (99.4%) [33] and on diagnoses (sensitivity to dementia 80%, heart failure 97% and tuberculosis 100%) [45]; however, outpatient diagnoses can have significantly lower sensitivity rates, which is attributed to the (quarterly) coding practices in Germany (sensitivity to tuberculosis 40%) [45].

In principal, validation should always be context-specific and take the external framework conditions of the target population into account [25]. A prerequisite for validation is a review of the plausibility and consistency of the data. This can for example include repeat diagnoses of chronic diseases and a comparison between drug and diagnostic data. It holds that greater severity of the endpoint is associated with higher validity, while less relevant observations have lower validity (inpatient sensitivity to hypertension 65%, cancer 91% and acute myocardial infarction 94% [44], outpatient sensitivity to back pain 74% and hypertension 81% [19]). Good validity can especially be assumed for mortality-related endpoints among the vascular patient cohort over the age of 65 years [15, 16, 25, 32, 38].

Adequate consideration of the diagnosis-free history (lookback) is recommended for incidence predictions. Czwickla et al. were able to show that by taking a lookback period of 7 years, the incidence of cancer is estimated to be about 10% lower (breast cancer 138.7 vs. 129.0 per 100,000 persons, prostate cancer 103.6 vs. 95.1 and colorectal cancer 42.1 vs. 38.3) [16].

For internal data validation, it can also be checked whether the same diagnosis codes predict comparable prevalences during different observational periods. It could for example be shown using routinely collected data from the BARMER health insurance provider that inclusion criteria for PAOD treatment and its clinical comorbidities as well as interventions correlated over a longer period [30]. Registry data can be used for the further validation of routinely collected data. One approach using model-based validation and a second approach using stratification-based validation have been developed within the IDOMENEO study funded by the German Federal Joint Committee (Gemeinsamer Bundesausschuss [G-BA]) [12]. In the model-based approach, the data is hierarchically ordered according to the hospital treating the patient, i.e. patients from one hospital form a cluster, both in the registry and in the routinely collected data from BARMER. Hospital-specific deviations can thus be taken into account and the multilevel models of the two data sources compared on the patient level. In the stratification-based approach, patients are divided into subgroups according to their characteristics, such as age, gender and the individual comorbidity profile. This them enables a comparison by means of subgroup analyses between the two data sources. The more similar the two data sources are, the higher the validity. Both approaches can ensure k-anonymity in compliance with data protection regulations [54].

Data protection and ethical considerations

The introduction of the General Data Protection Regulation (GDPR) of the European Union and its national implementation a few years ago led to the substantial restriction of healthcare research and quality improvement using real-world data sources. At the same time, however, legal grey areas that existed in the past have been replaced by

Indicator	Claims data of health insurance funds (e.g. BARMER, DAK-Gesundheit)	Inpatient care data in accordance with § 21 KHEntgG (InEK, Destatis)
Quality of results		
Mortality	High validity over entire duration of insurance up to 15 years (infrequent change of provider above the age of 65 years)	Discharge reason death limited to in-hospital deaths
Inpatient readmission	Possible with limitation in cause of readmission (e.g. emergency or identical main diagnosis as index case)	No data linkage possible (case-based analysis)
Reintervention	Possible for entire duration of insurance via procedure coding (limitations regarding side of body due to incomplete coding)	Limited to current hospital case
Amputation	Possible for entire duration of insurance via procedure coding (limitations regarding side of body due to incomplete coding)	Limited to current hospital case
Myocardial infarction	Possible via main or admission diagnosis or procedures	Limited to current hospital case
Stroke	Possible via main or admission diagnosis or procedures	Limited to current hospital case
Wound healing disorder	Possible via main or admission diagnosis or procedures	Limited to current hospital case
Acute limb ischemia	Possible via main or admission diagnosis or procedures	Limited to current hospital case
Bleeding complication	Possible via main or admission diagnosis or procedures	Limited to current hospital case
Decreasing kidney function and dialysis	Possible via main or comorbid diagnosis or procedures	Limited to current hospital case
Other levels of quality (quality of process and structure)		
Pharmaceutical prescription according to guidelines	All pharmaceutical prescriptions and filled prescriptions via ATC classification available quarterly	No outpatient pharmaceutical prescription data available
Diagnostics according to guidelines	Inpatient and outpatient care over entire duration of insurance	Limited to inpatient and reimbursed procedures in current hospital case
Case volume	Limited to population insured by provider	Complete data collection
Length of stay	Complete data collection	Complete data collection
Weekend effect	Complete data collection	Complete data collection
Adjusting factors and others		
Comorbidities	Previous outpatient and inpatient diagnoses and pharmaceutical prescriptions up to index case (e.g. prescription for insulin as indication of diabetes)	Limited to main and secondary diagnosis as well as current hospital case
Medication history prior to admission	Complete data collection (except for over-the-counter pharmaceuticals)	No outpatient pharmaceutical data available
Socioeconomic status	Limited. Can be inferred from co-payment exemptions or proxy variables	Not available in valid form
Diagnostic or screening prior to admission (outpatient)	If EBM-number is available (e.g. screening and consultation on abdominal aortic aneurysms)	No outpatient data available

ATC anatomical therapeutic chemical, *EBM* German uniform assessment standard, *InEK* Institute for the Hospital Remuneration System, *KHEntgG* German Hospital Reimbursement Act

legally verifiable regulations. The GDPR places high demands on the anonymization of data, as the purely hypothetical linking of datasets or individual characteristics is already explicitly described as a criterion for personal data. To ensure that the social data collected by social insurance institutions in compliance with the law can also be used by researchers for the purposes of quality improvement and healthcare research, the data owners perform comprehensive processing or de facto anonymization. Through appropriate aggregations, censoring and logically meaningful date shifts, these data are altered in such a way that they are no longer considered personal or social data but still allow valid evaluations. One alternative is to obtain the explicit informed consent of all insured persons; however, this would create a disproportionate amount of work for many projects. With respect to ethical considerations, the interdisciplinary good practice in secondary data analysis clearly states in its first guidelines that although secondary data analyses must be carried out in accordance with ethical principles, an ethics committee only needs to be consulted in individual cases [50].

Methodological considerations

From a methodological, statistical perspective, routinely collected data represent a particularly interesting source of secondary data and an important basis for health services research. Similar to information from vital statistics, these data contain relevant information on the entire target population rather than on just a random sample [35]. Another statistically interesting aspect is the process of data generation, which is standardized and independent of possible scientific questions. Both the recording of the population and the availability of data entail methodological necessities and must be taken into account in the research design accordingly. Since routinely collected data do not represent a random sample, inferential statistical methods and especially hypothesis testing with associated *p*-values play a far lesser role here than they do with primary data [57]. This is further emphasized by the high statis-

Table 3	Strengths and limitations of selected indicator of quality in various data sources	
Parameter	**Randomized control trial**	**Routine data**
Coverage of the population	Low, often only high-volume centers included and small selective samples	High, all service providers included but often only access to a single health insurance fund
Diversity of the population	Exceptionally low because study population tailored to a specific disease and intervention. Focus on typical cases	Exceedingly high because full real-world representation of provision of health care
Internal validity	Medium to high, due to randomization blinding and controlled intervention. But large variation of the quality of a specific trial CONSORT statement	Low to medium and dependent on methodology for adjustment for bias and validation of variables used for the analysis STROBE statement
External validity	Low, due to highly selected population	High, because data collected during routine care
Laboratory parameters	Partly included. Biobanking enables additional biomarker-based and genetic analyses	Usually not included. Could be approximated partly via ICD codes (hypertension, hypercholesterolemia)
Number of cases	Low. Usually minimum size for detecting treatment effects. Underpowered for (rare) long-term outcomes	Exceedingly high. Challenges arising from potential false positive findings. Sensitivity analyses necessary
Adverse events	High. Recorded according to protocol	Available for a long follow-up duration but dependent on quality of documentation (outpatient, inpatient)
Costs	High, additional effort for screening, recruitment, conduct and documentation	Low, no additional effort for documentation
Availability	Low. Data mostly analyzed within project partners and not generally publicly available	High. Simpler since innovation funding, consortium projects established
Estimation of prevalence	Yes	Yes
Estimation of incidence	Barely	Yes, if patient history is assessed
Follow-up	Mostly short follow-up duration	Many years of follow-up without additional effort of documentation
Cross-sectoral	Rarely	Yes, outpatient and inpatient data could be linked
Selection and recall bias?	Yes. Selection bias due to hospital and specifics of patients	Barely
Assessment of comparative effectiveness?	No	Yes
Suitable for quality control/ benchmarking?	Less suitable	Well suited for parameters like mortality, diagnoses and procedures
Assessment of causality?	Yes	Partly, rather correlations and associations. Quasi-experimental approaches possible
Patient-reported outcomes, lifestyle variables	Yes, depending on research question	No, only available via proxy variables
Unspecific diagnoses	Barely	Yes, especially in outpatient sector

CONSORT consolidated standards of reporting trials, *STROBE* strengthening the reporting of observational studies in epidemiology, *ICD* International Classification of Diseases

tical power due to the large number of cases and the possibility of multiple testing, which leaves analyses prone to false positive outcomes. For these reasons, standardized differences combined with considerations of medical relevance are better suited to the assessment of differences than the consideration of *p*-values alone. This applies to the assessment of gender disparities in secondary prevention following inpatient PAOD hospitalization [40]. Likewise, absolute differences should also always be considered in addition to any potentially statistically significant relative differences with respect to indicators of outcome quality.

Since routinely collected data are already available when new studies are designed, a deductive approach is not possible without restrictions and variables are measured based on the existing database in addition to prior theoretical considerations [52]. As a consequence, required target variables (e.g. laboratory parameters or patient-reported endpoints) might not be available or at least not in the desired format. Proxy variables must be used in this case. One example is the measurement of smoking status using the ICD-10 code F17 (mental and behavioral disorders due to use of tobacco) as a proxy for nicotine consumption [43]. This also applies to complex diagnoses and procedures, which must be compiled individually from a multitude of billing codes in order to approximate the target intervention as accurately as possible. This raises the potential statistical problem of omitted variable bias. Further bias arises from the lack of randomization of the data in relation to the primary hypothesis, which can lead to unmeasured confounding.

Critics of routinely collected data analyses, big data and data-driven research often generally doubt the potential of this research area to provide statements on relevant issues [2]; however, this perspective ignores the fact that findings from secondary data analyses do not intend to replace evidence from randomized trials but should rather be seen as complementary to this (◘ Table 3). For example, the idea, development, exploration, assessment, and long-term follow-up (IDEAL) framework even explicitly emphasizes the

complementary function of routinely collected data in the review of the surgical efficacy of interventions (especially for rare events and long-term outcomes) [47]. During pragmatic trials, the aim is increasingly to also use intervention studies in everyday care or to link routinely collected data with primary surveys via data linkage [53]. Finally, quasiexperimental study designs are used in billing data to approximate causal statements [62]. There is a growing realization that different types of data and analysis can each supply their own evidence building blocks whose synthesis can generate further knowledge.

The recent global debate on the safety of drug-coated stents and balloons in femoropopliteal arteries is one example of the interplay of different types of data and analysis strategies [10]. In December 2018, safety concerns were published regarding long-term survival following the use of paclitaxel-coated medical devices in the femoral region [28]. These findings led to a global response, with recruitment for studies halted and usage warnings issued by national authorities. Neither clinical studies nor registry data could subsequently be used to validate this initial suspicion as long-term data of a sufficient quantity and quality were not available. Drawing on routinely collected data from BARMER, two comprehensive analyses of routinely collected data were conducted in parallel and independently of each other, taking very different methodological approaches. On the one hand, propensity score matching was implemented to compensate for the lack of randomization at the time of the intervention [13]. On the other hand, information on multiple interventions per patient was used to model a possible dose-response relationship [21]. Despite the different methodologies, both studies came to the conclusion that in the unselected real-world samples, the suspected safety concerns ultimately did not arise. The same discrepancy between secondary evaluations of intervention studies and routinely collected data has recently been observed regarding the use of paclitaxel coating in the arteries of the lower leg [24, 29].

Risk prediction with routinely collected data

Due to their longitudinal data structure, abundance of variables and large number of cases, routinely collected data also provides numerous opportunities for long-term risk predictions. Within the RABATT project funded by the Federal Joint Committee (G-BA), a learning risk score was recently developed for 5-year amputation-free survival following invasive revascularization of symptomatic PAOD [46]. This score was calculated using routinely collected data from BARMER and validated using GermanVasc registry data. Machine learning and regularization procedures, such as the LASSO selection procedure with cross-validation were used to identify variables with a high association with survival [56]. With good discrimination, predictors could be identified that enable a pragmatic classification of patients into low-risk to high-risk groups. In this application, such data-based machine learning procedures may of course also only be considered complementary to existing evidence and must always be viewed in the respective medical context in order to prevent spurious correlation [31]. It is therefore particularly important to also pursue external validation and continuous development when using prognosis models.

Discussion

In recent decades, various factors including the digital revolution and the introduction of the DRG in global health systems have led to a significant increase in the scientific use of routinely collected data. The growing amount of clinical documentation that doctors and nursing staff must complete and the commonly known limitations of primary data from registries make this source of data a valuable complement to quality improvement [3, 42]. Suitable selection criteria in the available datasets not only enable the measurement of process and structural quality indicators, but also an analysis of outcome quality indicators in the long term. Here, the particular challenge lies in the selection of suitable criteria and their (inter)national comparability.

Critics of secondary data usage, not just routinely collected data, reiterate the insufficient internal validity as the purpose of data generation (billing, revenue generation) leads to systematic bias. Indeed, such coding effects must be assessed on a project-specific basis and adequately taken into account. The fact is that revenue incentives have a measurable impact on patient selection and the choice of treatment [4, 61]. While routinely collected data is at least subject to regular independent review by the MDK, virtually no external validation has taken place so far for primary data in registries, which limits its usefulness [60].

Given that the statutory and private health insurance provider market remains highly fragmented despite market adjustment, population-specific selection effects arise from their different morbidity structures [26]; however, this mainly relates to the representativeness of the data and less to its contribution to quality improvement. This again underlines the relevance of sound methods, such as adjustment and standardization.

The hypothetical disadvantages of routinely collected data are also outweighed by significant advantages. While strict inclusion and exclusion criteria with corresponding data validation in randomized controlled trials (RCT) result in a higher internal validity, the routinely collected data of social insurance institutions reflects the reality of care far better and thus result in higher external validity. Due to the large number of cases (several million patients), it is also possible to analyze symptoms and patient groups that it is difficult for primary surveys to access. This particularly applies for severely ill multimorbid patients and residents of nursing homes, which underlines the association with vascular medicine. Compared to registries and RCTs, routinely collected data continue to exhibit a particularly high level of completeness, both at the start of the study and during the recording of study endpoints. Moreover, the data are available comparatively quickly and at low costs [36]. Similarly, automated billing data do not experience problems caused by a lack of response or incorrect patient information. This is

particularly advantageous in studies involving pharmacological treatment over a longer period [39]. Even if routinely collected data do not contain certain information, such as patient-reported endpoints and lifestyle characteristics, newer approaches provide the possibility of identifying relevant proxy variables from the tremendous wealth of information in a data-based manner using machine learning algorithms [55]. In the near future an abundance of further applications for new analytical methods will arise in this research area. These must then be monitored critically by experienced clinicians.

Countless publications demonstrate the benefits of routinely collected data for quality improvement in vascular medicine. In current routinely collected data studies conducted by BARMER, indications of potential for improvement could be found in the provision of drugs. Only about 60% of all patients with invasive revascularization of symptomatic PAOD received statins after the inpatient stay despite the fact that current guidelines clearly recommend this [1]. The challenge is now to transform these findings into better care [40]. As part of a quality initiative between the GermanVasc research group and DAK-Gesundheit health insurance provider on the treatment of aortic aneurysms, established outcome quality indicators for the weekend treatment of emergencies [14], severe bleeding complications [7] and spinal cord ischemia [23] could be evaluated. The outcomes of this work support the current debate on the necessary centralization of vascular medicine services and lead to further studies, for example on the introduction of patient blood management and CSF drainage in complex endovascular aortic surgery.

Meanwhile, the international vascular medicine community has already recognized the complementary benefits of routinely collected data. Hence cooperation, such as the VASCUNET committee (www.vascunet.org) and the medical device epidemiology network (MDEpiNet; www.mdepinet.org) are already working intensively on methodological aspects and the enhanced harmonization of routinely collected data research

[48]. In the current VASCUNET quality report on the interdisciplinary care of PAOD in 11 countries, for example, routinely collected data have been received for more than 1.1 million hospital cases. The report is expected to lead to extensive debates on quality and programs in the participating countries.

It is clear that the findings from routinely collected data studies are not without repercussions for the doctor-patient relationship. The benchmarks of medical standards cannot ignore the knowledge acquired through such long-term studies without change. In itself this is not unusual; however, this legal aspect gains new weight when routinely collected data are combined with machine learning to develop algorithm-based forecasting tools. As already impressively demonstrated in the data protection considerations, this means that one is always moving within legally relevant areas and this must be kept in mind. Against this backdrop, it would appear crucial for medical and legal practitioners to face these new challenges together and in synergy, and to develop solutions to overcome them.

To conclude, it is important and correct to carefully weigh up the advantages and limitations of routinely collected data from health insurance providers and their importance in quality improvement. Challenges should not lead to a general rejection of this data basis but rather prompt a constructive exchange and methodological improvement.

Only through continuous usage and a controversial discussion will it be possible to improve the quality of routinely collected data research and thus also the quality of patient care.

Conclusion

- The use of routine data from social security institutions for secondary purposes in health services research and quality development has significantly increased in recent decades and a further increase is foreseeable.
- There are interdisciplinary and generally accepted guidelines and consensus recommendations for their use, linkage and reporting.

- Working with routine data requires not only medical but also a strong methodological expertise and can certainly provide high-quality complementary evidence for interventional studies.
- A representation of healthcare provision in longitudinal and cross-sectional perspectives is only possible via the social insurance institutions or in correspondingly merged data records of other data owners.
- The case-related data records of the Institute for the Remuneration System in Hospitals, the German Federal Statistical Office and the Central Institute for Statutory Health Insurance in Germany offer helpful insights into provision of healthcare, but are restricted to the assessment of the quality of the results due to the lack of longitudinal linkage.
- Depending on the research question, the restriction of the insured population to a single health insurance fund could potentially be a limitation for epidemiological questions, which requires appropriate methodological measures (e.g. standardization).
- Numerous routine data studies have already been published that have proven the benefits for quality development in interdisciplinary vascular medicine.
- Further context-specific validation studies for the German healthcare situation and methods for data protection-compliant validation of routine data are necessary.

Corresponding address

Dr. C.-A. Behrendt
Forschungsgruppe GermanVasc, Klinik und Poliklinik für Gefäßmedizin, Universitäres Herz- und Gefäßzentrum UKE Hamburg, Universitätsklinikum Hamburg-Eppendorf Martinistr. 52, 20246 Hamburg, Germany
behrendt@hamburg.de

Compliance with ethical guidelines

References

1. Aboyans V, Ricco JB, Bartelink MEL et al (2018) Editor's choice—2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European society for vascular surgery (ESVS). Eur J Vasc Endovasc Surg 55:305–368
2. Antes G (2016) Ist das Zeitalter der Kausalität vorbei? Z Evid Fortbild Qual Gesundhwes 112(1):S16–22
3. Bates DW, Gawande AA (2003) Improving safety with information technology. N Engl J Med 348:2526–2534
4. Beck AW, Sedrakyan A, Mao J et al (2016) Variations in abdominal aortic aneurysm care: a report from the international consortium of vascular registries. Circulation 134:1948–1958
5. Behrendt CA (2020) Routinely collected data from health insurance claims and electronic health records in vascular research—a success story and way to go. Vasa 49:85–86
6. Behrendt CA, Debus ES, Mani K et al (2018) The strengths and limitations of claims based research in countries with fee for service reimbursement. Eur J Vasc Endovasc Surg 56:615–616
7. Behrendt CA, Debus ES, Schwaneberg T et al (2020) Predictors of bleeding or anemia requiring transfusion in complex endovascular aortic repair and its impact on outcomes in health insurance claims. J Vasc Surg 71:382–389
8. Behrendt CA, Härter M, Kriston L et al (2017) IDOMENEO – Ist die Versorgungsrealität in der Gefäßmedizin Leitlinien- und Versorgungsgerecht? Gefasschirurgie 22:41–47
9. Behrendt CA, Heidemann F, Riess HC et al (2017) Registry and health insurance claims data in vascular research and quality improvement. Vasa 46:11–15
10. Behrendt CA, Peters F, Mani K (2020) The swinging pendulum of evidence—is there a reality behind results from randomized trials and real-world data? Lessons learned from the paclitaxel debate. Eur J Vasc Endovasc Surg. https://doi.org/10.1016/j.ejvs.2020.01.029
11. Behrendt CA, Pridohl H, Schaar K et al (2017) Clinical registers in the twenty-first century: balancing act between data protection and feasibility? Chirurg 88:944–949
12. Behrendt CA, Schwaneberg T, Hischke S et al (2020) Data privacy compliant validation of health insurance claims data: the IDOMENEO approach. Gesundheitswesen 82(S 02):S94–S100. https://doi.org/10.1055/a-0883-5098
13. Behrendt CA, Sedrakyan A, Peters F et al (2020) Editor's choice—long term survival after femoropopliteal artery revascularisation with paclitaxel coated devices: a propensity score matched cohort analysis. Eur J Vasc Endovasc Surg 59:587–596
14. Behrendt CA, Sedrakyan A, Schwaneberg T et al (2019) Impact of weekend treatment on short-term and long-term survival after urgent repair of ruptured aortic aneurysms in Germany. J Vasc Surg 69:792–799.e2
15. Czwikla J, Domhoff D, Giersiepen K (2016) ICD coding quality for outpatient cancer diagnoses in SHI claims data. Z Evid Fortbild Qual Gesundhwes 118–119:48–55
16. Czwikla J, Jobski K, Schink T (2017) The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data. BMC Med Res Methodol 17:122
17. Donabedian A (2005) Evaluating the quality of medical care. Milbank Q 83:691–729
18. Elixhauser A, Steiner C, Harris DR et al (1998) Comorbidity measures for use with administrative data. Med Care 36:8–27
19. Erler A, Beyer M, Muth C et al (2009) Garbage in – Garbage out? Validität von Abrechnungsdiagnosen in hausärztlichen Praxen. Gesundheitswesen 71:823–831
20. Ferber LV, Behrens J (1997) Public Health – Forschung mit Gesundheits- und Sozialdaten: Stand und Perspektiven. Memorandum zur Analyse und Nutzung von Gesundheits- und Sozialdaten
21. Freisinger E, Koeppe J, Gerss J et al (2019) Mortality after use of paclitaxel-based devices in peripheral arteries: a real-world safety analysis. Eur Heart J. https://doi.org/10.1093/eurheartj/ehz698
22. Gehring S, Eulenfeld R (2018) German medical informatics initiative: unlocking data for research and health care. Methods Inf Med 57:e46–e49
23. Heidemann F, Kolbel T, Kuchenbecker J et al (2020) Incidence, predictors, and outcomes of spinal cord ischemia in elective complex endovascular aortic repair: an analysis of health insurance claims. J Vasc Surg. https://doi.org/10.1016/j.jvs.2019.10.095
24. Heidemann F, Peters F, Kuchenbecker J et al (2020) Long-term outcomes after revascularizations below the knee with paclitaxel-coated devices—a propensity score matched cohort analysis. Eur J Vasc Endovasc Surg (In Press)
25. Hoffmann F, Andersohn F, Giersiepen K et al (2008) Validation of secondary data. Strengths and limitations. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 51:1118–1126
26. Hoffmann F, Icks A (2012) Structural differences between health insurance funds and their impact on health services research: results from the Bertelsmann health-care monitor. Gesundheitswesen 74:291–297
27. Ingenerf J (2007) Gesetzliche Krankenversicherung: Der Morbi-RSA soll den Wettbewerb um Gesunde beenden. Dtsch Arztebl Int 104:1564
28. Katsanos K, Spiliopoulos S, Kitrou P et al (2018) Risk of death following application of paclitaxel-coated balloons and stents in the femoropopliteal artery of the leg: a systematic review and meta-analysis of randomized controlled trials. J Am Heart Assoc 7:e11245
29. Katsanos K, Spiliopoulos S, Kitrou P et al (2020) Risk of death and amputation with use of paclitaxel-coated balloons in the infrapopliteal arteries for treatment of critical limb ischemia: a systematic review and meta-analysis of randomized controlled trials. J Vasc Interv Radiol 31:202–212
30. Kreutzburg T, Peters F, Riess HC et al (2020) Editor's choice—comorbidity patterns among patients with peripheral arterial occlusive disease in Germany: a trend analysis of health insurance claims data. Eur J Vasc Endovasc Surg 59:59–66
31. Kriston L (2020) Machine learning's feet of clay. J Eval Clin Pract 26:373–375
32. Langner I, Ohlmeier C, Haug U et al (2019) Implementation of an algorithm for the identification of breast cancer deaths in German health insurance claims data: a validation study based on a record linkage with administrative mortality data. BMJ Open 9:e26834
33. Langner I, Ohlmeier C, Zeeb H et al (2019) Individual mortality information in the German pharmacoepidemiological research database (GepaRD): a validation study using a record linkage with a large cancer registry. BMJ Open 9:e28223
34. March S, Andrich S, Drepper J et al (2019) Gute Praxis Datenlinkage (GPD). Gesundheitswesen 81:636–650
35. Mueller U (2000) Die Maßzahlen der Bevölkerungsstatistik. In: Mueller U, Nauck B, Diekmann A (eds) Handbuch der Demographie 1. Springer, Berlin, Heidelberg, pp 1–91
36. Neubauer S, Schilling T, Zeidler J et al (2016) Auswirkung einer leitliniengerechten Behandlung auf die Mortalität bei Linksherzinsuffizienz. Herz 41:614–624
37. Neubauer S, Zeidler J, Schilling T et al (2016) Eignung und Anwendung von GKV-Routinedaten zur Überprüfung von Versorgungsleitlinien am Beispiel der Indikation Linksherzinsuffizienz. Gesundheitswesen 78:e135–e144
38. Ohlmeier C, Langner I, Hillebrand K et al (2015) Mortality in the German pharmacoepidemiological research database (GepaRD) compared to national data in Germany: results from a validation study. BMC Public Health 15:570
39. Peters F, Kreutzburg T, Kuchenbecker J et al (2020) A retrospective cohort study on the provision and outcomes of pharmacological therapy after revascularization for peripheral arterial occlusive disease: a study protocol. BMJ Surg Interv Health Technol. https://doi.org/10.1136/bmjsit-2019-000020
40. Peters F, Kreutzburg T, Riess H et al (2020) Optimal pharmacological treatment of symptomatic peripheral arterial occlusive disease and evidence of female patient disadvantage: an analysis of

health insurance claims data. Eur J Vasc Endovasc Surg (In Press)

41. Quan H, Sundararajan V, Halfon P et al (2005) Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 43:1130–1139

42. Quiroz JC, Laranjo L, Kocaballi AB et al (2019) Challenges of developing a digital scribe to reduce clinical documentation burden. NPJ Digit Med 2:114

43. Reinecke H, Unrath M, Freisinger E et al (2015) Peripheral arterial disease and critical limb ischaemia: still poor outcomes and lack of guideline adherence. Eur Heart J 36:932–938a

44. Schneeweiss S, Avorn J (2005) A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 58:323–337

45. Schubert I, Ihle P, Köster I (2010) Interne Validierung von Diagnosen in GKV-Routinedaten: Konzeption mit Beispielen und Falldefinition. Gesundheitswesen 72:316–322

46. Schwaneberg T, Debus ES, Repgen T et al (2019) Entwicklung eines selbstlernenden Risikoscores an Real-World-Datenquellen. Gefässchirurgie 24:234–238

47. Sedrakyan A, Campbell B, Merino JG et al (2016) IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. BMJ 353:i2372

48. Sedrakyan A, Cronenwett JL, Venermo M et al (2017) An international vascular registry infrastructure for medical device evaluation and surveillance. J Vasc Surg 65:1220–1222

49. Swart E, Bitzer E, Gothe H et al (2016) STandardisierte BerichtsROutine für Sekundärdaten Analysen (STROSA) – ein konsentierter Berichtsstandard für Deutschland, Version 2. Gesundheitswesen 78:e145–e160

50. Swart E, Gothe H, Geyer S et al (2015) Good practice of secondary data analysis (GPS): guidelines and recommendations. Gesundheitswesen 77:120–126

51. Swart E, Heller G (2007) Nutzung und Bedeutung von (GKV-) Routinedaten für die Versorgungsforschung. Theoretische Ansätze, Methoden, Instrumente und empirische Befunde. In: Janßen C, Borgetto B, Heller G (eds) Medizinsoziologische Versorgungsforschung. Juventa, Weinheim, pp 93–112

52. Swart E, Ihle P, Gothe H (2014) Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden, und Perspektiven. Huber, Bern

53. Swart E, Stallmann C, Powietzka J et al (2014) Datenlinkage von Primär-und Sekundärdaten. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 57:180–187

54. Sweeney L (2012) k-anonymity: a model for protecting privacy. Int J Uncertainty, Fuzziness Knowledge-Based Syst 10:557–570

55. Thesmar D, Sraer D, Pinheiro L et al (2019) Combining the power of artificial intelligence with the richness of healthcare claims data: opportunities and challenges. PharmacoEconomics 37:745–752

56. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc Ser A 58:267–288

57. Tijssen JG, Kolm P (2016) Demystifying the new statistical recommendations: the use and reporting of p values. J Am Coll Cardiol 68:231–233

58. Trenner M, Eckstein HH, Kallmayer MA et al (2019) Secondary analysis of statutorily collected routine data. Gefässchirurgie 24:220–227

59. van Walraven C, Austin PC, Jennings A et al (2009) A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 47:626–633

60. Venermo M, Mani K, Kolh P (2017) The quality of a registry based study depends on the quality of the data—without validation, it is questionable. Eur J Vasc Endovasc Surg 53:611–612

61. Venermo M, Wang G, Sedrakyan A et al (2017) Editor's choice—carotid stenosis treatment: variation in international practice patterns. Eur J Vasc Endovasc Surg 53:511–519

62. Zhang X, Faries DE, Li H et al (2018) Addressing unmeasured confounding in comparative observational research. Pharmacoepidemiol Drug Saf 27:373–382