International Journal of Technology Assessment in Health Care, 32:1/2 (2016), 69–77.

doi:10.1017/S0266462316000131

USING CLAIMS DATA FOR EVIDENCE GENERATION IN MANAGED ENTRY AGREEMENTS

Alina Brandes, Larissa Schwarzkopf
Helmholtz Zentrum München; German Research Center for Environmental Health (GmbH), Institute of Health Economics and Health Care Management

Wolf H. Rogowski
Department of Health Care Management, Institute of Public Health and Nursing Research, Health Sciences, University of Bremen, Germany; Helmholtz Zentrum München; German Research Center for Environmental Health (GmbH), Institute of Health Economics and Health Care Management
rogowski@uni-bremen.de

Objectives: This study assesses the use of routinely collected claims data for managed entry agreements (MEA) in the illustrative context of German statutory health insurance (SHI) funds.
Methods: Based on a nonsystematic literature review, the data needs of different MEA were identified. A value-based typology to classify MEA on the basis of these data needs was developed. The typology is oriented toward health outcomes and utilization and costs, key components of a new technology's value. For each MEA type, the suitability of claims data in establishing evidence of the novel technology's value in routine care was systematically assessed. Assessment criteria were data availability, completeness, timeliness, confidentiality, reliability, and validity.
Results: Claims data are better suited to MEA addressing uncertainty regarding the utilization and costs of a novel technology in routine care. In schemes where safety aspects or clinical effectiveness are assessed, the role of claims data is limited because clinical information is not included in sufficient detail.
Conclusions: The suitability of claims data depends on the source of uncertainty and, in consequence, the outcome measures chosen in the agreements. In all schemes, the validity of claims data should be judged with caution as data are collected for billing purposes. This framework may support manufacturers and payers in selecting the most suitable contract type and agreeing on contract conditions. More research is necessary to validate these results and to address remaining medical, economic, legal, and ethical questions of using claims data for MEA.

Keywords: Technological innovation, Decision making, Insurance, health, reimbursement, Risk sharing, financial/economics, Data collection/methods

Effectiveness and costs, key components of a new medical technology's value, are usually uncertain at the point of coverage decision making. This is partly because coverage decision makers ask for different types of evidence than market approval agencies. Also, they may be subject to information asymmetry as the manufacturer may hold hidden information regarding the technology's characteristics and intended use that cannot be detected by the payer before market entry (1). Under such uncertainty, the decision maker may, on the one hand, reject coverage and deny patients access to beneficial technologies. A positive coverage decision, on the other hand, may lead to reimbursement of ineffective medical technologies and strain already tight budgets.

To address this uncertainty, market entry can be accompanied by arrangements whereby evidence regarding the performance and utilization of the technology is collected alongside use in clinical practice. That way, the final coverage decision can be delayed until sufficient evidence is available, while enabling early patient access to novel technologies. Also, the level of reimbursement can be conditioned on the performance of the novel technology, so that the financial risk is shared between payers and manufacturers (2). These schemes are described using differing terms in the literature. In the following, the term *managed entry agreements* (MEA) is used for all forms of contracts between manufacturers and payers where coverage is tied to the collection of further evidence to address uncertainty regarding the novel health technology's value during its introduction. MEA have received increasing attention in the scientific literature (3–5).

MEA are also of importance in the fifth German Social Code Book (Sozialgesetzbuch, SGB V), which largely determines coverage and reimbursement by the German statutory health insurance (SHI). Since the introduction of the Healthcare Provision Act (GKV-Versorgungsstrukturgesetz) in 2011, it is possible for the SHI to cover novel examination and treatment methods on the condition that evidence is collected alongside use in clinical practice (§137e SGB V) (6). This is a similar concept to coverage with evidence development (CED), a form of MEA conducted by the National Health Service in the UK and Centers for Medicare and Medicaid Services in the United States (7;8). Furthermore, in Germany, individual SHI funds may engage in selective contracts with healthcare providers

The research is carried out on behalf of the Helmholtz Zentrum München. The Helmholtz Zentrum München is an independent organization funded by the German and Bavarian governments. Research for this study was supported by a grant from the German Federal Ministry of Education and Research (grant number 13KQ0902V).

within the scope of integrated care or model schemes. These schemes are often risk-sharing agreements (RSA), which are also a form of MEA (5).

The success of a MEA relies on the collection of valid evidence on the performance and utilization of the novel technology in clinical practice; the question arises where to find this evidence. Original data collected within randomized controlled trials are seen as the gold standard for evidence collection. However, their ability to demonstrate (cost-)effectiveness in clinical practice is limited, for example, because of the short time horizons of many studies or because the settings in clinical trials do not necessarily correspond with clinical practice. Original data collected within registries may overcome some of these problems, but involve their own distinct limitations (9). In both cases, data collection is likely to incur high costs so that the costs of evidence collection can easily exceed the value associated with the additional evidence.

An alternative is routinely collected claims data which are readily available from insurance funds. Even if they are subject to several limitations (10;11), they might be a highly valuable data source because they are less costly and they provide information about the costs and effects associated with new technologies in clinical practice. Paragraphs §291, §295, §300, §301, and §302 of SGB V specify the information that is routinely collected within German SHI claims data. This includes, first, data that have to be transferred to SHI funds for the reimbursement of health care in various healthcare sectors (inpatient care, outpatient care, pharmaceuticals, physiotherapy, assistive technologies, and rehabilitative care). Second, it includes basic claims data on insurant characteristics that SHI funds may collect such as age and gender.

Using data collected by German SHI funds for illustration, this study assesses to what extent the claims data of health insurance funds are suitable sources of evidence for different types of MEA. The use of MEA in Germany is still novel, and there is no publicly accessible data source of existing contracts and the data they incorporate. Therefore, this assessment could not be based on either systematic literature reviews or empirical analysis of data from existing MEA. Instead, it is based on three steps: first, data needs for the different types of MEA are extracted from existing reviews; second, a set of criteria for assessing the suitability of claims data to meet these data needs was developed, which is reported in the Methods section; third, the suitability of German SHI for MEA was assessed on the basis of this set of criteria. Following this assessment, implications for current contracting practice and future applications in Germany are deduced, and the generalizability of these implications for other healthcare systems is discussed.

METHODS

To identify published literature describing types of MEA and their respective data requirements, a qualitative, nonsystematic literature review was conducted in Medline (for details, see Supplementary Table 1). Given that several recent reviews were identified, the first step in this study was mainly based on an analysis of these reviews. Studies were identified and data needs for different MEA were extracted by AB. Based on the identified studies and these data needs, a typology to present and classify the different types of MEA was developed inductively by A.B. and W.R.

To assess the suitability of claims data to meet these data needs, the following criteria were developed by A.B. and W.R. on the basis of the same literature search.

Availability: The availability of relevant pieces of information in claims data is one obvious criterion for their suitability to establish MEA. For example, if the value proposition of a new technology is to reduce a certain clinical event such as incontinence in prostate cancer care, this clinical endpoint needs to be included in the data to be assessed within a MEA. To assess availability, the information contained in SHI claims data according to §§291, 295, 300, 301, and 302 in SGB V were searched to detect information on the outcomes used in the MEA schemes.

Completeness: The relevant piece of information needs to be included completely, i.e. available for all patients and providers. This would not be the case, for example, if both urologists and general practitioners could treat incontinence, but data on incontinence treatment were available only for one group of physicians. To assess completeness, the data descriptions were searched for descriptions of relevant exclusions. Furthermore, it was assessed whether the data were available for all types of insurees (insurance holder and co-insured family member), types of healthcare providers (general practitioners and specialists in ambulatory care, inpatient care), and funding schemes (ambulatory reimbursement and inpatient DRGs).

Timeliness: Given that the generation of routine data may be time consuming, availability and completeness need to be complemented by timeliness to facilitate the use of this information within MEA (12). To assess timeliness, information regarding the reporting and reimbursement process was searched for relevant delays between the time of an event and the time information about the event is available in claims data.

Confidentiality: The information must not be subject to confidentiality, which is an important consideration in the use of social health insurance data (13). To detect limitations in the use of the data for confidentiality reasons, the relevant paragraphs in the Social Code Book (§§287, 303a–f SGB V, §75 SGB X) as well as the German privacy law ("Datenschutzgesetz", §§3, 3a, 4, 4a, 11, 13, 14) were searched for rules that set legal constraints on availability.

Reliability and Validity: Finally, standard criteria for using data in empirical research are reliability and validity. This would not be the case, for example, if diagnostic coding largely depended on the habits and diligence of different physicians. To assess this criterion, a second nonsystematic literature search was

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

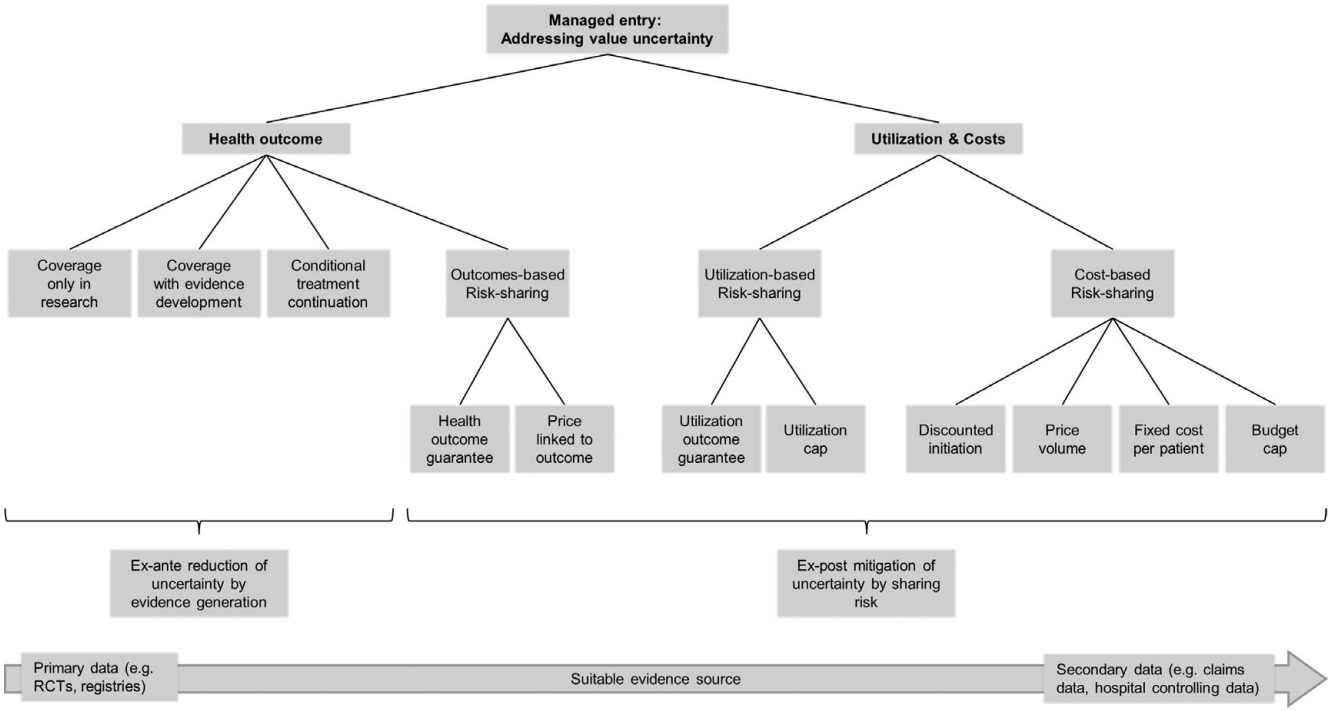

Figure 1. Value-based typology of managed entry agreements (MEA).

conducted to identify papers that provide information regarding the quality of claims data for research purposes in general and German SHI data in particular (see Supplementary Table 1). This literature was searched for relevant information regarding the reliability and validity of the information available in claims data as well as complementary information regarding the other criteria.

Face validity of the criteria and the assessment of the suitability of routine data for MEA with respect to the criteria were assessed by an independent third party (L.S.) who is experienced in the analysis of German SHI claims data.

RESULTS

Data Needs of Different MEA

Based on the corresponding data needs, the eleven types of MEA that were identified in the literature search could be structured as displayed in Figure 1. Similar to existing taxonomies (14;15), the structure is oriented at the component of the novel technology's value which is addressed by additional evidence collection—health outcome or utilization and costs.

MEA Using a Health Outcome

CED schemes with "in research" or "with research" arrangements as well as conditional treatment continuation (CTC) agreements are typically used to reduce uncertainty about the health outcome. Evidence is collected before the final coverage decision is made. This requires detailed evidence on clinical endpoints which are usually generated in a clinical trial or registry established in line with the MEA (2).

In contrast to CED schemes, risk-sharing agreements (RSA) mitigate uncertainty after the coverage decision has been made by distributing the financial risk associated with the coverage of the novel technology between contracting parties (5). In health outcomes-based RSA, the reimbursement level of a novel technology is dependent on its effectiveness. Health outcome is measured as intermediate endpoints (e.g., biomarkers, tumor shrinkage) or final endpoints (morbidity, survival, disease severity, quality of life) (14).

MEA Using Utilization- and Cost-based Outcomes

Utilization-based RSA address uncertainty around the utilization of a novel technology in routine care (4). Evidence is required on the type of healthcare provider, for example, if reimbursement is limited to certain specialists. Characteristics of patients receiving a health technology may be of interest if only certain patients (e.g., high-risk patients) are covered by a RSA. Also, the expertise of the healthcare provider in the utilization of the novel technology (e.g., surgery method) as well as the compliance of patients might be of interest in a utilization-based RSA. In case of utilization caps, data are required on the number of doses, treatment cycles, or devices used per patient.

In cost-based RSA, the payer might reimburse a novel technology for a discounted price at treatment initiation, whereas the price reverts to list price after an agreed number of treatment cycles or period of time (15). Similarly, price–volume

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

Table 1. Assessment of Claims Data in Health Outcome-Based MEA

Type of MEA		Data need	Availability	Completeness	Timeliness	Confidentiality	Reliability	Validity
Coverage only in research		Clinical data (e.g. biomarker)	No	NA	NA	NA	NA	NA
CED		Clinical data (e.g. biomarker)	No	NA	NA	NA	NA	NA
CTC		Clinical data (e.g. biomarker)	No	NA	NA	NA	NA	NA
Health outcome-based RSA	Intermediate endpoint	Biomarker/ laboratory values	No	NA	NA	NA	NA	NA
	Final endpoint	Morbidity	Proxy: Diagnostic codes	Main diagnosis: complete Secondary diagnosis: depending on relevance for DRG weight	3–6 months	Data from morbidity-based risk compensation scheme	Selection/ classifi-cation bias Confounding	Economic incentives in coding practice
			Proxy: Hospital episodes	Complete	3–6 months	Not confidential	High	Hospital episodes: Only SHI relevant episodes coded
			Proxy: Procedure codes following adverse events/ complications	See Table 2 – dose cap	3–6 months	Not confidential	High	See Table 2 – dose caps
		Survival	Date of death (basic claims data, inpatient care)	High	3–6 months	Not confidential	High	No information on cause of death
		Disease severity	No	NA	NA	NA	NA	NA
		Quality of life	No	NA	NA	NA	NA	NA

agreements regulate the price paid per treatment unit, whereas the price is tied to the overall amount of a product used. Both schemes require data on the number of doses, treatment cycles, or devices used per patient. Setting a fixed price per patient for the complete course of treatment is a common scheme as well; data on treatment costs at a patient level are required. A budget cap might also be used in a cost-based RSA, where the total expenditure on a treatment is limited (15).

Assessment of SHI Data

The following sections describe the extent to which the data needs of the MEA types are met by SHI data, according to the assessment criteria described above.

Availability

Clinical information such as laboratory values or tumor progression is not included in SHI claims data. There can be exceptions to this rule, for example, if sickness funds collect laboratory values within a disease management program. However, generally, SHI data are unlikely to be of use in conditional coverage agreements (CED or CTC) as well as in health outcome-based RSA using intermediate clinical endpoints such as biomarkers or tumor progression (10).

In health outcome-based RSA using morbidity as an outcome measure, inpatient and outpatient diagnostic codes, length of stay, and the number of hospital episodes may be used as proxies for morbidity in SHI claims data. Date of death is included in basic SHI claims (16); there is, however, no information

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

Table 2. Assessment of Claims Data in Utilization Outcome-Based MEA

Type of MEA	Data need	Availability	Completeness	Timeliness	Confidentiality	Reliability	Validity
Utilization outcome	Type of health care provider	Institutional codes: hospitals, physicians, pharmacies, physical therapists	High	3–6 months	Agreement of health care provider necessary	High	High
	Type of patient receiving treatment	Basic claims data	Date of birth, sex, address: all insured persons	No time lag	Date of birth, address: data protection measures	High	Address, occupation: often out of date
			Premium, family status, number of children, education, occupation: main insured persons only				Premium: because of assessable income limit, not always representative of income
		Diagnosis codes	See Morbidity	See Morbidity	See Morbidity	See Morbidity	See Morbidity
	Expertise	Proxy: no. of treatments per hospital/outpatient doctor	High	3–6 months	Not confidential	High	Other factors influencing expertise
	Compliance	Proxy: drug/device collection at provider	High	3 months	Not confidential	High	Not representative of drug intake/device use

available on cause of death, severity of disease (e.g., cancer staging), or quality of life. Severity of disease might, depending on the disease in question, be approximated by combining data from several healthcare sectors, e.g., diagnoses and pharmaceuticals.

Regarding utilization-based RSA, SHI claims data allow the identification of individual healthcare providers by institutional codes for physicians, hospitals, pharmacies, and physiotherapists. Physicians' institutional codes also allow the identification of specialist groups (10). Diagnostic codes from inpatient and outpatient treatment provide information on patient characteristics, as discussed previously. In basic claims data additional relevant information on classification of patients is included, particularly if age and gender are of interest in determining a patient's risk status. Reasons for dropout are also coded, which are death or transition to another insurance fund. RSA conditioned on expertise can be approximated by the number of treatments per hospital or provider coded in SHI claims data. Schemes based on patients' compliance might control patients picking up drugs or devices at a healthcare provider (10). Number of doses or treatment cycles can be identified in SHI claims data by inpatient operation and procedure codes (OPS) (17), outpatient doctor's fee scale codes (18), the central pharmaceutical number, or in the case of assistive technologies, a distinct position number (19).

For cost outcome-based RSA, cost data are available for inpatient (diagnosis related group, DRG) and outpatient (doctor's fee scale) care, as well as pharmaceuticals, physiotherapy, assistive technologies, and rehabilitation at a patient level in SHI claims data.

Completeness

Owing to the administrative nature of the data, only SHI relevant healthcare utilization is coded, which excludes, for example, out-of-pocket payments (drugs, medical devices) or coverage by other types of private or social insurance (e.g., rehabilitation by pension funds) (20). Claims data contain basic information on age, gender, and place of residence, whereas more sophisticated socio-demographic information (household size, income, education, occupation) is documented only in a limited manner.

Timeliness

The timeliness of SHI claims data varies by healthcare sector. Complete inpatient data are reported to the SHI fund immediately after discharge. Outpatient data are usually delayed by approximately 6 months as they are transferred over the Association of Statutory Health Insurance Physicians to insurance funds. Diagnostic information in outpatient data is, in general,

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

summarized on a quarterly basis per patient, whereas inpatient data are reported on a hospital-episode basis per case (§301 SGB V, §295 Abs 2 SGB V) (6).

Confidentiality

Confidentiality issues are to be considered when patient data are not anonymous or if pseudonymization is not possible. In such cases, ethics approval and/ or the approval of the data protection agency as well as informed consent from affected patients is necessary to use claims data for RSA (13). Costs controlled by selective contracts between manufacturers and insurance funds (e.g., rebates on pharmaceuticals) might be subject to confidentiality as well.

Reliability

Diagnostic codes in SHI claims data are coded using the ICD system. A study by Stausberg et al. (2008) shows that the reliability of diagnoses coding using ICD-10 is only fair to moderate, with agreement rates between coders of around 50 percent (21). One possible explanation for low agreement rates is the complexity of the German coding system, which is difficult to understand even for coding experts.

The central pharmaceutical number and the distinct position number of assistive technologies both allow the identification of individual health technologies. The central pharmaceutical number even allows distinction between products from different manufacturers. However, it is possible that codes change over time, for example when a technology is no longer on the market. The same applies to OPS codes and doctor's fee scale codes.

Overall, cost information from SHI claims data is reliable as actual spending on a broad range of services and technologies incurred by SHI is reported (20).

Validity

Regarding the coding quality of diagnostic codes, inpatient data can be assumed to be more comprehensive than outpatient data, and there is evidence that comorbidities tend to be underreported in outpatient care in Germany (22). This might be explained by the current reimbursement practice in Germany. In the hospital sector, treatment is reimbursed by means of a diagnosis-specific fixed remuneration, which is also weighted by the patient's comorbidity status, the DRG, whereas in outpatient care, the amount of comorbidity hardly affects reimbursement (22).

Length of stay and the number of hospital episodes, identified by date of admission or discharge, generally have a good coding quality in SHI claims data as these are used for billing purposes. Also, date of all-cause death is coded validly for all insured persons because of its importance in the management of insurees.

Intake and utilization of a technology do not appear in SHI claims data. Thus, validity of compliance measures should be assessed critically. In outpatient care, as a result of physicians' budget restrictions, only part of the provided medical services are incurred at the expense of SHI, and thus appear in claims data.

The validity of utilization and cost data can be influenced by patients' morbidity. If the costs of alternative treatments are compared in an RSA, the differing baseline morbidity of patients should be considered. Differences in costs might result from systematic differences in disease severity rather than the performance of medical technologies.

Tables 1 to 3 provide an overview of health, utilization, and cost outcome-based MEA, their respective data needs, and the assessment of SHI data regarding their suitability to meet these data needs.

DISCUSSION

Interpretation of Results

To our knowledge, this study is the first to assess the use of routinely collected claims data in MEA in a German context. Briggs et al. (2010) (7) state that "good routine data capture mechanisms have a potentially crucial role to play in the feasibility of any managed entry scheme". This study suggests that routinely collected claims data are a useful data source to address uncertainty about a new technology's costs and effectiveness at market entry; however, they are not suitable for every type of MEA. The usefulness of claims data depends on the source of uncertainty and, in consequence, the outcome measures chosen in the agreements.

The particular strength of claims data use in MEA is the data's ability to display utilization of medical technologies in routine care and real-world cost data incurred by SHI funds (10). As most RSA are conducted because there is uncertainty around utilization or the budget impact of a medical technology in clinical practice, this aspect of claims data is most important. Also, timeliness of evidence, compared with clinical trials and registries, is important in the coverage decision-making process (12). Use of claims data is less expensive than conducting clinical trials or establishing registries (10). This might be a particular advantage for cost-based RSA, as the high costs of primary data collection often exceed the cost savings incurred by an MEA (4). Furthermore, collection of claims data represents no extra burden for patients and clinicians compared with clinical trials and registries. This might enhance stakeholders' acceptance of MEA.

In schemes where safety aspects or clinical efficacy are assessed, SHI claims data are unlikely to be sufficient for most contracts as the sole source of evidence because clinical information is not included in sufficient detail. Also, using claims data is not possible if the technology of interest cannot be

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

Table 3. Assessment of Claims Data in Cost Outcome-Based MEA

Type of MEA	Data need	Availability	Completeness	Timeliness	Confidentiality	Reliability	Validity
Fixed costs per patient	Costs of treatment per patient	Inpatient: DRGs	Costs reimbursed by SHI: complete	3–6 months	Selective contracts with manufacturers	High	Costs incurred at insurance: high
		Outpatient: doctors' fee scale (EBM) Pharmaceuticals Physical therapy Assistive technologies	Out of pocket payments, co-payments, other social insurers: not included				Costs incurred at patient/health care provider: not reflected
Dose cap	No. of doses	Inpatient: OPS	Only claims included that are relevant for SHI	3–6 months	Not confidential	High	Economic incentives in coding practice
	treatment cycles/ devices per patient	Outpatient: EBM Pharmaceuticals: Central Pharmaceutical Number Assistive technologies: Position number	Coded procedures: complete				Budget restrictions: provided work not coded (especially outpatient)
Discounted initial price	No. of doses treatment cycles/ devices per patient in defined period	See dose cap	See dose cap	See dose cap	See dose cap	See dose cap	See dose cap
Budget cap	Total costs for treatment/device	See fixed costs per patient	See fixed costs per patient	See fixed costs per patient	See fixed costs per patient	See fixed costs per patient	See fixed costs per patient
Price related to volume	No. of doses treatment cycles/ devices in a certain period	See dose cap	See dose cap	See dose cap	See dose cap	See dose cap	See dose cap

identified distinctly in the data due to ambiguous coding. Their validity is also influenced by the multiple biases described in this study. Especially if comparative effectiveness and costs are assessed within a MEA, selection bias might arise as a result of nonrandomization. Matching techniques and other statistical methods can be used to adjust for confounding variables but are also subject to limitations (11). SHI claims data provide information on statutorily insured people, who constitute 85 percent of the German population; correspondingly, service utilization is restricted to utilization within the SHI system. This should be considered when data collected in a scheme must be representative of the whole population, as costs and effects might vary between insurance populations. In most cases, however, information on an insurance-specific population is useful for RSA because reimbursement of a novel technology is negotiated between an individual SHI fund and a manufacturer; the SHI fund is interested in the effectiveness and costs of the technology in its own insured population.

Limitations

A value-based typology of MEA is presented in this study, which might not accommodate all aspects of MEA discussed in the literature. For example, we did not specify how evidence collected in MEA is translated into price or revenue adjustments. Furthermore, published examples show that, in practice, hybrid forms of MEA are possible, such as CED schemes with a price adjustment arrangement. However, a clear distinction is needed in this study to assess the data requirements of different types of MEA. Even if different MEA types are combined into more complex agreements, decision makers in industry and health insurance should still be aware of the strengths and weaknesses of its single components.

This study assesses data that are generally available from SHI funds through German social law. The range and structure of the data are extracted from literature searching only; even if all the authors could additionally draw upon experience of working with sickness fund routine data, no systematic analysis

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press

of original data provided by an SHI fund was conducted in this study.

To our knowledge, there is no published standardized approach to assess the suitability of claims data for MEA. Owing to the novelty of this topic, no large number of published studies regarding MEA in Germany could be expected; furthermore, there is no publicly accessible data source of existing contracts and the data they are based on. Therefore, this assessment had to be based on a rather exploratory methodology. Furthermore, work is needed to confirm the results of this study. Also, while this framework provides an overview of general issues in the development of MEA, there are still many context-specific aspects that have to be taken into account as well.

Generalizability of Results

The assessment criteria are selected in a generic way so that they can be applied to the assessment of routinely collected data in any healthcare system. The suitability of routine data for different types of MEA may differ from the results presented here in the German healthcare context. The extent of clinical information included—such as laboratory values available, for example, in the General Practitioners Research Database (UK)—has a particularly strong influence on the suitability of using claims data in these schemes (23). In this case, MEA involving clinical outcomes can be based on claims data, which is not possible with German SHI claims data. Also, the patient population available in the data set might bias results, especially if insurance eligibility is dependent on the employment situation or age, as is common in U.S. health care (24).

Other aspects discussed here, for example confounding and data protection issues as well as disease-specific difference in the validity of claims data, are applicable to all insurance systems. Also, claims data are collected for billing purposes, and coding practice might be influenced by economic incentives in all insurance systems.

Implications

MEA appear not to be used frequently yet in the management of medical innovation in Germany. This study may support contracting parties when selecting the most suitable contract type and agreeing on contract conditions. To use MEA in the management of product innovation, manufacturers might consider conducting early economic evaluation in the planning phase of a MEA. Early evaluation might be used to define outcome measures for the MEA, which are the best indicators that the promised value is indeed delivered. Claims data gathered over the duration of the scheme can be incorporated into decision models for the calculation of costs and cost-effectiveness measures (25).

There is a need for more research on MEA. In particular, the assessment criteria described here need to be applied to case studies to validate our results. Further research is also necessary to assess the legal and ethical issues involved with claims data use in MEA, as well as the suitability of different methods of data analysis for different contract types. The assessment criteria proposed in this study might also be used to evaluate the extent to which data generated in hospitals are suited for MEA. Hospitals usually adopt novel technologies first and are an attractive contract partner for manufacturers.

CONCLUSIONS

To our knowledge, this study is the first to develop a framework for the systematic assessment of using claims data in MEA. This framework supports manufacturers as well as payers in selecting the most suitable contract type and agreeing on contract conditions. The suitability of SHI claims data to validate the value proposition of a new technology depends on which component of value, effects or utilization and costs, are most uncertain. Generally, claims data are better suited to addressing uncertainty around utilization and budget impact than uncertainty about effectiveness. Further research is needed to apply the framework in case studies, to assess the use of MEA by institutions other than health insurance funds, and to explore the ethical and legal requirements for MEA in Germany.

SUPPLEMENTARY MATERIAL

Supplementary Table 1
http://dx.doi.org/10.1017/S0266462316000131

CONFLICTS OF INTEREST

The authors have no conflict of interest with regard to this project. Publication of the study results was not contingent on the funding party's approval or censorship of the manuscript.

REFERENCES

1. Rogowski WH. An economic theory of the fourth hurdle. *Health Econ*. 2013;22:600-610.
2. Hutton J, Trueman P, Henshall C. Coverage with evidence development: An examination of conceptual and policy issues. *Int J Technol Assess Health Care*. 2007;23:425-432.
3. Adamski J, Godman B, Ofierska-Sujkowska G, et al. Risk sharing arrangements for pharmaceuticals: Potential considerations and recommendations for European payers. *BMC Health Serv Res*. 2010;10:153.
4. Stafinski T, McCabe CJ, Menon D. Funding the unfundable: Mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics*. 2010;28:113-142.
5. Garrison LP Jr, Towse A, Briggs A, et al. Performance-based risk-sharing arrangements-good practices for design, implementation, and evaluation: Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force. *Value Health*. 2013;16:703-719.
6. Fünftes Buch Sozialgesetzbuch - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477), das durch Artikel 3 des Gesetzes vom 22. Dezember 2011 (BGBl. I S. 3057) geändert worden ist. (1988).

7. Briggs A, Ritchie K, Fenwick E, Chalkidou K, Littlejohns P. Access with evidence development in the UK: Past experience, current initiatives and future potential. *Pharmacoeconomics*. 2010;28:163-170.
8. Mohr PE, Tunis SR. Access with evidence development - the US experience. *Pharmacoeconomics*. 2010;28:153-162.
9. Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: Defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force. *Value Health*. 2009;12:1044-1052.
10. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58:323-337.
11. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: Analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report–Part III. *Value Health*. 2009;12:1062-1073.
12. Garrison LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR Real-World Data Task Force Report. *Value Health*. 2007;10:326-335.
13. Willison DJ. Health services research and personal health information: Privacy concerns, new legislation and beyond. *Can Med Assoc J*. 1998;159:1378-1380.
14. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy*. 2010;96:179-190.
15. Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. *Value Health*. 2012;15:570-579.
16. Kuepper-Nybelen J, Hellmich M, Abbas S, et al. Association of long-term adherence to evidence-based combination drug therapy after acute myocardial infarction with all-cause mortality. *A prospective cohort study based on claims data. Eur J Clin Pharmacol*. 2012;68:1451-1460.
17. Deutsches Institut für Medizinische Dokumentation und Information (DIMDI). Operationen- und Prozedurenschlüssel Version 2013. 2013 [updated 2013]. http://www.dimdi.de/static/de/klassi/ops/kodesuche/onlinefassungen/opshtml2013/index.htm (accessed July 8, 2013).
18. Kassenärztliche Bundesvereinigung. Einheitlicher Bewertungsmaßstab für ärztliche Leistungen. 2013 [updated 2013]. http://www.kbv.de/ebm2013/EBMGesamt.htm (accessed July 8, 2013).
19. Spitzenverband der gesetzlichen Krankenversicherungen. Hilfsmittelverzeichnis des GKV-Spitzenverbandes. 2007 [updated 2007]. https://hilfsmittel.gkv-spitzenverband.de/home.action (accessed July 8, 2013).
20. Schwarzkopf L, Menn P, Leidl R, et al. Excess costs of dementia disorders and the role of age and gender - An analysis of German health and long-term care insurance claims data. *BMC Health Serv Res*. 2012;12:165.
21. Stausberg J, Lehmann N, Kaczmarek D, Stein M. Reliability of diagnoses coding with ICD-10. *Int J Med Inform*. 2008;77:50-57.
22. Erler A, Beyer M, Muth C, Gerlach FM, Brennecke R. Garbage in - Garbage out? Validitat von Abrechnungsdiagnosen in hausarztlichen Praxen [Garbage in - garbage out? Validity of coded diagnoses from GP claims records]. *Gesundheitswesen*. 2009;71:823-831.
23. García Rodríguez LA, Pérez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *Br J Clin Pharmacol*. 1998;45:419-425.
24. Hennessy S, Bilker WB, Weber A, Strom BL. Descriptive analyses of the integrity of a US Medicaid claims database. *Pharmacoepidemiol Drug Saf*. 2003;12:103-111.
25. O'Malley SP, Selby WS, Jordan E. A successful practical application of coverage with evidence development in Australia: Medical Services Advisory Committee interim funding and the PillCam Capsule Endoscopy Register. *Int J Technol Assess Health Care*. 2009;25:290-296.

https://doi.org/10.1017/S0266462316000131 Published online by Cambridge University Press