

Das Gesundheitswesen

Evaluation von neuen Versorgungsformen für Menschen mit psychischen Erkrankungen: Pro & Contra zum Einsatz von GKV-Routinedaten – Ein Erfahrungsbericht

Anne Neumann, Roman Kliemt, Fabian Baum, Martin Seifert, Denise Kubat, Stefanie March, Ines Weinhold, Enno Swart, Jochen Schmitt.

Affiliations below.

DOI: 10.1055/a-2297-5347

Please cite this article as: Neumann A, Kliemt R, Baum F et al. Evaluation von neuen Versorgungsformen für Menschen mit psychischen Erkrankungen: Pro & Contra zum Einsatz von GKV-Routinedaten – Ein Erfahrungsbericht. Gesundheitswesen Das 2024. doi: 10.1055/a-2297-5347

Conflict of Interest: The authors declare that they have no conflict of interest.

This study was supported by Konsortium bestehend aus mehr als 70 GKV-en, keine

Abstract:

Der vorliegende Artikel stellt die Besonderheiten in der Evaluation von Versorgungsformen bei Menschen mit psychischen Erkrankungen in Deutschland dar. Es werden die Erfahrungen der Evaluation von Modellvorhaben in der psychiatrischen Versorgungsforschung basierend auf Routinedaten gesetzlicher Krankenversicherungen (GKV) von über 70 gesetzlichen Krankenkassen (EVA64-Studie) systematisch dargestellt.

Am Beispiel der EVA64-Studie werden Empfehlungen für den Einsatz von GKV-Routinedaten im Bereich psychiatrischer Versorgungsforschung und allgemein abgeleitet. (1) Zunächst wird die Studie und der Einsatz der genannten Daten in dieser beschrieben und eingeordnet. (2) Dann werden die einzelnen Outcomes der Studie dargestellt und bewertet, (3) um danach Kriterien, Stärken und Einsatzmöglichkeiten dieser Daten darzustellen. (4) Abschließend werden Empfehlungen für die Weiterentwicklung von GKV-Routinedaten als Bewertungsgrundlage formuliert.

The article presents the special features of an evaluation within mental health services research in Germany. The experiences of the evaluation of flexible and integrated treatment options with a global treatment budget in psychiatric hospitals based on routine data of more than 70 statutory health insurance (SHI) funds (EVA64 study) are systematically presented.

Using the EVA64 study as an example, recommendations for the use of claims data in the field of mental health services research and in general are derived. (1) First, the study and the use of claims data in this study is described and classified. (2) Then, the individual outcomes of the study are presented and evaluated to (3) derive criteria, strengths and potential uses of claims data. (4) Finally, recommendations for the further development of claims data from SHI funds as a basis for evaluation are described.

Corresponding Author:

Dr. Anne Neumann, Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus der Technischen Universität Dresden, Zentrum für Evidenzbasierte Gesundheitsversorgung, Dresden, Germany, anne.neumann@ukdd.de

Affiliations:

Anne Neumann, Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus der Technischen Universität Dresden, Zentrum für Evidenzbasierte Gesundheitsversorgung, Dresden, Germany

Roman Kliemt, Wissenschaftliches Institut für Gesundheitsökonomie und Gesundheitssystemforschung, WIG2, Leipzig, Germany

Fabian Baum, Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus der Technischen Universität Dresden, Zentrum für Evidenzbasierte Gesundheitsversorgung, Dresden, Germany

[...]

Jochen Schmitt, Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus der Technischen Universität Dresden, Zentrum für Evidenzbasierte Gesundheitsversorgung, Dresden, Germany



This article is protected by copyright. All rights reserved.

Accepted Manuscript

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Tabelle 1: kritische Bewertung der Outcomes in der EVA64-Studie

Outcome	Risiko für Bias		Stärke der Abbildbarkeit mit GKV-Routinedaten		Relevanz für Versorgungskontext (Menschen mit psychischen Erkrankungen)	
	Bewertung 1 = hoch; 2 = mäßig; 3 = gering; 4 = kein	Erläuterung	Bewertung 1 = nicht da; 2 = gering; 3 = mäßig; 4 = sehr gut	Erläuterung	Bewertung 1 = keine; 2 = gering; 3 = mäßig; 4 = sehr gut	Erläuterung
Operationalisierung						
voll- und teilstationäre psychiatrische Behandlungsdauer	4		4		4	
Zählung der Tage anhand von Abrechnungen für alle aufgegriffenen Personen (auch ohne stationärem Aufenthalt)						
ambulante Leistungen, inkl. PIA	3	Nur Betrachtung innerhalb SGB V möglich	3	Einteilung in Haupt- und Nebendiagnosen fehlen; Dokumentation der Diagnosen sind nicht erlösrelevant; lückenlose sektorenübergreifende Erfassung	3	Es werden nur Häufigkeiten gezählt und Kontakte nicht qualitativ bewertet
Zählung der Tage mit Kontakte						
Behandlungskontinuität	3	Nur Betrachtung innerhalb SGB V möglich	3	Keine Abbildung der Behandlung im selben Setting bzw. vom gleichen Behandelnden bzw. Team möglich	4	
Anteil derer mit mindestens einer ambulanten Behandlung (im						

Krankenhaus (oder außerhalb) innerhalb eines definierten Zeitraums nach stationärer Behandlung						
Kontaktabbrüche Anteil mit mindestens einem Kontakt zum psychiatrischen Versorgungssystem innerhalb eines definierten Zeitraumes für Menschen mit schweren psychischen Störungen	3	Nur innerhalb SGB V	3	Ein gemessener Kontaktabbruch stellt nicht zwangsweise wirklich einen Kontaktabbruch dar (falscher Alarm)	3	Geringe Fallzahl der Menschen mit schweren psychischen Störungen
Arzt- und Krankenhaus-Hopping Anteil der Versicherten mit mindestens drei unterschiedlichen Leistungserbringern im psychiatrischen Sektor	3	Nur innerhalb SGB V	4		3	Betrifft sehr wenige Versicherte
Arbeitsunfähigkeit Anteil mit mindestens einer und Dauer der Arbeitsunfähigkeit	2	Nur auf einen Teil der Menschen mit psychischen Erkrankungen anwendbar (Krankengeldanspruchsberechtigte z.B. nicht Studierende, verrentete Personen, Personen ohne Erwerbseinkommen) und kurze Krankmeldungen ggf. über	3	Alle gemeldeten AU-Fälle erfasst, jedoch zwei Einschränkungen (siehe Bias)	3	Teilweise nur indirekt vom Modellvorhaben beeinflussbar

		„andere Erkrankungen“ oder über arbeitgeberinterne „Krank-ohne-Schein“-Tage abgefangen (Untererfassung Kurzzeit-AU)				
Wiederaufnahmerate Anteil der Personen, die innerhalb von einem Jahr nach stationärer Entlassung erneut stationär aufgrund einer psychischen Diagnose aufgenommen wurden	4		4		4	
Komorbiditäten Score zur Messung von Begleiterkrankungen (basierend auf Elixhauser Score)	3	Abhängig von ICD-Kodierqualität (stationär: beeinflusst von Abrechnungslogik; ambulant: Diagnosen nicht erlösrelevant)	2	Score wurde jedoch für Wahrscheinlichkeit, im Krankenhaus zu versterben, ermittelt	2	Wenig Relevanz für Versorgungskontext, da Elixhauser-Score somatisch ausgerichtet
Mortalität Anteil der Personen, die innerhalb eines definierten Zeitraumes versterben	4		4		2	Mortalität nur geringfügig direkt durch Modellvorhaben beeinflussbar (Ausnahme Suizid), keine Todesursache verfügbar
Erkrankungsprogression Anteil derer, die eine schwere depressive Episode (aus leichter/mittelgradiger),	2	Gut über ICD-10 abbildbar, jedoch Kodierung einer Erkrankungsprogression bei Menschen mit psychischen Erkrankungen sehr eingeschränkt möglich bzw. individuelle	2	Nur über drei Wege in der ICD-10-Kodierung abbildbar; keine PROs, keine diagnostischen Maße/Screenings	2	Da abhängig von Kodierung; nur drei Pfade; Erkrankungsschwere psychisch Erkrankter nicht mit GKV-

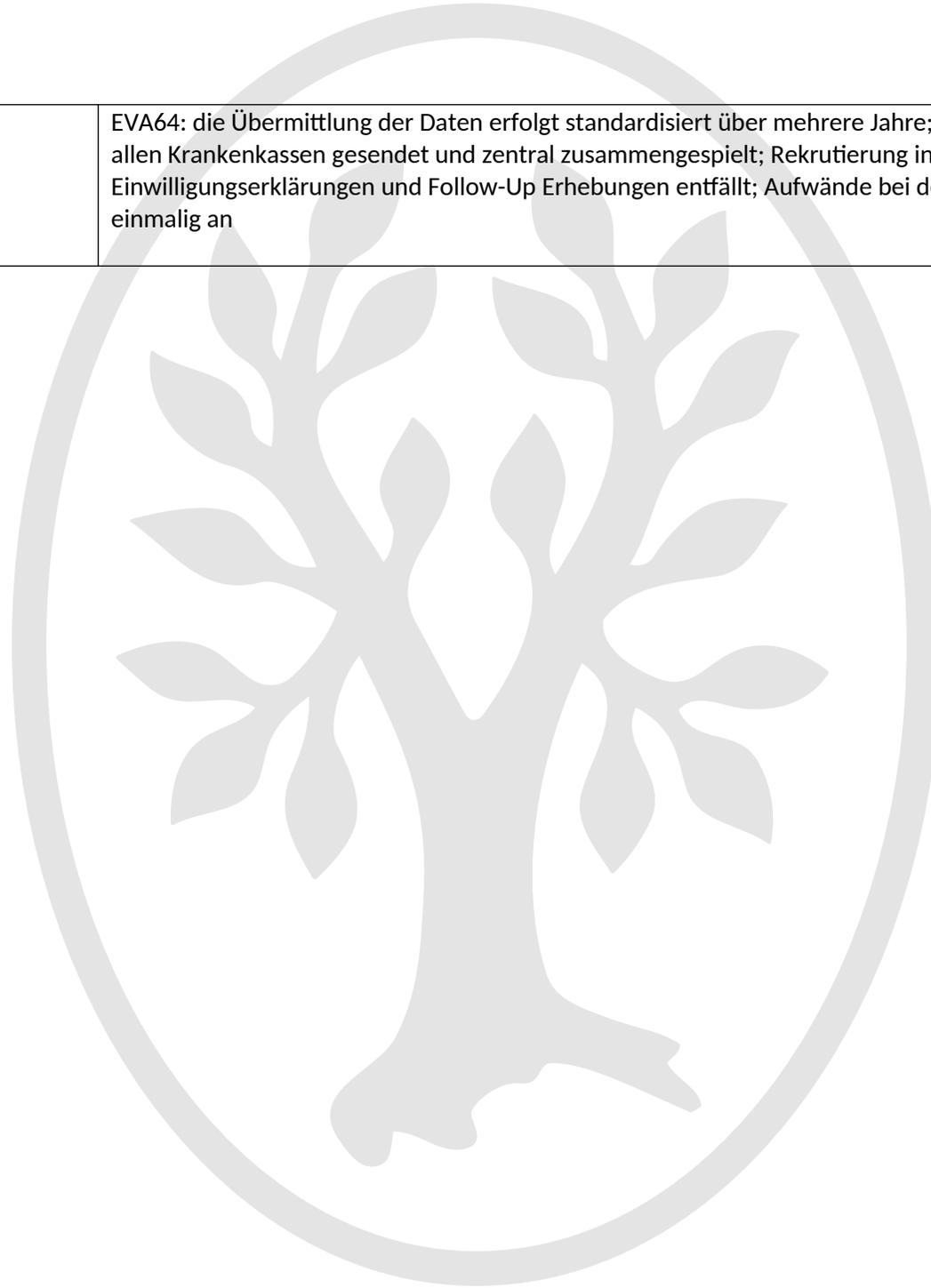
eine rezidivierende depressive Störung (aus jeglicher depressiver Episode) oder eine psychische oder Verhaltensstörung durch psychotrope Substanzen innerhalb eines definierten Zeitraumes entwickeln		Entscheidung zwischen Kodierenden				Routinedaten messbar
Leitlinienadhärenz Abbildung von 10 Qualitätsindikatoren abgeleitet von Großimlinghaus et al.	1	Aspekte wie Dokumentation in der PIA oder Medikation im Krankenhaus mit GKV-Routinedaten nicht abbildbar, daher betrachtete Personen eingeschränkt	1	Da Nenner oft eingeschränkt und Angebot und Gründe für Nichtinanspruchnahme nicht ersichtlich	2	Wäre sehr wichtig, kann jedoch schwer abgebildet werden
Direkte Kosten	2	Preiseffekte wie Ausgleichszahlungen aus Vorjahren nicht ersichtlich; daher Verzerrung möglich	3	Nur Kosten aus Sicht der GKV	3	
Kosteneffektivität	3	Abhängig von Effektivität und Kostenmessung	2	Da Effektivität (primäre Outcomes, gemessen an GKV-Daten) und Kosten der GKV sich bedingen	3	

Tabelle 2: Bewertung von GKV-Routinedaten zur Evaluation (psychiatrischer) Modellvorhaben

Kriterium	Studien basierend auf GKV-Routinedaten, Relevanz EVA64-Studie
Datenbasis	<p>Fest, kaum zusätzliche settingbezogene Parameter erfassbar; Datenfelder bekannt; teilweise sind Sondervereinbarungen wie Selektivverträge zu berücksichtigen;</p> <p>EVA64: um Modellvorhaben evaluieren zu können, sind Kenntnisse zur modellhaften Behandlung und Abrechnung essentiell; auch ist ein Verständnis der Versorgung in der Regelversorgung (kollektiv wie selektiv) wichtig, um diese korrekt und ganzheitlich abbilden zu können (z.B. Abbildung der Psychiatrischen Institutsambulanz, Abbildung der Modellbestandteile in den Daten, Veränderungen der Dokumentation über die Zeit)</p>
Perspektive	<p>Sicht der Datenbereitstellenden, z.B. Leistungserbringenden, GKV;</p> <p>EVA64: die Einordnung der Perspektive der Datenquelle ist wichtig; weitere wichtige Perspektiven, wie die der Patientinnen und Patienten und deren individuelle Risikofaktoren und Ressourcen können nicht abgebildet werden und erfordern andere Datenquellen bzw. das Erweitern von GKV-Routinedaten; die klare Kommunikation von Möglichkeiten und Grenzen ist vorzunehmen</p>
Erreichbarkeit	<p>Alle Versicherte erreichbar (Einschränkung: privat Krankenversicherte, Nichtversicherte); große Fallzahl möglich;</p> <p>EVA64: um Modellvorhaben evaluieren zu können, ist der Einbezug und die Nachbetrachtung aller behandelten Menschen am Krankenhaus wichtig; ein selektiver Einschluss verschiedener Settings, wie oft in Primärstudien aufgrund der Versorgungssituation in der psychiatrischen Versorgung zu vermuten, ist für eine valide Evaluation hinderlich</p>
Generalisierbarkeit/Repräsentativität	<p>Hoch, sofern bundesweite Daten mehrerer großer Krankenkassen eingehen;</p> <p>EVA64: für die Ableitung der Effektivität in der Evaluation ist der hier gegebene hohe Versorgungsanteil und der Einschluss von vielen Krankenkassen essentiell</p>
Inhalte Forschungsfragen	<p>Inanspruchnahme, Versorgung, (eingeschränkt) Outcome; eingeschränkt auf Variablen, die in GKV-Routinedaten zur Verfügung stehen;</p> <p>EVA64: für die Evaluation muss die Einschränkung der Möglichkeit der Abbildung patient:innenbezogener Outcomes bewusst sein, d.h. GKV-Routinedaten eignen sich sehr gut für die Abbildung der Inanspruchnahme und Versorgung, aber nicht für die Darstellung patient:innenzentrierter Ergebnisse</p>

Abbildung des Versorgungsgeschehens	Inanspruchnahme, Versorgungsprozesse, Diagnosen/Morbidität, Monitoring Versorgungsgeschehen, sektorenübergreifend, Zugang zum Versorgungssystem, Kosten aus Perspektive der Datenbereitstellenden EVA64: obenstehende Aspekte können mit den GKV-Routinedaten gut bzw. sehr gut abgebildet werden
Abbildung der Qualität der Behandlung	Kaum möglich, solange sich diese nicht in der leitliniengemäßen Erbringung standardisiert dokumentierter Leistungen (über EBM- und OPS-Klassifikation) abbilden lässt. EVA64: diese Einschränkung muss bei der Einordnung der Ergebnisse berücksichtigt werden
Abbildung der Präferenz der Patientinnen und Patienten	Nicht möglich EVA64: die Präferenzen der Patientinnen und Patienten bzw. der Leistungserbringer kann mit den GKV-Routinedaten nicht betrachtet werden; hier sind weitere Daten notwendig
Bias	Ungenügende Kodierung der Diagnosen und Leistungen möglich; u.a. potenzielle Bias: <ul style="list-style-type: none"> - Missclassification Bias möglich, da nur dokumentationsrelevante Informationen vorliegen (z.B. weniger gute Passfähigkeit zwischen zwei zu vergleichenden Gruppen, da nur GKV-Routinedaten); - Inanspruchnahmebias (keine Inanspruchnahme trotz Behandlungsbedarf) möglich; - Kein Apprehension Bias (Verhaltensänderung durch Kenntnis der Studienteilnahme); - Kein Attrition Bias (systematischer Unterschied im Ausscheiden aus der Studie zwischen beiden Gruppen); - Kein Information Bias (ggf. unterschiedliche Ansprache in beiden Gruppen), z.B. Erinnerungsbias (stationäre Aufenthalte sind präsenter und häufiger in Kontrollgruppe); - Kein Non-response Bias (systematisch unterschiedliche Unterschiede zwischen Antwortenden und Nicht-Antwortenden); - Geringes Risiko für Selektionsbias bei der Auswahl der Studienteilnehmenden (ausgenommen Nicht- und Privatkrankenversicherte): (keine selektierte Nichtteilnahme bzw. Lost-to-Follow-Up), u.a. Neyman Bias, Healthy Worker Effect; - Kein Recall Bias (Gruppen erinnern sich unterschiedlich); - Kein Volunteer bias (Freiwillige sind verschieden von der Allgemeinbevölkerung) EVA64: eine explizite Bewertung dieser Bias im konkreten Setting der Anwendung muss vorgenommen werden, wenn Projekte mit GKV-Routinedaten im Feld der psychiatrischen Versorgungsforschung umgesetzt werden
Rekrutierungszeit & -kosten	Geringer als primärdatengestützte Evaluation (Kosten für Infrastruktur, jährliche Datenaufbereitung sowie Berichterstellung)

EVA64: die Übermittlung der Daten erfolgt standardisiert über mehrere Jahre; alle Daten werden zentral von allen Krankenkassen gesendet und zentral zusammengespielt; Rekrutierung in Zentren über Einwilligungserklärungen und Follow-Up Erhebungen entfällt; Aufwände bei den Kassen fallen im Wesentlichen einmalig an



Ausgangslage

In Deutschland werden, im Vergleich zu anderen europäischen Ländern, weiterhin sehr viele Menschen stationär versorgt, denen häufig auch ambulant geholfen werden könnte (1). Als wesentliche Ursachen hierfür werden die starken Sektorengrenzen, die mengenbasierte Vergütungssystematik, die fehlende Steuerung und die hohe Kapazität stationärer Betten in Deutschland diskutiert (1). Die derzeitige, im Jahr 2023, vom Bundesministerium für Gesundheit angestrebte Krankenhausreform zielt darauf ab, die Versorgungsstrukturen zu verändern, um eine bedarfsgerechte, qualitativ hochwertige Versorgung der Bevölkerung auf dem Land und in der Stadt sicherzustellen. Aufgrund ihrer anderen und komplexen Struktur wurde die psychiatrische Versorgung bei den bisherigen Bestrebungen bzgl. der Krankenhausreform zunächst ausgeschlossen (2), in einer eigenen Stellungnahme zu Ende September 2023 jedoch separat berücksichtigt (3). Die geschilderte Problematik trifft auch in diesem Feld zu. So wies Deutschland im Jahr 2017 128 stationäre Betten pro 100.000 Einwohnerinnen und Einwohner für die Behandlung psychisch Erkrankter auf, den zweithöchsten Wert in der Europäischen Union; im EU-Durchschnitt waren es 2018 nur 73 (4).

Die Versorgung von Menschen mit einer psychischen Erkrankung erfordert häufig die Einbindung multiprofessioneller Akteure verschiedener Sektoren und liegt in der Zuständigkeit mehrerer Sozialgesetzbücher (SGB) (5). Die ausgeprägte Fragmentierung führt daher besonders in der Versorgung von Menschen mit psychischen Erkrankungen zu einem komplexen System, das insbesondere Menschen mit schweren sowie multiplen psychischen (und somatischen) Erkrankungen die notwendige personelle sowie die Behandlungs- und Informationskontinuität deutlich erschwert (5). Dabei benötigt es vor allem hier neue Ansätze, das für Menschen mit psychischen Störungen zu rigide und unflexible System einer starren Trennung in stationäre und ambulante Versorgung aufzuweichen. Alternative Behandlungsformen, wie stationsäquivalente Behandlung (Stäb), Assertive Community Treatment (ACT) oder akut-ambulante Behandlung (AAB) finden bei der Behandlung von Menschen mit psychischen Erkrankungen daher zunehmend Anwendung. Dennoch werden weiterhin in Deutschland sehr viele Menschen mit psychischen Erkrankungen vorwiegend stationär versorgt (6).

Vor dem Hintergrund dieser Herausforderungen und den positiven Erfahrungen (7-11) der Psychriebudgets in Schleswig-Holstein (12) sowie verschiedenen Verträgen der Integrierten Versorgung (13, 14) wurde 2012 ein Gesetz zur Weiterentwicklung der Versorgung psychisch erkrankter Menschen basierend auf Modellvorhaben, der § 64b SGB V, verabschiedet. Laut Gesetzgebung zielt dieser auf „die Weiterentwicklung der Versorgung psychisch kranker Menschen“ ab, „die auf eine Verbesserung der Patientenversorgung oder der sektorenübergreifenden Leistungserbringung ausgerichtet ist, einschließlich der komplexen psychiatrischen Behandlung im häuslichen Umfeld“ (15). Die inhaltliche Ausgestaltung dieses Gesetzes wird durch die Gesetzgebung nicht geregelt. Seit 2013 wurden verschiedene Modellvorhaben etabliert. Fast alle bis heute etablierten Modellvorhaben wurden ausschließlich an psychiatrischen (Erwachsene und/oder Kinder- und Jugendliche) bzw. psychosomatischen Kliniken umgesetzt. Dabei verfolgten diese ein globales Behandlungsbudget, d.h. ein krankenhausbasiertes Gesamtbudget für alle Formen der stationären und krankenhausbasierten ambulanten Versorgung basierend auf eine vorab definierte Anzahl an zu behandelnden Personen unabhängig von der Leistungsmenge (7, 16, 17). Zwei Ausnahmen bilden das Modellvorhaben am Zentralinstitut für Seelische Gesundheit in Mannheim sowie das über den Innovationsfonds des Gemeinsamen Bundesausschusses (G-BA) finanzierte multizentrische Projekt Gemeindepsychiatrische Basisversorgung (18).

Modellvorhaben, die im Rahmen von § 64b SGB V umgesetzt wurden, müssen gemäß § 65 SGB V von unabhängigen Sachverständigen wissenschaftlich evaluiert werden. Die gesetzlichen Krankenkassen (Gesetzliche Krankenversicherung (GKV)) oder ihre Verbände haben eine „wissenschaftliche Begleitung und Auswertung der Modellvorhaben im Hinblick auf die Erreichung der Ziele der Modellvorhaben [...] nach allgemein anerkannten wissenschaftlichen Standards zu veranlassen“ (19). Die „Bundesweit einheitliche Wissenschaftliche Evaluation von Modellvorhaben nach § 64b SGB V“ (EVA64) (20) startete im Jahr 2015 und ist bis heute die einzige Studie in Deutschland mit einer solch hohen Anzahl an beteiligten gesetzlichen Krankenkassen (>70 Krankenkassen). Insgesamt werden in der Evaluationsstudie EVA64 18 Modellvorhaben betrachtet. Die Datengrundlage für EVA64 stellen ausschließlich GKV-Routinedaten der beteiligten Krankenkassen dar. Im Jahr 2025 wird die EVA64-Studie abgeschlossen sein. Bisherige Ergebnisse weisen auf eine Verkürzung vollstationärer Behandlungstage in Modellvorhaben bei einer erhöhten Inanspruchnahme teilstationärer Behandlungstage und/oder ambulanter Leistungen im Krankenhaus hin (21-23). Weitere Ergebnisse werden nach Abschluss der Studie erwartet. Eine Folge-Studie (EVA64.2) untersucht weitere bis Juni 2026 neu abzuschließende Modellvorhaben sowie die weitere Entwicklung bestehender Verträge (24).

Im Folgenden wird der auf die EVA64-Studie bezogene Einsatz von GKV-Routinedaten kritisch bewertet und es werden Empfehlungen für den Einsatz dieser Daten im Bereich psychiatrischer Versorgungsforschung und allgemein abgeleitet. (1) Zunächst wird die EVA64-Studie und der Einsatz von GKV-Routinedaten in dieser beschrieben und eingeordnet. (2) Im darauffolgenden Abschnitt werden die einzelnen Outcomes der EVA64-Studie dargestellt und bewertet, (3) um danach Kriterien, Stärken und Einsatzmöglichkeiten von GKV-Routinedaten darzustellen. (4) Abschließend werden Empfehlungen für die Weiterentwicklung von GKV-Routinedaten als Bewertungsgrundlage formuliert.

Evaluation psychiatrischer Modellvorhaben mittels GKV-Routinedaten

Zunächst wird nun der Einsatz von GKV-Routinedaten in der EVA64-Studie allgemein beschrieben und diese eingeordnet. Dabei werden vor allem deren Potenziale und Limitationen im Kontext der Evaluation psychiatrischer Modellvorhaben adressiert; für allgemeine Ausführungen zur wissenschaftlichen Nutzung von GKV-Daten sei auf Standardliteratur verwiesen (25).

Die EVA64-Studie ist eine **GKV-routinedatenbasierte kontrollierte Kohortenstudie** (20). Zum Nachweis der Effektivität von Versorgungsmodellen sind prinzipiell randomisierte kontrollierte Studien als Goldstandard zu betrachten (26). Im EVA64 Studiendesign war jedoch eine Randomisierung in Modellvorhaben und kein Modellvorhaben nicht möglich, da die Evaluation erst nach dem Start eines Modellvorhabens einsetzte und keinen Einfluss auf den Abschluss eines Modellvertrages hatte. Auch eine generell mögliche Clusterrandomisierung und der Einsatz eines Stepped-wedged-Designs (27) ist mit GKV-Routinedaten generell möglich, war jedoch in der EVA64-Studie aufgrund der nicht möglichen Einflussnahme auf den Start eines Modellvorhabens ausgeschlossen. Der Abschluss eines Modellvorhabens war nicht Gegenstand eines Randomisierungsprozesses, sondern Resultat von Verhandlung.

Unter **GKV-Routinedaten** werden die Abrechnungsdaten der gesetzlichen Krankenkassen verstanden. Diese beinhalten neben Stammdaten wie Geschlecht, Geburtsdatum, Wohnort und Versichertenstatus Angaben zu Diagnosen, stationär und ambulant erbrachten Leistungen,

Arbeitsunfähigkeit (AU) sowie zu ärztlich verordneten Arznei-, Heil und Hilfsmittel. Diese Sekundärdaten dienen primär der Dokumentation und Abrechnung von Leistungen der Krankenkassen und werden zunehmend sekundär im Rahmen wissenschaftlicher Untersuchungen genutzt (28). Die Verwendung dieser standardisierten Datenbasis ermöglicht es, im Rahmen von EVA64 über alle Modellvorhaben hinweg ein einheitliches Studiendesign umzusetzen. Da Patientinnen und Patienten nicht explizit einwilligen müssen und nicht von der Teilnahme ausgeschlossen werden, ist ein möglicher Selektionsbias bei der Selektion der Studienteilnehmenden stark minimiert. Auch Versicherte, die mit Primärdatenerhebungen häufig nicht erreicht werden können, sind eingeschlossen. Ausgenommen davon sind privat Krankenversicherte und Menschen ohne Krankenversicherung (z.B. aufgrund von Obdachlosigkeit) sowie gesetzlich Krankenversicherte von nicht an der Evaluation teilnehmenden Krankenkassen. Innerhalb der Modellregionen sind bei den an der Evaluation beteiligten Kassen mindestens 85 Prozent aller Patientinnen und Patienten versichert, so dass unter Berücksichtigung der genannten Gruppe konservativ die Gesamtabdeckung auf 70 bis 75 Prozent geschätzt werden kann. Aufgrund der großen Fallzahlen sind auch umfassende Subgruppenanalysen möglich. Insgesamt konnten in der EVA64-Studie mehr als 131.800 psychiatrisch versorgte Patientinnen und Patienten (davon über 16.200 Kinder und Jugendliche) über mindestens drei Jahre und mehr als 50.200 (davon über 5.800 Kinder und Jugendliche) über mindestens sieben Jahren beobachtet werden.

Eine wesentliche Stärke der Evaluation ist zudem, dass eine **Vergleichsgruppe** über ein zweistufiges Matching herangezogen wird. Zunächst wurden vergleichbare Kliniken über ein standardisiertes Verfahren (29) ermittelt. Dazu wurden gewichtete Ähnlichkeitsscores basierend auf a priori definierten Kriterien (Angaben aus Strukturierten Qualitätsberichten gemäß § 137 SGB V sowie zugeordneter Daten des Bundesinstituts für Bau-, Stadt und Raumforschung (INKAR)) je Modellvorhaben ermittelt. Im zweiten Schritt wurden zu den behandelten Personen in den Modellvorhaben vergleichbare Menschen der Regelversorgung identifiziert. Ein Lost-to-Follow-up („in der Nachbeobachtung verlorengegangen“), ein häufiges Problem in primärdatenbasierten Studien, erfolgt hier nur, wenn Personen die Krankenkasse gewechselt haben; was aufgrund der anonymen Datenlieferung jeder Krankenkasse nicht beobachtet werden kann. Dies tritt bei der hier betrachteten Personengruppe aber selten auf (ca. 3 Prozent pro Jahr). Auch ein differentieller Informationsbias (d.h. ein zwischen Gruppen unterschiedliches „Messverfahren“ zur Bestimmung des Endpunktes) kann mit GKV-Daten ausgeschlossen werden.

Das Zusammenführen von Daten aus **über 70 gesetzlichen Krankenkassen** stellt methodisch eine Herausforderung dar (30). Trotz gesetzlicher Vorgaben und dadurch gut definierter Datenfelder und Prozesse benötigt eine kassenübergreifende Entgegennahme, individuelle Prüfung und zusammenführende Aufbereitung erhebliche zeitliche und personelle Ressourcen. Auch in langjährigen Evaluationen, wie EVA64, ergeben sich über die Zeit Änderungen bzgl. Dokumentation und Klassifizierungssystemen (ICD, EBM, OPS, ATC) kassenspezifisch und allgemein, die im Projektverlauf kontinuierlich überprüft werden müssen (30). Zudem erfordern Fusionen zwischen Krankenkassen und strukturelle Veränderungen weitere Abstimmungen. Dennoch ermöglicht die langjährige Perspektive von EVA64 den Aufbau von Routinen und Erfahrungen für eine kompetente und zeitnahe Aufbereitung von GKV-Routinedaten einer Vielzahl an gesetzlichen Krankenkassen, analog der NAKO Gesundheitsstudie (31).

Die **Repräsentativität** der Daten ist ihr wesentlicher Vorteil, besonders da auch Gruppen von Personen inkludiert werden können, die für die Forschung sonst schwer zu erreichen sind (z.B. schwer erkrankte oder kognitiv eingeschränkte Menschen). Die Repräsentativität ist jedoch auf gesetzlich krankenversicherte Personen beschränkt und schließt daher Menschen ohne Krankenversicherung bzw. mit einer privaten Krankenversicherung nicht ein. Die GKV-

Routinedatengrundlage ermöglicht es in EVA64, Daten von Versicherten sektoren- und einrichtungsübergreifend zu betrachten. Dies erfolgt unabhängig vom Behandelnden und sowohl im stationären also auch im ambulanten Bereich. Jedoch können, wie bei allen Sekundärdatenanalysen, nur Daten betrachtet werden, die auf Grundlage der Leistungsbereiche gemäß SGB V erhoben werden. Informationen aus anderen Sozialgesetzbüchern, wie „Rehabilitation und Teilhabe von Menschen mit Behinderungen“ (SGB IX) oder „Soziale Pflegeversicherung“ (SGB XI) sind mit den GKV-Routinedaten nicht abgebildet. Nicht-Inanspruchnahme trotz medizinischen Behandlungsbedarfs (Inanspruchnahmebias) und selbst finanzierte Leistungen, wie frei verkäufliche Arzneimittel, werden ebenfalls nicht erfasst.

Ein wichtiger Aspekt, der anhand von GKV-Daten nur sehr eingeschränkt untersucht werden kann, sind **patient:innenzentrierte und -berichtete Outcomes (PRO)**, wie Lebensqualität, Behandlungszufriedenheit, diagnostische Maße sowie Outcome der Therapie. Eine damit verbundene weitere erhebliche Einschränkung ist die begrenzte Abbildbarkeit der Erkrankungsschwere psychischer Erkrankungen in den GKV-Routinedaten. Im Matching der Patientinnen und Patienten wurde versucht, diese durch das Matching der Parameter *Kombination aller Indexdiagnosen des Referenzfalls*, Unterscheidung von der Klinik aus betrachteten *neuen und bekannten Patientinnen und Patienten* sowie *Anzahl von voll-, teilstationären und ambulanten Behandlungsfällen im Jahr vor Studieneintritt* zu erfassen und zwischen den Gruppen vergleichbar zu machen. Dennoch verbleibt eine Unsicherheit bezüglich der vergleichbaren Erkrankungsschwere. Auch die Untersuchung einer Veränderung in der Symptomlast bzw. Erkrankungsschwere ist ein wichtiger Parameter, welcher mit den GKV-Routinedaten derzeit nur sehr eingeschränkt (bspw. über ICD-10 Kodierung, u.a. durch Übergang von einer leichten (F32.0) zu einer mittelgradigen (F32.1) depressiven Episode) untersucht werden kann. Diese hängt jedoch stark von der Qualität der Diagnosekodierung ab. Generell ist es bei Studien zu Menschen mit psychischen Erkrankung schwer, ausschließlich anhand von verfügbaren Kodierungen und ggf. Surrogatparametern auf die Krankheitsschwere zu schließen. Dafür benötigt es dringend die Verfügbarkeit von diagnostischen Maßen. Eine Verbindung dieser mit GKV-Routinedaten würde eine Bereicherung sein.

Der **Bedarf** einer Versorgung von Menschen mit psychischen Störungen ist generell schwer, und insbesondere mit GKV-Routinedaten, kaum erfassbar. Diese Daten gelten vorrangig für die Abrechnung und Dokumentation in der GKV und werden nicht für Zwecke der Evaluation oder Versorgungsforschung erhoben (32). Nicht erlösrelevante Daten bieten einen geringeren Anreiz für eine vollständige Dokumentation. Zum Beispiel hat die Kodierung der Diagnosen, im Gegensatz zum stationären Setting, im ambulanten Bereich kaum Einfluss auf die Abrechnung, so dass hier von einer geringeren Dokumentationsqualität ausgegangen werden muss (33). Diese Einschränkungen sind allgemeiner sekundärdatenbezogener Art und nicht begrenzt auf die Gruppe der Menschen mit einer psychischen Erkrankung.

Ein weiterer Aspekt der GKV-Routinedaten ist der nicht unerhebliche **Zeitverzug** (34). Berichte aus der EVA64-Studie, die beispielweise im Jahr 2020 erstellt wurden, beruhen auf Daten, die 2019 geliefert wurden und umschließen abgeschlossene Fälle bis Ende 2018. Trotz dieses Verzuges ist bei Betrachtung von epidemiologischen Studien mit Routinedaten bzw. einem Linkage dieser (z.B. die NAKO Gesundheitsstudie (35)), Registern (z.B. klinische Krebsregister oder TREATgermany (36)) oder klinischen Studien der Weg über die einzelnen gesetzlichen Krankenkassen oft der schnellste. Auch hat die COVID-19-Pandemie gezeigt, dass GKV-Routinedaten für Zwecke des Forschungs- und Versorgungsmonitorings beschleunigt verfügbar gemacht werden müssen (34). Da die Datenaufbereitung von über 70 gesetzlichen Krankenkassen sehr umfassend ist und alle datenliefernden Stellen einbezogen werden müssen, können in EVA64 nur einmal pro Jahr Daten neu zur Verfügung gestellt werden. Aufgrund des Datenschutzes erfolgen die Datenlieferungen zudem in

jedem Jahr komplett neu, um eine anonymisierte Lieferung zu gewährleisten. Dies bedeutet auf Seiten der gesetzlichen Krankenkassen eine umfangreiche Sammlung von Daten über einen längeren Zeitraum (Daten von 2011 bis teilweise 2022). In der EVA64-Studie werden die Daten zunächst bei den Krankenkassen für den gesamten Zeitraum kontinuierlich gespeichert. Einmal im Jahr bis zum Ende der Laufzeit der Studie wird der aktuelle Stand an die Auswertenden in anonymisierter Form zur Verfügung gestellt. Die Speicherung von GKV-Routinedaten über einen so langen Zeitraum wurde durch die rechtliche Grundlage des § 64b SGB V und den damit verbundenen gesetzlichen Auftrag zur Evaluation ermöglicht. Die für die Auswertenden anonymisierten Daten erlauben gleichwohl eine patient:innenbezogene längsschnittliche Betrachtung von Versorgungsverläufen. Eine anonymisierte Datenlieferung erspart dafür einen Antrag nach § 75 SGB X (Übermittlung von Sozialdaten für die Forschung und Planung) und erfordert ggf. nur eine Anzeige nach § 80 SGB X (Datenverarbeitung im Auftrag).

Bewertung der Outcomes in EVA64

Im Folgenden werden die einzelnen Outcomes der EVA64-Studie kritisch bewertet und das Risiko für Bias, die Qualität der Abbildbarkeit mit GKV-Routinedaten sowie die Relevanz für den Versorgungskontext (Versorgungsforschung zu Menschen mit psychischen Erkrankungen) zusammengefasst (Error: Reference source not foundTabelle 1).

Bestimmte Outcomes können auf Basis von GKV-Routinedaten sehr gut abgebildet werden, während andere Outcomes diversen Einschränkungen unterliegen. Gut abbildbar sind die **voll- und teilstationären psychiatrischen Behandlungsdauern**. Diese sind in den GKV-Routinedaten lückenlos auf Grundlage der Abrechnungs- und Dokumentationsrelevanz mit Haupt- und Nebendiagnosen erfasst. Auf der anderen Seite ist die in der EVA64-Studie gemessene Behandlungsdauer kaum mit anderen Publikationen vergleichbar, in denen eine bestimmte Gruppe, z.B. definiert anhand bestimmter Diagnosen, rekrutiert wird. In EVA64 berechnet sich dieses Outcome auf Basis aller pro Klinik in die Studie eingeschlossenen Personen. Das heißt, hier werden auch Patientinnen und Patienten gänzlich ohne stationären Aufenthalt erfasst. Dies ist für die Bewertung des Erfolgs von Modellvorhaben wesentlich, um auch eine mögliche vollständige Vermeidung von stationären Aufenthalten zu berücksichtigen.

Auch die Inanspruchnahme **ambulanter Leistungen** (aus dem Bereich des SGB V) ist über GKV-Routinedaten gut erfasst. Bei diesem Outcome wird die Häufigkeit der Tage mit ambulantem Kontakten gezählt, ohne eine qualitative Bewertung vorzunehmen. Die Zuordnung zu psychischen Erkrankungen muss häufig über kodierte Diagnosen, Facharztgruppen oder einzelne Abrechnungen erfolgen. Die Kodierung von Diagnosen im ambulanten Bereich unterliegt jedoch nicht, wie oben erwähnt, der Dokumentationslogik und ist nicht erlösrelevant. Zudem können zum Beispiel Diagnosen teilweise kodiert sein, die für die Behandlung sekundär sind, aber z.B. die Medikation beeinflussen. Eine Einteilung in Haupt- und Nebendiagnosen, wie im stationären Bereich, gibt es im ambulanten Bereich nicht. Dies trifft auch auf den **ambulantem Bereich am Krankenhaus** (wie bspw. die Psychiatrische Institutsambulanz (PIA)) zu. Die PIA-Dokumentationsverordnung von 2012 erlaubte bundesweit einheitlich lediglich die Abbildung von PIA-Leistungen ohne vs. mit Kontakt von Ärztin oder Arzt bzw. Psychologin oder Psychologe und mit vs. ohne aufsuchende Behandlung. Einzelne Bundesländer bzw. Einrichtungen rechneten zwar mit Einzelleistungsvergütung ab, ein bundesweit einheitliches Vorgehen war in EVA64 jedoch nur über die vier bundeseinheitlichen Parameter möglich. Die neue PIA-Dokumentationsvereinbarung von 2018 erweiterte die Kodierung um die

Leistungsart, den Umfang und die Berufsgruppen. Diese neue Dokumentation wird in der EVA64.2-Studie aufgegriffen, um die PIA-Behandlung qualitativ besser bewerten zu können. Die ambulanten Kontakte im Krankenhaus über die PIA werden über PIA-Leistungsschlüssel bzw. Entgeltschlüssel in Bundesländern mit Einzelleistungsvergütung taggenau erfasst (PIA-Doku-Vereinbarung bundesweit bzw. separate Landesvereinbarungen) (37). Im niedergelassenen Bereich rechnen bspw. Psychotherapeuten neben der Grundpauschale in der Regel das psychotherapeutische Gespräch oder gruppentherapeutische Leistungen zusätzlich je Sitzung ab. Somit gehen wir davon aus, dass eine Frequenzanalyse im ambulanten Bereich auch valide umgesetzt werden kann und dass Pauschalisierung, besonders im psychiatrischen/psychotherapeutischen Bereich, die Häufigkeit der Kontakt nur gering unterschätzt. Weiterhin wäre eine mögliche Unterschätzung in beiden Gruppen (Modellklinik und Regelversorgung) gegeben.

Während in anderen Studien die **Behandlungskontinuität** als Behandlung im gleichen Setting bzw. vom gleichen Behandelnden bzw. Team definiert wird (38, 39), kann dies so mit GKV-Routinedaten nicht abgebildet werden. Daher wurde in EVA64 als Behandlungskontinuität erfasst, ob innerhalb eines definierten Zeitraums nach stationärer Behandlung eine ambulante Behandlung (im Krankenhaus oder außerhalb) erfolgt ist. Modellvorhaben legen jedoch häufig ihren Fokus auf Behandlungskontinuität im Sinne der ersten Definition. Auf der anderen Seite kann auch die mit GKV-Routinedaten abgebildete Behandlungskontinuität die stärkere Einbindung der PIA messen und diese als kontinuierliche Behandlung ausweisen.

Das Outcome **Kontaktabbrüche** in EVA64 misst fehlende Kontakte zum psychiatrischen Versorgungssystem innerhalb eines bestimmten Zeitraumes für Menschen mit schweren psychischen Störungen. Schwere psychische Störungen sind als schwerwiegende psychische Diagnosen mit einer bestimmten Dauer definiert (40). Das Outcome Kontaktabbrüche ist generell mit den GKV-Routinedaten gut messbar, jedoch sind die Fallzahlen von Menschen mit schweren psychischen Störungen oft sehr gering. Im Outcome **Arzt- und Krankenhaus-Hopping** wird auf der anderen Seite der Anteil der Versicherten mit mindestens drei unterschiedlichen psychiatrischen Leistungserbringenden in einem Bereich je Sektor untersucht. Dieses Outcome kann mit GKV-Routinedaten gut abgebildet werden; jedoch betrifft Arzt- und Krankenhaus-Hopping generell sehr wenige Menschen mit psychischen Diagnosen.

Die Abbildung von **Arbeitsunfähigkeit (AU)** mit GKV-Routinedaten ist generell sehr gut möglich, da alle gemeldeten AU-Fälle mit Diagnosen erfasst sind. In der EVA64-Studie werden diese zum einen über die Anzahl der stationären Behandlungstage plus vertragsärztlich dokumentierter AU-Scheine berechnet. Für die vertragsärztlichen AU-Bescheinigungen gibt es in diesem Rahmen zwei Einschränkungen. Erstens können Aussagen über AU nur für einen Teil der Menschen mit psychischen Erkrankungen getroffen werden. Nicht für alle Versicherte sind AU-Bescheinigungen relevant, wie zum Beispiel verrentete Personen, Studierende oder Menschen ohne Erwerbseinkommen. Zur Abschätzung der Teilpopulation der Erwerbstätigen unter den Mitgliedern gibt es bei den gesetzlichen Krankenkassen das Datenfeld „krankengeldanspruchsberechtigt“. Jedoch konnte dies in EVA64 nicht von allen beteiligten Kassen geliefert werden. Daher wurde in EVA64 auf den Versichertenstatus „Mitglied“ gefiltert und nur diese Population betrachtet (d.h. selbst versichert und nicht-verrentet oder familienversichert). Dies kann zu einer Überschätzung der erwerbsfähigen Personen führen, da auch Personen mit Versichertenstatus „Mitglied“ ggf. nicht erwerbsfähig oder -tätig sind. Zweitens ist davon auszugehen, dass gerade bei Menschen mit psychischen Erkrankungen kurze Krankmeldungen als „andere somatische Erkrankungen“ deklariert oder über arbeitsinterne „Krank-ohne-Schein“-Tage abgefangen werden, die dann in den GKV-

Routinedaten nicht ersichtlich sind. Für einen Vergleich von Modell- und Regelversorgung könnte hier ein (kleiner) systematischer Unterschied entstehen, da bei einer möglichen Verschiebung der Behandlung in den ambulanten Bereich in der Modellversorgung diese Einschränkungen eher die Modellkliniken betreffen würden. Daher könnte der mögliche Effekt einer Verringerung einer AU im Modellvorhaben etwas unterschätzt werden, sollte die Behandlung im Modellvorhaben in den ambulanten Bereich verlagert worden sein.

Die **Wiederaufnahmerate**, also der Anteil der Personen, die innerhalb von einem Jahr nach stationärer Entlassung erneut stationär aufgrund einer psychischen Diagnose aufgenommen wurden, kann mit GKV-Routinedaten sehr gut abgebildet werden. Zur Ermittlung von **Komorbiditäten** wurde der Elixhauser Score herangezogen (41). Dieser wurde ursprünglich zur Abschätzung der Wahrscheinlichkeit, im Krankenhaus zu versterben, ermittelt (42). Für diesen Parameter werden definierte somatische und psychische Erkrankungen basierend auf ICD-10 Kodierungen in den definierten Clustern erfasst und die durchschnittliche Anzahl der Cluster des Elixhauser Scores gegenübergestellt. Dieser Parameter lässt sich generell sehr gut auf GKV-Routinedaten anwenden. Die Relevanz für den Versorgungskontext der psychischen Erkrankungen ist jedoch eingeschränkt, da im Elixhauser Score nur wenige psychische und überwiegend somatische Diagnosen berücksichtigt sind. Weiterhin wurde die **Mortalität** in EVA64 untersucht. Auch diese lässt sich sehr gut mit GKV-Routinedaten abbilden. Die Gründe für eine erhöhte oder niedrigere Mortalität lassen sich jedoch nicht mit den Daten aufdecken, denn die Gründe für ein Versterben bzw. die Todesursache sind nicht hinterlegt.

Um **Erkrankungsprogression** mit GKV-Routinedaten abzubilden, wurden in EVA64 drei Pfade definiert, die generell gut anhand der ICD-10 Kodierung – unabhängig von deren Qualität – abgebildet werden können. Diese umfassen erstens die Entwicklung von einer leichten oder mittelgradigen depressiven Episode zu einer schweren depressiven Episode; zweitens die Entwicklung von einer depressiven Episode zu einer rezidivierenden depressiven Störung; drittens die Entwicklung einer psychischen oder Verhaltensstörung durch psychotrope Substanzen. Diese sind die einzigen Erkrankungsprogressionen psychischer Erkrankungen, die mit GKV-Routinedaten (mittels ICD-10-Kodierung) abgebildet werden können. Die Dokumentation der Erkrankungsschwere fehlt in den GKV-Routinedaten. Diese ist jedoch notwendig, um eine „wahre“ Progression abbilden zu können. Hier kann vermutlich die derzeit sich in Arbeit befindende ICD-11 Struktur, die besser an den diagnostischen und statistischen Leitfaden psychischer Störungen (DSM-V) angepasst werden soll (43), langfristig Verbesserung erzielen. Durch die Anwendung von Stamm- und Erweiterungs-codes soll eine bessere und genauere Erfassung der Symptomatik erreicht werden (44). Entscheidend ist und wird die Qualität der Diagnosestellung und Kodierung bleiben.

Weiterhin wurde versucht, **Leitlinienadhärenz** mit GKV-Routinedaten abzubilden. Dazu wurden zu Beginn der Studie EVA64 basierend auf einem Projekt der Deutschen Gesellschaft für Psychiatrie, Psychosomatik und Nervenheilkunde (DGPPN) zur Entwicklung von Qualitätsindikatoren (QIs) (45) für vier psychiatrische Krankheitsbilder verschiedene QIs anhand von GKV-Routinedaten herangezogen. Aus den 18 in der Publikation vorgestellten QIs wurden 10 final aufgenommen, da in den restlichen QIs verschiedene Aspekte, wie Dokumentation in der PIA oder Medikation im Krankenhaus, mit GKV-Routinedaten nicht abbildbar waren. Zudem ist zu unterstreichen, dass individuelle Entscheidungen basierend auf Erkrankungsschwere oder personenindividuellen Merkmalen bzw. Präferenzen der Patientinnen und Patienten mit GKV-Routinedaten generell nicht abgebildet werden können. Leitlinienadhärenz kann daher mit den derzeit vorliegenden GKV-Routinedaten nur schwer abgebildet werden.

Die **Kosten**, die in der EVA64-Studie abgebildet werden konnten, umfassen ausschließlich direkte Kosten aus Perspektive der GKV. Weitere Kosten konnten im Rahmen dieser Studie nicht erhoben werden. Um Kosten aus einer gesellschaftlichen Perspektive abbilden zu können, sind weitere Datenquellen, wie z.B. aus Befragungen der Betroffenen und deren Angehörigen, notwendig. Dieser Ansatz wurde in den Studien PsychCare (46) und EVA_TIBAS (47) verfolgt. Vergleiche zwischen Angaben aus GKV-Routinedaten und Primärbefragungen zeigen, dass Angaben zu stationären Aufenthalten relativ gut von den Betroffenen (bezogen auf sechs Monate vor Befragung) erinnert wurden, während Angaben zu ambulanten in Anspruch genommenen Leistungen große Abweichungen zwischen Selbstangaben und Angaben über GKV-Routinedaten aufzeigen (48). Kosten aus GKV-Routinedaten stellen eine umfangreiche und aus Sicht der GKV lückenlose Datenquelle dar, die gut herangezogen werden kann. Einschränkend bleibt zu erwähnen, dass nur das ausgewertet bzw. aufgezeigt werden kann, was dokumentiert ist. Ausgleichszahlungen aus Vorjahren, d.h. Budgetüberschüsse oder -defizite vom Vorjahr im Vergleich zum vereinbarten Gesamtbudget, konnten für einige Modellvorhaben nicht vom regulären Budget getrennt werden, so dass hier Unsicherheiten und ggf. Überschätzungen der Kosten aufgetreten sind. Diese können jedoch kommuniziert und entsprechend in den Auswertungen diskutiert werden. Bei der Betrachtung der **Kosten-Effektivität** auf Basis von GKV-Routinedaten werden den Kosten Effektivitätsergebnisse gegenübergestellt. In EVA64 wurden die Gesamtkosten den vollstationären Behandlungstagen sowie den AU-Tagen gegenübergestellt. Zu berücksichtigen ist für dieses Outcome, dass beide Seiten, besonders die stationären Behandlungstage als Hauptfaktor der Gesamtkosten, einander bedingen.

Die Aussagekraft von Ergebnissen basierend auf GVK-Routinedaten hängt von deren Dokumentationsqualität ab. Da in Modellvorhaben ein prospektives Gesamtbudget vorliegt, kann ein geringerer Anreiz bestehen, sämtliche Leistungen zu dokumentieren. Die Evaluation kann jedoch nur das betrachten, was auch dokumentiert wurde. Zudem ist es wichtig, dass Umsteuerungen in alternative Behandlungsformen, z.B. ACT oder AAB, ebenfalls dokumentiert und diese Dokumentation den Evaluierenden zur Verfügung gestellt werden. Nur so können diese neuen Formen adäquat berücksichtigt werden. Generell werden die dargestellten Outcomes in der EVA-Studie zunächst quantitativ mittels Tagen, Ereignissen, Raten oder Kosten gemessen, um eine angenommene Veränderung einer Inanspruchnahme zu betrachten. Eine qualitative Bewertung der Modellvorhaben kann mit GKV-Routinedaten nicht vorgenommen werden und bedarf andere Studien.

Tabelle 1 fasst die kritische Bewertung der Outcomes mit einer Einschätzung für ein Risiko für Bias, der Stärke der Abbildung mit GKV-Routinedaten allgemein und einer Einschätzung für den Versorgungskontext (Menschen mit psychischen Erkrankungen) aus der Sicht der Autorinnen und Autoren zusammen.

[Hier Tabelle 1 einfügen.]

Kriterien, Stärken und Einsatzgebiete GKV-Routinedaten

Basierend auf den Erfahrungen und Ausführungen in EVA64 wurden von dem Team der beteiligten Wissenschaftlerinnen und Wissenschaftler Kriterien zu Studien mit GKV-Routinedaten für die Versorgungsforschung im Bereich psychische Gesundheit abgeleitet (Tabelle 2). Die Stärken und primären Einsatzgebiete in Abgrenzung zu Primärdaten wurden zudem visuell aufbereitet (Abbildung

1). GKV-Routinedaten eignen sich sehr gut zum Monitoring von Veränderungen der Inanspruchnahme über die Zeit bzw. mit dem Eintritt einer neuen Versorgungsform, der Betrachtung aller versorgten Versicherten sowie der sektorenübergreifenden längsschnittlichen Betrachtung von Versorgungspfaden. Für die Betrachtung der Qualität der Behandlung sind Studien mit Primärdaten zur Abbildung der Sicht von Patientinnen und Patienten, Leistungserbringenden und Angehörigen essentiell.

[Hier Tabelle 2 einfügen.]

GKV-Routinedaten als Grundlage für Versorgungsforschung haben ihre Berechtigung und Schwerpunkte. Diese eignen sich primär dazu, das Versorgungsgeschehen und die Inanspruchnahme abzubilden. Die Abbildung von Behandlungsergebnissen und der Präferenz von Patientinnen und Patienten, Leistungserbringenden und Angehörigen ist jedoch nur mit der Untersuchung von Primärdaten möglich. Die PsychCare-Studie, eine vom Innovationsfonds finanzierte multizentrische Studie, untersuchte dazu Modellvorhaben nach § 64b SGB V prospektiv und kontrolliert und umfasste u.a. eine Primärdatenerhebung mittels Fragebögen sowie eine Datenintegration als Datenlinkage von Fragebogendaten und GKV-Routinedaten (46). Datenlinkage ist der beste Weg, um die Stärken von GKV-Routinedaten und weiteren Datenquellen, wie die der Fragebogen, miteinander zu verbinden und Synergien zu nutzen, welches großes wissenschaftliches Potential zur Verbesserung der Gesundheitsversorgung aufweist (34) (Abbildung 1).

[Hier Abbildung 1 einfügen]

Entsprechende Standards eines Datenlinkages von Primär- und Sekundärdaten sind formuliert (49). Erfahrungen aus der PsychCare-Studie zeigen, dass auch in der Population von Menschen mit psychischen Erkrankungen ein Linkage durchgeführt werden konnte und die Akzeptanz bei den Patientinnen und Patienten mit rund 90 Prozent vergleichbar hoch zu anderen Datenlinkage-Studien, wie der NAKO Gesundheitsstudie, war (50). Eine Limitation in der PsychCare-Studie war jedoch, dass nur eine begrenzte Anzahl gesetzlicher Krankenkassen teilgenommen hatte, was die Größe der gelinkten Datenbasis stark reduzierte (50). Eine Einwilligung zur wissenschaftlichen Nutzung der GKV-Routinedaten ist jedoch derzeit zwingend notwendig, bevor Daten an eine wissenschaftliche Einrichtung übermittelt werden können. Dies setzt im Vorfeld viel Überzeugungsarbeit auf Seiten des Studienteams und Kooperationsbereitschaft bzw. Zeit und Priorität auf Seiten der Krankenkassen voraus (50). Ein Datenlinkage als zukünftiger Goldstandard sollte, falls möglich, unabhängig von der kassenindividuellen Zustimmung nach Einwilligung der Versicherten ermöglicht werden; das neue Gesundheitsdatennutzungsgesetz (GDNG) sollte hier eine entsprechende Rechtsgrundlage schaffen (34).

Konklusion: Empfehlungen für die Weiterentwicklung der Bewertungsgrundlage

Um die Effektivität, den Nutzen sowie die Kosten und Kosten-Effektivität neuer Versorgungsmodelle für die Bevölkerung bewerten zu können, ist eine begleitende Evaluation unabdingbar. Wir begrüßen daher die Gesetzgebung mit dem § 65 SGB V, die eine solche Evaluation von einem unabhängigen Sachverständigen basierend auf einem aktuell wissenschaftlichen Standard vorgesehen hat. Die Rahmenbedingung dieser Gesetzgebung wurde von den gesetzlichen Krankenkassen, die laut § 65 SGB V eine Evaluation beauftragen müssen, anhand einer Ausschreibung (Outcomes & Datenquelle) weitgehend vorgegeben und konnten von den Evaluierenden nur noch bedingt beeinflusst werden. Diese gemeinsame Ausschreibung basierend auf § 65 SGB V hat aus Sicht der Autorinnen und Autoren eine umfassende Evaluation sowie die Ausrichtung weiterer Entscheidungen/Entwicklungen basierend auf wissenschaftlicher Evidenz ermöglicht. Übergreifende Evaluationen von Modellvorhaben und neuen Versorgungsformen sollten auch in anderen Bereichen der Gesundheitsversorgung fest etabliert werden.

Basierend auf den vorab geschilderten Erfahrungen empfehlen wir ein Monitoring mit GKV-Routinedaten als wichtigen Bestandteil einer begleitenden Evaluation, wie auch eines Monitorings im Falle einer Überführung von Modellvorhaben in die Regelversorgung. GKV-Routinedaten sind die beste Grundlage für eine kontinuierliche, kosteneffiziente Evaluation der Inanspruchnahme und der Abbildung des Versorgungsgeschehens mit geringem Potential für Bias. Die aufwändige Rekrutierung in Kliniken und in allen Settings ist ressourcenintensiv und birgt Verzerrungsrisiken. GKV-Routinedaten können hier einen vergleichsweise kostengünstigen und ungefilterten Zugang zu relevanten Informationen liefern, besonders auch für schwer erreichbare Personengruppen, welche oft in der Versorgungsforschung von Menschen mit psychischen Erkrankungen vorkommen. Wir empfehlen zudem GKV-Routinedaten mit weiteren Informationen zu bereichern, um eine validere Evaluation zu ermöglichen. Es fehlen u.a. Angaben zu Symptomschwere; hier könnte die Einführung der ICD-11 Verbesserungen schaffen. Die Einführung einer neuen ICD Kodierung wird jedoch auch zu neuen Herausforderungen der Umstellung auf eine neue Kodierung (Kodierqualität etc.) und eine teils eingeschränkte Vergleichbarkeit mit ICD-10 führen. Dies könnte dazu führen, dass Diagnosen vor und nach Einführung nur bedingt vergleichbar wären. Die genaue Ausgestaltung und der Zeitpunkt der Einführung der ICD-11 sind jedoch unklar. Der ICD-Katalog oder eine geeignete Kodierung sollte aus unserer Sicht die Abbildung der Schwere einer psychischen Erkrankung ermöglichen können. Ein entsprechendes Kodiersystem sollte weiterentwickelt werden. Ein weiterer Aspekt, um den die GKV-Routinedaten ergänzt werden sollten, sind diagnostische bzw. klinische Scores, um die Erkrankungsschwere und die Vergleichbarkeit zwischen Populationen besser einschätzen zu können. Die einheitliche Dokumentation der erbrachten Leistungen wird durch die Umstellung der ICD jedoch nicht tangiert.

Um die Qualität einer Behandlung oder einer neuen Versorgungsform in GKV-Routinedaten messen zu können, ist die Abbildung von Qualitätsindikatoren bzw. die Einschätzung von Leitlinienadhärenz wichtig. Diese sind derzeit in GKV-Routinedaten nur dann abbildbar, wenn sie konkret als EBM- oder OPS-dokumentierte Leistungen adressiert werden. Die Erhebung von sektoren- und diagnoseunabhängigen Qualitätsindikatoren für Menschen mit psychischen Erkrankungen wurde im Projekt PsychCare (46) gestartet und sollte weiterverfolgt werden.

Zur Abbildung von Nutzen für Patientinnen und Patienten, Leistungserbringenden und Angehörigen oder des Behandlungsergebnisses der Versorgung ist jedoch der Einsatz von Primärdaten (über Fragebögen, Interviews, Fokusgruppen) unabdingbar. Mögliche Arten von Bias müssen beschrieben und möglichst umgangen werden. Eine kontinuierliche, bürokratiearme Form der Erhebung von Parametern zu PROMs (patient-reported outcome measures) und PREMs (patient-reported experience measures) sollte daher aus Sicht der Autorinnen und Autoren eingeführt werden, beispielsweise im Rahmen der datengestützten Qualitätssicherung. Ein Zusammenspielen beider

Datenquellen ist möglich (50), wichtig und wird von den Versicherten größtenteils begrüßt, auch von Menschen mit psychischen Erkrankungen. Ein Linkage mit Primärdaten oder anderen versorgungsnahen Daten birgt großes Potential, Kausalzusammenhänge besser zu erkennen (34).

Abschließend bleibt anzumerken, dass die Auswirkungen der EVA64-Studie und weiterer Folgestudien erkennbar sind (u.a. durch Diskussion in Politik zur Neuausrichtung psychiatrischer Versorgung, Berücksichtigung bei Vertragsverhandlungen, bei Ausrichtung der Modellversorgung etc.), jedoch weitreichende Schritte für eine Übertragung in die Regelversorgung bzw. auf das Gesamtversorgungssystem bisher ausgeblieben sind. Solange eine veränderte Finanzierung und Versorgung nur in einzelnen Modellvorhaben umgesetzt sind, wird sich die angestrebte Verbesserung für alle Menschen mit psychischen Erkrankungen nicht realisieren lassen. Daher begrüßen wir den aktuellen Diskussionsbeitrag der gesetzlichen Krankenkassen um Lehmann et al. für eine grundlegende Reform der Planung, Vergütung und Qualitätssicherung in der psychiatrischen Versorgung (51). Die drei primär genannten Ziele der sektorenunabhängigen Bedarfsplanung, der sektorenunabhängigen Vergütung und der Qualitätssicherung sowie Steuerung auf Basis von Routinedaten unterstützen wir. Hierfür müssen GKV-Routinedaten gestärkt, mit weiteren Daten angereichert und schneller verfügbar gemacht werden.

Literatur

1. Busse R, Karagiannidis C, Augurzky B, Schmitt J, Bschor T. Der Vorschlag der Regierungskommission für eine grundlegende Reform der Krankenhausvergütung. Berlin; 2023.
2. Regierungskommission für eine moderne und bedarfsgerechte Krankenhausversorgung. Grundlegende Reform der Krankenhausvergütung. Berlin; 2022 6.12.2022.
3. Regierungskommission für eine moderne und bedarfsgerechte Krankenhausversorgung. Achte Stellungnahme und Empfehlung der Regierungskommission für eine moderne und bedarfsgerechte Krankenhausversorgung Psychiatrie, Psychosomatik und Kinder- und Jugendpsychiatrie („Psych-Fächer“): Reform und Weiterentwicklung der Krankenhausversorgung. Berlin; 2023 29.09.2023.
4. eurostat. Mental health care - psychiatric hospital beds 2023 [Available from: <https://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/edn-20201009-1>].
5. Schwarz J, von Peter S, Baumeister H, Dahling V, Guhne U, Gouzoulis-Mayfrank E, et al. [DNVF-Discussion paper - Specificities, Challenges and Aims of Mental Health Service Research in Germany]. Gesundheitswesen. 2021;83(7):541-52.
6. Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen. Koordinierte Versorgung von Menschen mit psychischen Erkrankungen [Coordinated care for people with mental illness]. 2018.
7. König HH, Heinrich S, Heider D, Deister A, Zeichner D, Birker T, et al. [The regional psychiatry budget (RPB): a model for a new payment system of hospital based mental health care services]. Psychiatr Prax. 2010;37(1):34-42.
8. Roick C, Deister A, Zeichner D, Birker T, König HH, Angermeyer MC. [The regional budget for mental health care: a new approach to combine inpatient and outpatient care]. Psychiatr Prax. 2005;32(4):177-84.
9. Deister A, Zeichner D, Witt T, Forster HJ. [Changes in mental health care by a regional budget: results of a pilot Project in Schleswig-Holstein (Germany)]. Psychiatr Prax. 2010;37(7):335-42.
10. Karow A, Bock T, Daubmann A, Meigel-Schleiff C, Lange B, Lange M, et al. [The Hamburg-model of integrated care for patients with psychosis: Part 2. Results of the clinical course over 2- and 4-years of treatment]. Psychiatr Prax. 2014;41(5):266-73.

11. Mueller-Stierlin AS, Helmbrecht MJ, Herder K, Prinz S, Rosenfeld N, Walendzik J, et al. Does one size really fit all? The effectiveness of a non-diagnosis-specific integrated mental health care program in Germany in a prospective, parallel-group controlled multi-centre trial. *Bmc Psychiatry*. 2017;17(1):283.
12. Deister A, Zeichner D, Roick C. Ein Regionales Budget für die Psychiatrie. Erste Erfahrungen aus einem Modellprojekt [A Regional Budget for Psychiatry. First experiences from a model project]. *psychoneuro*. 2004;30(5):285-8.
13. Stierlin AS, Herder K, Helmbrecht MJ, Prinz S, Walendzik J, Holzmann M, et al. Effectiveness and efficiency of integrated mental health care programmes in Germany: study protocol of an observational controlled trial. *Bmc Psychiatry*. 2014;14:163.
14. Lambert M, Bock T, Daubmann A, Meigel-Schleiff C, Lange B, Lange M, et al. [The Hamburg-model of integrated care for patients with psychosis: Part 1. Rationale, treatment concept and results of the pre-study]. *Psychiatr Prax*. 2014;41(5):257-65.
15. Federal Ministry of Justice and Consumer Protection. § 64b SGB V: Modellvorhaben zur Versorgung psychisch kranker Menschen [Modell projects for the care of mentally ill people] 2012 [cited 2023. Available from: https://www.gesetze-im-internet.de/sgb_5/_64b.html].
16. Jacobs R, Chalkley M, Aragon MJ, Bohnke JR, Clark M, Moran V. Funding approaches for mental health services: Is there still a role for clustering? *BJPsych Adv*. 2018;24(6):412-21.
17. Schroder B, Flessa S. [Regional Budgets in Psychiatry: An Alternative to Hospital per Diem Charges and the New Reimbursement System? - A Case Study from the District of Dithmarschen]. *Psychiatr Prax*. 2017;44(8):446-52.
18. Mueller-Stierlin AS, Meixner F, Kohlmann A, Schumacher M, Hansel A, Pouwels M, et al. Effectiveness and cost-effectiveness of a community-based mental health care programme (GBV) for people with severe mental illness in Germany: study protocol for a randomised controlled trial. *Trials*. 2020;21(1):598.
19. Federal Ministry of Justice and Consumer Protection. § 65 SGB V: Auswertung der Modellvorhaben [Evaluation of model projects] 2012 [cited 2023. Available from: <https://www.sozialgesetzbuch-sgb.de/sgbv/65.html>].
20. Neumann A, Swart E, Häckl D, Kliemt R, March S, Kuster D, et al. The influence of cross-sectoral treatment models on patients with mental disorders in Germany: study protocol of a nationwide long-term evaluation study (EVA64). *Bmc Psychiatry*. 2018;18.
21. Baum F, Schmitt J, Seifert M, Kliemt R, Kubat D, March S, et al. Lengths of inpatient stay and sick leave of patients with mental diseases: disorder-specific effects of flexible and integrated treatment programs in Germany. *Translational psychiatry*. 2022;12(1):370.
22. Neumann A, Baum F, Seifert M, Schoffer O, Kliemt R, March S, et al. [Reduction of Days in Inpatient Care in Psychiatric Hospitals with Flexible and Integrated Treatment for Patient-Centered Care with a Global Budget - Results with Three-Year Follow-up from the Evaluation Study EVA64]. *Psychiatr Prax*. 2021;48(3):127-34.
23. Baum F, Schoffer O, Neumann A, Seifert M, Kliemt R, March S, et al. Effectiveness of Global Treatment Budgets for Patients With Mental Disorders—Claims Data Based Meta-Analysis of 13 Controlled Studies From Germany. *frontiers in Psychiatry*. 2020;11(131).
24. EVA64.2 study group. EVA64.2 - Projektbeschreibung Dresden2023 [cited 2023 19.04.2023]. Available from: https://www.uniklinikum-dresden.de/de/das-klinikum/universitaetscentren/zegv/projekte/projektbeschreibung/copy_of_eva%2064.2.
25. Swart E, Ihle P, Gothe H, Matusiewicz D. Routinedaten im Gesundheitswesen. *Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern: Hans Huber Verlag, Hogrefe AG; 2014.
26. Machin D, Campbell MJ. *The Design of Studies for Medical Research*. Chichester, West Sussex, England: John Wiley & Sons, Ltd; 2005.
27. Köberlein-Neu J, Hoffmann F. Das Stepped Wedge Design: Stufenlos regelbar? *Z Evid Fortbild Qual Gesundh wesen (ZEFQ)*. 2017;126:1-3.

28. Swart E, Gothe H, Geyer S, Jaunzeme J, Maier B, Grobe T, et al. Gute Praxis Sekundärdatenanalyse (GPS): Leitlinien und Empfehlungen [Goode Practice Secondary Data Analysis: Guidelines and Recommendations]. *Gesundheitswesen*. 2015;77:120-6.
29. Petzold T, Neumann A, Seifert M, Kuster D, Pfennig A, Weiss J, et al. [Identification of Control Hospitals for the Implementation of the Nationwide and Standardized Evaluation of Model Projects According to section sign 64b SGB V: Analysis of Data from Structured Quality Reports]. *Gesundheitswesen*. 2019;81(1):63-71.
30. March S, Zimmermann L, Kubat D, Neumann A, Schmitt J, Baum F, et al. Methodische Herausforderungen bei der Nutzung von Daten von mehr als 70 gesetzlichen Krankenkassen - Ein Werkstattbericht aus der EVA64-Studie. *Das Gesundheitswesen*. 2020;82:S4-S12.
31. German National Cohort C. The German National Cohort: aims, study design and organization. *Eur J Epidemiol*. 2014;29(5):371-82.
32. Veit C, Bierbaum T, Wesselmann S, Stock S, Heidecke CD, Apfelbacher C, et al. [Routine Practice Data for Health Care Analyses: Part 3 of the Manual]. *Gesundheitswesen*. 2022.
33. IGES Institut GmbH. Bewertung der Kodierqualität von vertragsärztlichen Diagnosen Berlin; 2012 3 December 2012.
34. Schmitt J, Bierbaum T, Geraedts M, Gothe H, Harter M, Hoffmann F, et al. [Das Gesundheitsdatennutzungsgesetz - Potenzial für eine bessere Forschung und Gesundheitsversorgung]. *Gesundheitswesen*. 2023.
35. NAKO Gesundheitsstudie study group. NAKO Gesundheitsstudie 2023 [cited 2023. Available from: <https://nako.de/>].
36. TREATGermany study group. TREATgermany - Deutsches Klinisches Register: Therapie und medizinische Versorgung von Patienten mit moderater bis schwerer Neurodermitis 2023 [Available from: <https://treatgermany.org/>].
37. Neubert O, Richter M. Psychiatrische Institutsambulanzen: Erste Schritte zur Transparenz im Rahmen der Psych-Entgeltreform. Stuttgart; 2016.
38. de Cruppé W, Assheuer M, Geraedts M, Beine K. Association between continuity of care and treatment outcomes in psychiatric patients in Germany: a prospective cohort study. *Bmc Psychiatry*. 2023;23(1):52.
39. Assheuer M, Beine K, Mehl C, Kellner M, Agelink M, Sieberer M, et al. [Implementation of Continuity of Care in Everyday Care - A Comparison Between two Psychiatric Hospitals]. *Psychiatr Prax*. 2021;48(3):143-8.
40. Falkai PH. S3-Leitlinie Psychosoziale Therapien bei schweren psychischen Erkrankungen [S3-Guideline Psycho-Social Therapies for Severe Mental Illnesses]. Berlin Heidelberg: Springer; 2013 2013.
41. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *MedCare*. 2005;43(11):1130-9.
42. Elixhauser A, Steiner C, Harris DR, Coffey RN. Comorbidity measures for use with administrative data. *Medical care*. 1998;36(1):8-27.
43. Sievers C. ICD-11: Mehr als nur ein Update. Berlin: BARMER Institut für Gesundheitssystemforschung (bifg); 2021 August 2021.
44. Jungclaussen I, Hauten L, Lentzen G. Was ist neu in der ICD-11? 2022 [20.04.2023]. Available from: <https://psylife.de/magazin/aenderungen-icd11>.
45. Grossimlinghaus I, Falkai P, Gaebel W, Janssen B, Reich-Erkelenz D, Wobrock T, et al. [Developmental process of DGPPN quality indicators]. *Nervenarzt*. 2013;84(3):350-65.
46. Soltmann B, Neumann A, March S, Weinhold I, Hackl D, Kliemt R, et al. Multiperspective and Multimethod Evaluation of Flexible and Integrative Psychiatric Care Models in Germany: Study Protocol of a Prospective, Controlled Multicenter Observational Study (PsychCare). *Front Psychiatry*. 2021;12:659773.
47. Neumann A, Hense H, Baum F, Kliemt R, Seifert M, Harst L, et al. Evaluation of a flexible and integrative psychiatric care model in a department of child and adolescent psychiatry in Tübingen, Germany: study protocol (EVA_TIBAS). *BMC health services research*. 2021;21(1):1262.

48. Garcia TB, Kliemt R, Claus F, Neumann A, Soltmann B, Baum F, et al. Self-reported versus administrative data records: implications for assessing healthcare resource utilization of mental disorders. *Research Square*; 2023.
49. March S, Andrich S, Drepper J, Horenkamp-Sonntag D, Icks A, Ihle P, et al. [Good Practice Data Linkage]. *Gesundheitswesen*. 2019;81(8-09):636-50.
50. Pfennig A, Soltmann B, Neumann A, Baum F, von Peter S, Ignatyev Y, et al. Ergebnisbericht - Wirksamkeit sektorenübergreifender Versorgungsmodelle in der Psychiatrie - eine prospektive, kontrollierte multizentrische Beobachtungsstudie (PsychCare). 2023 12.05.2023.
51. Lehmann G, Eger K, Potent A, Conrad P, Garmann-Rolfes B, Garre P, et al. Psych-Reform aus einem Guss. *f&w*. 2023;4/2023.



