Hochschule für Angewandte Wissenschaften Hamburg

Fakultät Life Sciences

Das Gesundheitsdatennutzungsgesetz: Untersuchung des Zugangs- und Nutzungsprozesses am Beispiel von „Orphan Diseases“ bei Kindern – Ein Scoping Review

Bachelorarbeit

Department Gesundheitswissenschaften

Vorgelegt von

Lena Rittmeister [redacted]

[redacted]

Erstprüfer

Prof. Dr. York Zöllner (HAW Hamburg)

Zweitprüfer

Dipl.-Psych. Franziska Kath (DAK-Gesundheit)

Wintersemester 2024/25

Abgabedatum: 17.01.2025

Die Abschlussarbeit wurde betreut und erstellt in Zusammenarbeit mit der DAK-Gesundheit

Abstract

Einleitung: In Deutschland sind relevante Gesundheitsdaten auf verschiedene Silos verteilt, die Antragsprozesse intransparent und analog, der Zugriff komplex, unsicher und kaum reproduzierbar. Zudem ist die Repräsentativität der Daten nicht gegeben. Das 2024 in Kraft getretene Gesundheitsdatennutzungsgesetz soll die wissenschaftliche Nutzung der Gesundheitsdaten in Einklang mit der Datenschutz-Grundverordnung verbessern. Ein einheitlicher Zugang zu den Abrechnungsdaten der gesetzlichen Krankenversicherung soll durch das Forschungsdatenzentrum-Gesundheit am Bundesinstitut für Arzneimittel und Medizinprodukte gewährleistet werden. Diese Arbeit untersucht die Zugangs- und Nutzungsprozesse von Gesundheitsdaten im Rahmen des Gesundheitsdatennutzungsgesetzes sowie deren Potenziale und Limitationen für die Forschung zu seltenen Erkrankungen.

Methode: Diese Arbeit ist als Scoping Review konzipiert. Die Suchrecherche auf dem LinkedIn Profil des Forschungsdatenzentrum-Gesundheit und den Datenbanken Google Scholar und LIVIVO ergibt 854 Ergebnisse. Nach Ausschluss der Duplikate und nicht relevanten Artikel wurden 17 Artikel in diese Arbeit eingeschlossen.

Ergebnisse: Der Zugang zu den Gesundheitsdaten wird über das Forschungsdatenzentrum Gesundheit und die Datenzugangs- und Koordinierungsstelle geregelt. Sichere Analyseräume und Datenschutzmaßnahmen gewährleisten eine zweckgebundene Nutzung. Klinische Studien zu seltenen Erkrankungen sind oft zeitaufwendig und ressourcenintensiv. Große Datenmengen bieten jedoch ein erhebliches Potenzial, ermöglichen eine ressourcenschonende Nutzung, eine nahezu unverzerrte Abbildung des Versorgungsgeschehens und die Möglichkeit zur Verknüpfung verschiedener Datenquellen. Herausforderungen bestehen in begrenzten klinischen Informationen, eingeschränkten soziodemografische Daten und methodischen Einschränkungen wie „Verrauschungstechniken“ und Upcoding.

Diskussion/ Schlussfolgerungen: Trotz einiger Herausforderungen überwiegt der Nutzen für die Patientinnen und Patienten, der sich aus dem neuen Gesetz ergibt. Die Forschenden sind folglich dazu angehalten, im Vorfeld der Antragsstellung zu eruieren, ob die Daten für den jeweiligen Forschungszweck geeignet sind.

Schlüsselwörter: Gesundheitsdatennutzungsgesetz, „Orphan Diseases“, Sekundärdatennutzung

Inhaltsverzeichnis

Abstract I
Tabellenverzeichnis IV
Abbildungsverzeichnis IV
Abkürzungsverzeichnis IV
1 Relevanz der Gesundheitsdatennutzung 1
2 Zielsetzung 3
2.1 Eingrenzung des Themas und Fragestellung 4
2.2 Struktur der Bachelorarbeit 5
3 Theoretischer Hintergrund 5
3.1 Überblick über die aktuelle Gesundheitsdatennutzung in Deutschland 5
3.2 Einführung in das Gesundheitsdatennutzungsgesetz 7
3.3 Seltene Krankheiten bei Kindern 8
4 Methode 9
4.1 Scoping Review 9
4.2 Ein- und Ausschlusskriterien 10
4.3 Suchstrategie 12
5 Ergebnis 12
5.1 Auswahl der Literatur 12
5.2 Der Zugangs- und Nutzungsprozess gemäß des Gesundheitsdatennutzungsgesetz 13
5.3 Möglichkeiten der Datennutzung durch das Gesundheitsdatennutzungsgesetz 18
5.4 Fallbeispiel: Zugangs- und Nutzungsprozess für Forschung zu „Orphan Diseases“ bei Kindern 22
5.5 Limitationen und Herausforderungen der Datennutzung durch das Gesundheitsdatennutzungsgesetz 23
6 Diskussion 26
6.1 Zusammenfassung der Ergebnisse 26
6.2 Interpretation der Ergebnisse 30
6.3 Beantwortung der Fragestellungen 32
6.3.1 Wie erfolgt der Zugang zu Gesundheitsdaten und deren Nutzung im Rahmen des Gesundheitsdatennutzungsgesetzes? 32
6.3.2 Welche spezifischen Möglichkeiten und Limitationen ergeben sich daraus für die Forschung zu seltenen Erkrankungen bei Kindern? 33
6.4 Stärken und Einschränkungen des Scoping Reviews 34
7 Fazit 35

7.1 Auswirkungen aus gesundheitswissenschaftlicher Sicht 35
7.2 Handlungsempfehlungen 36
8 Literaturverzeichnis 39
9 Anhang 44
9.1 Anhang A: Suchstring 44
9.2 Anhang B: Datenextraktion in Form einer Tabelle 44
10 Eidesstattliche Erklärung 48

Tabellenverzeichnis

Tabelle 1 Ein- und Ausschlusskriterien bei Literatursuche- und Auswahl 11
Tabelle 2 Datenextraktion in Form einer Tabelle 44

Abbildungsverzeichnis

Abbildung 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (eigene Darstellung, 2024) 13
Abbildung 2 Zugangs- und Nutzungsprozess gemäß des GDNG (eigene Darstellung, 2024) 18

Abkürzungsverzeichnis

BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
DAK-G	DAK-Gesundheit
DK-Stelle	Datenzugangs- und Koordinierungsstelle
FDZ-Gesundheit	Forschungsdatenzentrum Gesundheit
GDNG	Gesundheitsdatennutzungsgesetz
GKV	Gesetzliche Krankenversicherung
GKV-SV	GKV-Spitzenverband
VPN	Virtual Private Network

1 Relevanz der Gesundheitsdatennutzung

Die Relevanz der Gesundheitsdatennutzung wurde durch die Coronapandemie deutlich hervorgehoben. Sie hat exemplarisch aufgezeigt, dass eine schnelle Verfügbarkeit von Daten aus der Gesundheitsversorgung erforderlich ist, um den Bedarf an intensivmedizinischer Versorgung vorherzusagen und zu planen (March, et al., 2023, S. 146). Diesbezüglich ist zu kritisieren, dass Daten zu den wichtigen Kenngrößen der Coronapandemie wie der Hospitalisierung, dem Impfstatus und den Risikofaktoren in Deutschland nur unzureichend vorlagen (Lampl, et al., 2024, S. 444). Während der Bewältigung der COVID-19 Pandemie haben wir insbesondere von ausländischen Studien aus Israel und Großbritannien profitiert (Schmitt, et al., 2023, S. 216).

In der deutschen Presse wurde diese Problematik wiederholt thematisiert. Es wurden über die unzureichende Datenlage für Analysen im Kontext der Corona-Pandemie berichtet. Einer dieser veröffentlichten Artikel führt aus, dass die Bewertung der Pandemie-Maßnahmen durch Expertinnen und Experten aufgrund der unzureichenden Datenlage in Deutschland nicht möglich war (Busch, 2022). Ein weiterer Artikel legt dar, dass die Coronapandemie die Defizite der Digitalisierung bis zum gegenwärtigen Zeitpunkt offengelegt hat. Unter anderem wurden die Infektionszahlen per Fax übermittelt und Impftermine mussten anfänglich telefonisch über eine Hotline vereinbart werden (Digital Chiefs, 2022). Die Covid-19-Pandemie wird daher als "Datenerhebungskatastrophe" bezeichnet. Die Daten über den Verlauf dieser Pandemie weisen eine solch geringe Aussagekraft auf, dass sich Politik, Wissenschaft und Verwaltung in einer peinlichen Situation befinden (Wiarda, 2021). Die geringe Datenverfügbarkeit und die damit eingeschränkten Handlungsmöglichkeiten durch die Politik, werden auch wissenschaftlich diskutiert (Weber & Kuhn, 2023, S. 103; Deutsche Forschungsgemeinschaft, 2022).

Parallel dazu führt die fortschreitende Digitalisierung zu einer stetigen Zunahme des Volumens an Daten im Gesundheitswesen (Prasser, et al., 2024, S. 171). Gesundheitsdaten sind Daten, die im Rahmen der routinemäßigen Abläufe im Gesundheitswesen generiert werden. Diesbezüglich sind insbesondere Diagnose-, Behandlungs- und Abrechnungsdaten zu nennen (Heisig, et al., 2021, S. 363). Diese Gesundheitsdaten werden als Routinedaten bezeichnet (Heisig, et al., 2021, S. 367). Die korrekte und angemessene Nutzung zu Forschungs-, Beobachtungs- und Berichtszwecken kann einen erheblichen gesellschaftlichen Nutzen stiften. Aus diesem Grund ist es empfehlenswert, eine grundsätzlich positive Nutzungserlaubnis zu erteilen (Heisig, et al., 2021, S. 363). Die Routinedaten der gesetzlichen

Krankenversicherung (GKV) beinhalten eine Vielzahl von Informationen bezüglich des gesundheitlichen Zustands der Versicherten. Dazu zählen beispielsweise Daten über die stationäre und ambulante Versorgung, die Stammdaten der Versicherten sowie Informationen über die Pflege. Die GKV-Routinedaten stellen eine geeignete Grundlage für die Beschreibung der Morbiditätslast dar, da sie ambulante Diagnosen und stationäre Hauptentlassungsdiagnosen sowie stationäre Aufnahme- und Nebendiagnosen umfassen. Aufgrund des vorhandenen Populationsbezuges lassen sich auf dieser Basis Inzidenz- und Prävalenzermittlungen vornehmen. Die Eignung der GKV-Routinedaten für Forschungszwecke ist demnach gegeben (Ohlmeier, et al., 2014, S. 464f.).

Angesichts gesellschaftlicher Wandlungsprozesse steht die gesundheitliche Versorgung vor neuen Herausforderungen. Um sozialer und gesundheitlicher Ungleichheit entgegenzuwirken und die physische sowie psychische Gesundheit aller Bevölkerungsgruppen zu erhalten, Krankheiten zu behandeln oder zu verhindern, ist ein umfassendes Verständnis der vielfältigen Bedürfnisse der einzelnen Akteure und Patientinnen und Patienten notwendig. Subgruppenspezifische Gesundheitsdaten sind hierfür unerlässlich, da sie Prozesse und Strukturen im Gesundheitskontext nachvollziehbar machen und der Erforschung von sozialer Ungleichheit eine detaillierte Perspektive hinzufügen. Sie bilden die Grundlage für die zielgruppenorientierte Konzeption, Umsetzung und Evaluation von Versorgungstherapien. Im Rahmen der aktuellen Digitalisierungsprozesse im Gesundheitswesen eröffnen sich neue Möglichkeiten hinsichtlich der Bereitstellung, Erhebung und Auswertung von Daten zur Verbesserung der Patientinnen- und Patientenversorgung. Die Erfassung von Gesundheitsdaten erfolgt zunehmend in digitaler Form, wobei ein wachsender Anteil dieser Daten als Routinedaten verfügbar gemacht wird (Veit, et al., 2019, S. 20).

Ein weiteres Problem, das in diesem Zusammenhang berücksichtigt werden muss, ist das Braindrain, also die Emigration von Arbeitskräften (Klein, 2018). Um die Abwanderung von Forschenden zu verhindern und Deutschland für internationale Expertinnen und Experten im Gesundheitswesen attraktiver zu machen, ist die Schaffung eines verbesserten Datenzugangs von essenzieller Bedeutung. Insbesondere der Aspekt der Verfügbarkeit von Daten ist von entscheidender Bedeutung. Ein umfassender und einfacher Zugriff auf Daten kann dabei als Anreiz für Forschende dienen, sich in Deutschland niederzulassen (Fischer, et al., 2023, S. 4). Zudem prognostizieren Expertinnen und Experten, dass sich die Medizin durch die intelligente Vernetzung und Analyse von Daten in den Bereichen Forschung,

Diagnostik und Therapie erheblich verbessern wird (Kostera & Thranberend, 2018, S. 2).

Die FAIR-Leitprinzipien stellen eine umfassende Grundlage für moderne Datenveröffentlichungsumgebungen bereit, welche sowohl manuelle als auch automatisierte Prozesse unterstützen. Die Prinzipien der FAIR-Leitlinien umfassen die Grundsätze der Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit. Die Auffindbarkeit von Daten wird durch die Speicherung in durchsuchbaren Repositorys mit hinlänglichen Metadaten gewährleistet. Der Begriff der Zugänglichkeit zielt darauf ab, dass Daten unter klar definierten Bedingungen abrufbar sein müssen. Dies beinhaltet die Festlegung von Berechtigungen und Authentifizierungen. Unter dem Prinzip der Interoperabilität ist die Speicherung von Daten in einem Format zu verstehen, welches die Integration mit anderen Datensätzen ermöglicht. Dabei ist die Verwendung gemeinsamer Standards und Ontologien erforderlich. Schließlich impliziert die Wiederverwendbarkeit, dass Daten in einer Weise aufbereitet werden sollten, die ihre Nutzbarkeit für künftige Forschungsvorhaben gewährleistet. Dies erfordert eindeutige Lizenzbedingungen sowie eine umfassende Dokumentation. Die Implementierung dieser Prinzipien zielt darauf ab, die Effizienz und Effektivität der Datennutzung zu optimieren und die Wiederverwendbarkeit von Daten innerhalb der wissenschaftlichen Gemeinschaft zu fördern (Wilkinson, et al., 2016, S. 1f.).

2 Zielsetzung

Die Zielsetzung dieses Scoping Reviews besteht in einer umfassenden und detaillierten Analyse des Zugangs- und Nutzungsprozesses von Gesundheitsdaten im Rahmen des Gesundheitsdatennutzungsgesetz (GDNG). Im Fokus der Untersuchung steht dabei die Erforschung seltener Erkrankungen im Kindesalter. Im Rahmen dessen werden sowohl die Möglichkeiten als auch die Limitationen dieses neuen gesetzlichen Rahmens aufgezeigt und dessen Auswirkungen auf das Forschungsfeld untersucht.

Ein wesentlicher Bestandteil der Untersuchung stellt die detaillierte Beschreibung und Analyse sämtlicher Phasen des Zugangs- und Nutzungsprozess gemäß des GDNG dar. Dies umfassen die Identifikation und Erläuterung der erforderlichen Unterlagen, die Beschreibung der Analyseräume sowie die Skriptentwicklung und dem Prozess einer Publikation. Ein besonderes Forschungsinteresse gilt der Rolle und Funktion des Forschungsdatenzentrums Gesundheit (FDZ-Gesundheit) im Kontext des Antragsprozesses.

Zur Veranschaulichung der praktischen Implikationen des Zugangs- und Nutzungsprozesses wird ein Beispiel für einen Antrag eines Forschungsprojekts zu „Orphan Diseases“ bei Kindern entwickelt und Schritt für Schritt durchlaufen. Im Rahmen dessen erfolgt eine Identifikation potenzieller Herausforderungen im praktischen Antragsverfahren.

Die vorliegende Arbeit hat zum Ziel, die durch das GDNG eröffneten neuen Forschungsmöglichkeiten für seltene Erkrankungen bei Kindern zu identifizieren. Dabei wird das Potenzial für Datenverknüpfungen und deren Bedeutung für die Erforschung seltener Erkrankungen untersucht sowie Möglichkeiten zur Verbesserung der Datenqualität und -quantität für dieses spezielle Forschungsfeld analysiert. Ein weiterer Untersuchungsgegenstand liegt auf der kritischen Untersuchung der Limitationen und Herausforderungen. Dies beinhaltet die Untersuchung technischer, inhaltlicher und methodischer Herausforderungen bei der Verarbeitung und Analyse von Daten zu seltenen Erkrankungen bei Kindern.

Basierend auf diesen Analysen werden konkrete Vorschläge zur Optimierung des Antragsprozesses entwickelt. Es werden Möglichkeiten zur Überwindung der festgestellten Limitationen identifiziert und Ansätze zur Verbesserung der Datenqualität und -verfügbarkeit für dieses spezielle Forschungsfeld erarbeitet.

Durch die Erreichung dieser Ziele soll die Arbeit einen Beitrag zum Verständnis der praktischen Umsetzung des GDNG leisten und gleichzeitig die spezifischen Herausforderungen und Chancen für die Forschung zu seltenen Erkrankungen bei Kindern untersuchen. Die vorliegende Arbeit zielt darauf ab, nicht nur eine Analyse des Status quo durchzuführen, sondern auch zukunftsweisende Perspektiven für die Optimierung der Forschungsmöglichkeiten in diesem bedeutenden Bereich zu präsentieren.

2.1 Eingrenzung des Themas und Fragestellung

Das Thema des GDNG ist ein Arbeitsfeld, welches eine Vielzahl von Aspekten umfasst, darunter den Datenschutz, die technische Infrastruktur sowie die Nutzung von Gesundheitsdaten für die Forschung. Um das Thema dieser Bachelorarbeit angemessen zu behandeln, ist eine Eingrenzung erforderlich. Die vorliegende Bachelorarbeit widmet sich in erster Linie dem Zugangs- und Nutzungsprozess gemäß des GDNG sowie den Einschränkungen bei der Nutzung von Daten. Die Fokussierung erfolgt, da die DAK-Gesundheit (DAK-G) als Praktikumsgeberin eine Antragsstellung vorbereitet und die Möglichkeiten und Grenzen der Datennutzung

evaluieren möchte. Es gibt bislang kein veröffentlichtes Review, welches dieses Thema bearbeitet hat.

Die vorliegende Bachelorarbeit hat zum Ziel, den Zugangs- und Nutzungsprozess gemäß des GDNG zu untersuchen und die Möglichkeiten und Limitationen der Datenarbeit für eine Krankenkasse zu untersuchen. Die auf den Zielen basierende Forschungsfrage lautet wie folgt: Wie erfolgt der Zugang zu Gesundheitsdaten und deren Nutzung im Rahmen des GDNG? Welche spezifischen Möglichkeiten und Limitationen ergeben sich daraus für die Forschung zu seltenen Erkrankungen bei Kindern?

2.2 Struktur der Bachelorarbeit

Um die Fragestellung zu beantworten, wird zunächst der theoretische Hintergrund beschrieben. Im nachfolgenden Abschnitt wird das methodische Vorgehen der Bachelorarbeit dargelegt. Dies beinhaltet das Konzept eines Scoping Reviews, die Auswahlkriterien für die Literatur sowie den Prozess zur Erstellung des Suchstrings. Im nachfolgenden Kapitel erfolgt die Darstellung der Ergebnisse. Darin wird zunächst die gefundene Literatur vorgestellt, im Anschluss wird der Antragsprozess gemäß des GDNG beschrieben. Nachfolgend erfolgt ein exemplarischer Zugangs- und Nutzungsprozess am Beispiel seltener Erkrankungen bei Kindern, aus der sich die Möglichkeiten sowie die Limitationen ableiten lassen, die sich aus dem GDNG ergeben. Im Kapitel der Diskussion erfolgt eine Zusammenfassung der Ergebnisse sowie deren Interpretation und die beiden Forschungsfragen werden beantwortet. Des Weiteren werden in diesem Artikel die Einschränkungen dieser Arbeit aufgezeigt. In dem letzten Kapitel, dem Fazit erfolgt die Formulierung von Empfehlungen für die Praxis, die Forschung und die Politik. Des Weiteren erfolgt eine Darstellung der Auswirkungen des GDNG aus der Perspektive der Gesundheitswissenschaften.

3 Theoretischer Hintergrund

Im vorliegenden Kapitel werden wesentliche Informationen über die Nutzung von Gesundheitsdaten in Deutschland, die Einführung des Gesundheitsdatennutzungsgesetzes sowie seltene Erkrankungen bei Kindern dargelegt.

3.1 Überblick über die aktuelle Gesundheitsdatennutzung in Deutschland

Deutschland liegt in einem internationalen Vergleich der Digital-Health-Entwicklung trotz einiger gesetzlicher Initiativen auf einem der letzten Plätze. In einem Vergleich

mit 16 Nationen belegt Deutschland den vorletzten Platz (Kostera & Thranberend, 2018, S. 2). Bereits im Jahr 2003 wurden mit dem GKV-Modernisierungsgesetz erste Schritte unternommen, um die bei den Krankenkassen vorhandenen Gesundheitsdaten zu bündeln und für Forschungszwecke zugänglich zu machen. Diese Bestrebungen wurden in den Paragrafen 303a ff. SGB V verankert, wobei zugleich die Einrichtung einer sogenannten Datentransparenzstelle gesetzlich festgeschrieben wurde. Die Datentransparenzstelle hat die Aufgaben einer Vertrauens- und Datenaufbereitungsstelle zu erfüllen. Mit dem GKV-Versorgungsstrukturgesetz von 2011 wurden diese Vorschriften weiterentwickelt (Schulte, 2024, S. 134). In den vergangenen Jahren wurden seitens des Bundesministeriums für Gesundheit mehrere Gesetze verabschiedet, welche die Digitalisierung im Gesundheitswesen vorantreiben sollen. Im Jahr 2015 trat das E-Health-Gesetz in Kraft, welches die Basis für eine sichere digitale Kommunikation und Anwendungen im Gesundheitswesen schuf. Es legte den Grundstein für die Etablierung der Telematikinfrastruktur sowie die Implementierung der elektronischen Gesundheitskarte. Infolgedessen wurde im Jahr 2019 das Digitale-Versorgung-Gesetz erlassen, welches die Einführung digitaler Gesundheitsanwendungen fördert und Ärztinnen und Ärzten die Verschreibung von Apps auf Rezept ermöglicht. Des Weiteren wurde die Telemedizin sowie die Nutzung von Videosprechstunden gestärkt. Das Patientendaten-Schutz-Gesetz definiert die Rahmenbedingungen für den Datenschutz im Kontext der elektronischen Patientenakte sowie des E-Rezeptes, die Verarbeitung von Patientinnen- und Patientendaten wurde dadurch sichergestellt (Bundesministerium für Gesundheit, o.J.).Das im Jahr 2023 erlassene Digitalgesetz zielt darauf ab, den Prozess der Digitalisierung zu beschleunigen. Die elektronische Patientenakte wird als Standard für alle gesetzlich Versicherten eingeführt und das E-Rezept als verbindlich erklärt. Das übergeordnete Ziel besteht in der Verbesserung des Versorgungsalltags sowie der Unterstützung der Forschung (Digital-Gesetz, 2024, S. 1).

Trotz der bestehenden gesetzlichen Grundlagen sind die Daten der in der gesetzlichen Krankenversicherung Versicherten aktuell weiterhin auf rund 100 unterschiedliche Krankenkassen aufgeteilt. Im Rahmen einer vom Verein für Socialpolitik durchgeführten Online-Umfrage wurde die Zufriedenheit von 500 seiner Mitglieder mit dem Zugang zu Forschungsdaten untersucht. Über zwei Drittel der Befragten äußerten eine „eher unzufriedene“ oder „gar unzufriedene“ Einstellung. Das letzte Drittel der Befragten äußerte sich mit "eher zufrieden". Es konnte keine einzige positive Rückmeldung registriert werden. Als wesentliche Gründe für die

Unzufriedenheit wurden die unzulängliche Bereitstellung bereits vorhandener Daten, die fehlende Verknüpfbarkeit sowie die restriktiven Datenschutzvorgaben genannt (Fischer, et al., 2023, S. 8).

Versorgungsforschende die Daten aller gesetzlich Versicherten analysieren möchten, müssen daher bei jeder einzelnen Krankenkasse einen Zugang zu den Daten beantragen und die erhaltenen Ergebnisse anschließend selbst zusammenführen (Schulte, 2024, S. 135). Derzeit ist es den Krankenkassen lediglich möglich, die Daten ihrer eigenen Versicherten auszuwerten. Ein Beispiel hierfür ist der Kinder- und Jugendreport der DAK-Gesundheit (Witte, et al., 2023, S. 1). Dieses dezentrale System erschwert umfassende Forschungsprojekte, insbesondere zu seltenen Erkrankungen, da bestimmte Krankheiten möglicherweise bei einer einzelnen Krankenkasse nicht ausreichend repräsentiert sind. Diese Konstellation demonstriert, dass trotz der gesetzlichen Initiativen zur Optimierung der Datentransparenz und -nutzung in der Praxis noch Herausforderungen bei der zentralen Zusammenführung und Nutzung von Gesundheitsdaten für Forschungszwecke bestehen. Zur Erfassung verlässlicher Daten zur Prävalenz seltener Erkrankungen wurde mit Wirkung zum 1. April 2023 eine gesetzliche Verpflichtung für Krankenhäuser eingeführt, seltene Erkrankungen im stationären Bereich zu kodieren. Diese Kodierung ist von entscheidender Bedeutung für die Sichtbarkeit seltener Erkrankungen in den Informationssystemen von Gesundheits- und Forschungseinrichtungen (Bundesministerium für Gesundheit, 2024a).

Die bereits erörterten FAIR-Prinzipien werden hinsichtlich der Gesundheitsdaten in Deutschland nicht erfüllt. Die Daten sind auf verschiedene Silos verteilt, die Antragsprozesse sind intransparent und analog, der Zugriff ist komplex, unsicher und kaum reproduzierbar. Zudem ist die Repräsentativität der Daten nicht gegeben. Ein einheitlicher Zugang von zu den Abrechnungsdaten aus der GKV soll zukünftig durch das FDZ-Gesundheit am Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) gewährleistet werden. Des Weiteren befindet sich gegenwärtig ein europäischer Datenraum für Gesundheitsdaten im Aufbau, welcher den ersten gemeinsamen Datenraum der Europäischen Union in einem spezifischen Fachgebiet darstellt (Gassner, 2022, S. 739).

3.2 Einführung in das Gesundheitsdatennutzungsgesetz

Im Jahr 2021 wurde das GDNG im Koalitionsvertrag festgelegt. Das Gesetz soll zu einer besseren wissenschaftlichen Nutzung der Gesundheitsdaten in Einklang mit der Datenschutz-Grundverordnung führen. Dafür wird die Schaffung einer dezentralen

Forschungsdateninfrastruktur thematisiert (SPD, Bündnis 90/ Die Grünen, & FDP, 2021, S. 65). Das GDNG ist im ersten Artikel des Gesetzes zur verbesserten Nutzung von Gesundheitsdaten festgelegt und wurde am 26.03.2024 verabschiedet. Das GDNG legt besonderen Wert auf die Nutzung von Gesundheitsdaten für Forschungszwecke, die dem Gemeinwohl dienen, und auf die datenbasierte Verbesserung des Gesundheitswesens. Die Intention dieses Gesetzes ist eine Optimierung der Nutzung von Gesundheitsdaten sowie eine Regulierung derselben. Ziel ist es, den Zugang zu und die Nutzung von Gesundheitsdaten für Forschung und Gesundheitsversorgung zu erleichtern, während gleichzeitig der Datenschutz gewährleistet wird (Gesundheitsdatennutzungsgesetz, 2024, S. 1).

3.3 Seltene Krankheiten bei Kindern

In dieser Arbeit wird die europäische Definition einer seltenen Krankheit verwendet, wie sie in der Verordnung der Europäischen Union über Arzneimittel für seltene Leiden festgelegt ist. Eine seltene Krankheit betrifft nicht mehr als fünf Personen pro 10.000 Einwohner in der europäischen Bevölkerung (Europäisches Parlament und Rat, 1999, S. 1). Diese Krankheiten, die auch als sogenannte „Orphan Diseases“ bezeichnet werden, sind häufig genetisch bedingt und mit einem schweren chronischen Verlauf verbunden (Fehr & Prütz, 2023, S. 3). Der Begriff "Orphan Diseases" leitet sich von den "Orphan Drugs" ab, wobei es sich um Medikamente handelt, die zur Behandlung von seltenen Erkrankungen hergestellt werden. Die Besonderheiten von "Orphan Diseases" und "Orphan Drugs" machen sie zu einem Thema mit hoher Priorität für politische Entscheidungsträger (Denis, et al., 2009 zitiert nach Denis, et al., 2010, S. 344).

Die überwiegende Anzahl der über 8.000 bekannten seltenen Erkrankungen manifestiert sich im Kindes- und Jugendalter. Schätzungen zufolge sind in Deutschland etwa 4–6 % der Bevölkerung von einer seltenen Erkrankung betroffen, wobei der Anteil bei Kindern und Jugendlichen wahrscheinlich höher ist. Dies bedeutet, dass von den 13,75 Millionen Personen unter 18 Jahren in Deutschland etwa 700.000 betroffen sind. Diese Erkrankungen spielen sowohl in der stationären als auch in der ambulanten pädiatrischen Versorgung eine bedeutsame Rolle (Hoffmann & Hebestreit, 2022, S. 10).

Die Durchführung klinischer Studien zu seltenen Erkrankungen ist aufgrund der geringen Fallzahlen oft mit einem langen Zeitraum und einem hohen Aufwand verbunden. Für die forschenden Unternehmen sind sie daher in der Regel

unwirtschaftlich und werden häufig vernachlässigt (Bundesministerium für Bildung und Forschung, o.J.).

Patientinnen und Patienten mit seltenen Krankheiten warten oft jahrelang auf eine Diagnose und eine geeignete Behandlung. Die medizinische Expertin oder der medizinische Experte, der eine solche Krankheit diagnostizieren kann, praktiziert möglicherweise in einer anderen Region oder sogar in einem anderen Land. Darüber hinaus sind die wissenschaftlichen Erkenntnisse über bestimmte seltene Krankheiten oft unzureichend und verstreut (European Commision, 2014, S. 3).

Im Rahmen einer in Deutschland durchgeführten Umfrage gaben 69 % der befragten Patientinnen und Patienten mit seltenen Erkrankungen an, aufgrund einer Fehldiagnose eine falsche Behandlung erhalten zu haben. Dabei handelte es sich in 26 % der Fälle um eine medizinische, in 14 % der Fälle um eine chirurgische und in 5 % der Fälle um eine psychologische Behandlung. Zudem hat ein Fünftel der Befragten während der Diagnose fünf oder mehr Ärztinnen oder Ärzte aufgesucht (Kole & Faurisson, 2009, S. 242).

Zu den häufigsten seltenen Erkrankungen in Deutschland zählt die Lungenerkrankung Mukoviszidose. Die Mukoviszidose stellt eine angeborene Multiorganerkrankung dar. Die Anzahl der bekannten Fälle in Deutschland beläuft sich auf etwa 8.000 (DGP & GPP, 2022, S. 11). Ein weiteres Beispiel bietet die seltene und genetisch bedingte Nierenerkrankung, das Alport-Syndrom. Die meisten Betroffenen werden bereits im Jugend- oder jungen Erwachsenenalter dialysepflichtig (Bundesministerium für Bildung und Forschung, o.J.).

4 Methode

Im nachfolgenden Abschnitt folgt eine detaillierte Erläuterung des methodischen Vorgehens der Bachelorarbeit. Zudem wird die Wahl des Scoping Reviews als Design begründet. Im Folgenden werden die Ein- und Ausschlusskriterien dargelegt und der Rechercheprozess veranschaulicht. Abschließend wird der Suchstring definiert.

4.1 Scoping Review

Der Scoping Review wird durchgeführt, um eine detaillierte Beschreibung der Ergebnisse und dem Umfang der Forschung zu erreichen. Es bietet damit einen Mechanismus zur Zusammenfassung und Verbreitung des aktuellen Wissenstandes. Mit einem Scoping Review ist es möglich, einen Überblick über die vorhandene Evidenz zu geben und das unabhängig von ihrer Qualität und Methode (Arksey & O'Malley, 2005, S. 20).

Folgende Schritte sind Teil der Recherche:

1. Definition der Fragestellung

2. Identifizierung relevanter Studien

3. Auswahl der Studien

4. Darstellung der gesammelten Daten

5. Zusammenfassen der Ergebnisse (Arksey & O'Malley, 2005, S. 22)

4.2 Ein- und Ausschlusskriterien

Die Auswahl der Literatur erfolgte anhand der folgenden Kriterien: In die Literaturauswahl wurden Publikationen einbezogen, die nach dem 24.11.2021 erschienen sind. An diesem Tag wurde das GDNG im Koalitionsvertrag festgehalten. Des Weiteren ist erforderlich, dass die Texte vor dem 5. Dezember 2024 publiziert wurden. Die Texte müssen in deutscher Sprache verfasst sein. Die Berücksichtigung von Texten in der vorliegenden Arbeit setzt deren Zuordnung zu einer der folgenden Dokumentenkategorien voraus: Gesetzestext oder Verordnung, wissenschaftlicher Artikel oder Studie zur Nutzung von Gesundheitsdaten, Stellungnahme oder Positionspapier einer relevanten Organisation, Meldung des FDZ-Gesundheit oder Leitlinie zur Antragstellung für Gesundheitsdaten sowie Präsentation des FDZ-Gesundheit. Thematisch müssen die Texte die aktuelle Sekundärdatennutzung als Gegenstand haben, sich mit der Beschreibung des Antragsprozess, der potenziellen Datennutzung der Daten gemäß des GDNG oder den Limitationen der Datennutzung befassen. Da sich die Arbeit mit der Sekundärdatennutzung in Deutschland befasst, müssen die eingeschlossenen Texte in Deutschland veröffentlicht worden sein oder Bezug zu dem deutschen Gesundheitssystem nehmen. Die Texte müssen aus den folgenden Quellen stammen: Wissenschaftliche Datenbanken, Webseiten von Bundesbehörden und dem FDZ-Gesundheit oder Gesetzesbücher.

Eine Übersicht der Ein- und Ausschlusskriterien bietet die Tabelle 1.

Tabelle 1 Ein- und Ausschlusskriterien bei Literatursuche- und Auswahl

	Einschlusskriterium	Ausschlusskriterium
Zeitraum	Veröffentlicht zwischen dem 24.11.2021 und dem 04.12.2024	Veröffentlicht vor dem 24.11.2021 oder nach dem 04.12.2024
Sprache	Deutsch	Alle Sprachen außer Deutsch
Dokumententyp	Gesetzestexte und Verordnungen, wissenschaftliche Artikel und Studien zur Nutzung von Gesundheitsdaten, Stellungnahmen und Positionspapiere relevanter Organisationen, Meldungen des FDZ-Gesundheit, Leitlinien zur Antragstellung für Gesundheitsdaten, Präsentationen des Forschungsdatenzentrum	Blogs, Foren, Laienpresse, Webseiten von Kanzleien, Präsentationen von anderen Institutionen als dem FDZ- Gesundheit
Thematischer Fokus	Beschreibung des Antragsprozess, Nutzung der Daten, Limitationen der Daten	Themen die sich nicht primär mit dem GDNG oder der Sekundärdatennutzung beschäftigen
Geographischer Fokus	Deutschland	Anderer Länderbezug als Deutschland
Quellen	Wissenschaftliche Datenbanken, Webseiten von Bundesbehörden und FDZ-Gesundheit und Gesetzestexte	Laienpresse, Webseiten von Kanzleien, Schulungsunterlagen

4.3 Suchstrategie

Nach Vorlage von Elm und Kollegen wird eine dreistufige Suchstrategie angewendet. Zuerst findet eine eingeschränkte Suche in der Datenbank Google Scholar statt. In den identifizierten Artikeln werden die im Titel und Abstract enthaltenen Schlüsselwörter sowie die in den Datenbanken verwendeten Indexbegriffe analysiert. Folgende Schlagwörter wurden gefunden: Gesundheitsdatennutzungsgesetz, Forschungsdatenzentrum Gesundheit, Versorgungsnahe Daten, Sekundärdatennutzung, Gesundheitsdatennutzung, Datenzugangs- und Koordinierungsstelle, Forschungsdaten im Gesundheitswesen, Datenzugang und Routinedaten. Als zweiter Schritt wurde dann eine Suche mit allen identifizierten Stichwörtern und Indexbegriffen in Datenbanken und Suchmaschinen durchgeführt. Das waren die Suchmaschinen LIVIVO und Google Scholar. Außerdem wurden die Posts der LinkedIn Seite des FDZ-Gesundheit nach weiteren Artikeln und Informationen durchsucht. Aufgrund der Tatsache, dass das FDZ-Gesundheit in erster Linie mittels LinkedIn kommuniziert, wurde letztere Datenquelle für die Analyse herangezogen. In dem dritten Schritt wurden die Referenzlisten der gefundenen Artikel nach zusätzlichen Veröffentlichungen durchsucht (Von Elm, et al., 2019, S. 5). Alle schlussendlich verwendeten Suchbegriffe sind in dem Anhang A aufgeführt.

5 Ergebnis

Im folgendem werden die Ergebnisse detailliert dargestellt. Die einbezogenen Artikel sind als Tabelle im Anhang B ersichtlich. Aus den insgesamt 852 Artikeln, die in den Datenbanken gefunden wurden, werden 17 in den Scoping Review einbezogen.

5.1 Auswahl der Literatur

Die Abbildung 1 dient als Übersicht über das Auswahlverfahren der relevanten Artikel. Die Suche in Google Scholar, LIVIVO und den LinkedIn Posts fand am 04.12.2024 statt und identifizierte 853 Artikel. Die Duplikate wurden händisch entfernt. Bei den 812 recherchierten Artikeln erfolgte die Überprüfung der Relevanz durch Sichtung des Titels und Abstracts anhand der festgelegten Ein- und Ausschlusskriterien (siehe Tabelle 1.). So konnten 719 Artikel ausgeschlossen werden. Um die Eignung der verbliebenden 89 Artikel zu prüfen, mussten sie als Volltext gelesen werden. Entlang der Ein- und Ausschlusskriterien konnte entschieden werden, ob ein Artikel in den Scoping Review einbezogen wurde oder nicht. Ein Artikel wurde ausgeschlossen, da er vor dem 24.11.201 publiziert wurde, 59 Artikel wurden nicht in den Scoping Review einbezogen, da sich der thematische Fokus nicht mit dem Zugangs- und Nutzungsprozess des GDNG befasst. Weitere 12 Artikel wurden ausgeschlossen,

weil die Volltextartikel unauffindbar waren. Insgesamt erfüllten 17 Artikel die festgelegten Voraussetzungen und wurden in den Scoping Review einbezogen.

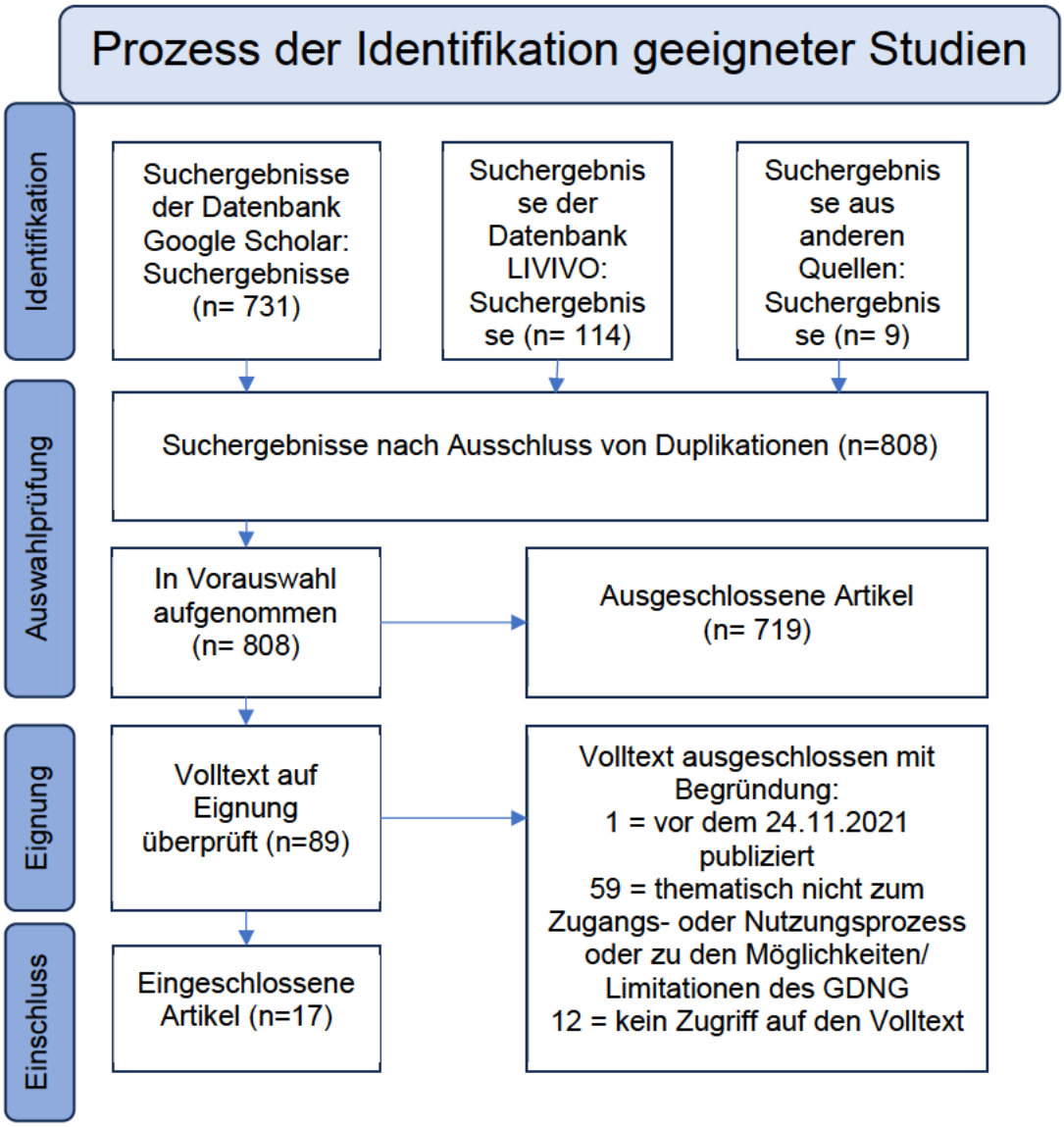

Abbildung 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (eigene Darstellung, 2024)

5.2 Der Zugangs- und Nutzungsprozess gemäß des Gesundheitsdatennutzungsgesetz

Der Zugangs- und Nutzungsprozess von Gesundheitsdaten gemäß dem GDNG erfolgt über das FDZ-Gesundheit und die zentrale Datenzugangs- und Koordinierungsstelle für Gesundheitsdaten (DK-Stelle) beim BfArM. Diese beiden Institutionen wurden eingerichtet, um Gesundheitsdaten sicher, transparent und zentralisiert für legitime Forschungsvorhaben zugänglich zu machen (Henke & Schneider, 2024, S. 221).

Zunächst wird beschrieben, wie der Datensatz gemäß des GDNG technisch entsteht. In einem komplexen Verfahren gemäß §§ 303 a bis f SGB V sollen die Krankenkassen quartalsweise ihre Leistungsdaten an den Spitzenverband der GKV (GKV-SV) übermitteln. Diese Daten enthalten ein „Versichertenpseudonym“, das eine kassenübergreifende eindeutige Identifizierung ermöglicht, ein sogenanntes Lieferpseudonym. Der GKV-SV prüft diese Daten und überträgt sie ohne das Lieferpseudonym, stattdessen mit einer Arbeitsnummer und pseudonymisierten Informationen an das FDZ-Gesundheit. Gleichzeitig werden die Daten mit dem Lieferpseudonym und der Arbeitsnummer an eine Vertrauensstelle weitergegeben. Laut dem Bundesministerium für Gesundheit wird dies gemäß der Datentransparenzverordnung an die Vertrauensstelle des Robert Koch-Instituts erfolgen. Die Vertrauensstelle übernimmt die Aufgabe, die Lieferpseudonyme in „periodenübergreifende Pseudonyme“ umzuwandeln. In Kooperation mit dem Bundesamt für Sicherheit in der Informationstechnik wird ein „schlüsselabhängiges Verfahren zur Pseudonymisierung" entwickelt. Dieses Verfahren gewährleistet, dass jedem Versicherten stets dasselbe Pseudonym zugewiesen wird, ohne dass Rückschlüsse auf das ursprüngliche Lieferpseudonym oder die Identität des Versicherten möglich sind. Die neu erstellten Pseudonyme werden an das FDZ-Gesundheit übermittelt. Nach der Datenübermittlung werden die Lieferpseudonyme und Arbeitsnummern gelöscht. Das FDZ-Gesundheit bewertet das „spezifische Reidentifikationsrisiko“ und trifft geeignete Maßnahmen, um dieses Risiko unter Berücksichtigung des wissenschaftlichen Nutzens zu minimieren. Die Daten sollen nach 100 Jahren gelöscht werden (Dannenberg, 2024, S. 16f.).

Der bisher geltende Akteursbezug (d.h. die Beschränkung des Nutzerkreises auf bestimmte Institutionen) wird durch einen Zweckbezug ersetzt. Das heißt, der Nutzerkreis ist nicht mehr auf bestimmte Institutionen beschränkt. Die an das FDZ-Gesundheit übermittelten Daten dürfen von den Datennutzenden verarbeitet werden, wenn sie zum Beispiel für die Verbesserung der Versorgungsqualität, Prävention und Pflege, gesundheitswissenschaftliche Forschung, die Analyse sektorenübergreifender Versorgungsformen, die Nutzenbewertung von Arzneimitteln und Medizinprodukten oder für das Testen und Trainieren von KI-Anwendungen im Gesundheitswesen verwendet werden. (Henke & Schneider, 2024, S. 221).

Zunächst müssen interessierte Institutionen, wie Krankenkassen, Universitäten oder Forschungseinrichtungen, ihre Institution im Portal des FDZ-Gesundheit registrieren. Nach erfolgreicher Authentifizierung und Freischaltung kann die Institution einen Antrag auf Datennutzung stellen. Dabei muss im Antrag detailliert beschrieben

werden, zu welchem Zweck die Daten genutzt werden sollen und welche Auswertungsmethoden vorgesehen sind. Die beantragten Daten dürfen nur für spezifische Nutzungszwecke verwendet werden, die gesetzlich zugelassen sind, beispielsweise zur Verbesserung der Versorgungsqualität oder die Krankenhausplanung. Zudem ist es erforderlich, die beteiligten Personen und die Struktur sowie den Umfang der benötigten Daten anzugeben. Dies ermöglicht eine sorgfältige und datenschutzkonforme Prüfung des Bedarfs (Ludwig, et al., 2024, S. 134).

Nach Eingang eines Antrags in das FDZ-Gesundheit erfolgt zunächst eine formale Prüfung desselben, wobei insbesondere auf die Vollständigkeit der Unterlagen geachtet wird. Im Anschluss wird eine inhaltliche Evaluierung durchgeführt, wobei unter anderem die angegebenen Nutzungszwecke einer kritischen Prüfung unterzogen werden. In der Vergangenheit wurde seitens der Forschenden oftmals eine zu lange Dauer des Prüfzeitraums sowie eine unzureichende Transparenz desselben kritisiert (Ihle, et al., 2022, S. 41). Dem Wunsch nach Planbarkeit wird nun dadurch entsprochen, dass innerhalb eines Zeitraums von drei Monaten (maximal vier bei aufwendiger Prüfung) über den Antrag entschieden werden muss (Datentransparenzverordnung, 2020, S. 7). Außerdem sollen die Antragstellenden Informationen über den Antragsfortschritt und die nächsten Schritte erhalten (Ihle, et al., 2022, S. 41). Zur Gewährleistung von Transparenz hinsichtlich der Datennutzung wird seitens des FDZ-Gesundheit ein öffentliches Register geführt. Dieses Register beinhaltet für jeden Antrag Informationen bezüglich der Nutzungsberechtigten, des Titels, der Zwecke, des Jahres der Antragsentscheidung sowie eine kurze Ergebnisdarstellung. (Datentransparenzverordnung, 2020, S. 7). Gemäß dem GDNG ist die Verarbeitung von Gesundheitsdaten im Rahmen eines Forschungsvorhabens ohne die Einwilligung der betroffenen Personen nur unter der Voraussetzung zulässig, dass die Verantwortlichen das Vorhaben vor Beginn der Datenverarbeitung in einem von der Weltgesundheitsorganisation anerkannten Primärregister für klinische Studien registrieren, sofern dies möglich ist. Eine solche Registrierung ist nicht erforderlich, wenn das Forschungsvorhaben bereits aufgrund anderer gesetzlicher Vorgaben an anderer Stelle registriert wurde (Gesundheitsdatennutzungsgesetz, 2024, S. 7).

Der Antragsprozess wird im Vergleich zu bisherigen Möglichkeiten effizient und schnell ablaufen. Datennutzende brauchen nach erfolgreicher Registrierung lediglich den digitalen Antrag zu stellen und erhalten anschließend bei Genehmigung Zugriff auf einen Zuschnitt der gewünschten Daten in einem zu diesem Zweck erzeugten

virtuellen Analyseraum. Der restliche Prozess wird im Hintergrund durch Mitarbeitende des FDZ-Gesundheit bearbeitet (Ludwig, et al., 2024, S. 134f.).

Im Falle einer positiven Prüfung des Antrags wird den Nutzungsberechtigten für einen begrenzten Zeitraum ein Scientific-Use-File zur Verfügung gestellt, welches speziell auf das jeweilige Projekt zugeschnittene Daten beinhaltet. Es handelt sich dabei nur um einen vollständig verfremdeten Auszug der Daten. Der Datenauszug ermöglicht die Entwicklung und Anpassung von Analyseskripten, um einen iterativen Forschungsprozess zu unterstützen. Die Bereitstellung der Scientific-Use-Files erfolgt in gesicherten, virtuellen Analyseräumen, sodass ein unberechtigtes Herunterladen der sensiblen Gesundheitsdaten verhindert wird. Die mittels der Analyseskripte auf dem Originaldatensatz generierten Resultate werden seitens des FDZ-Gesundheit einer Prüfung unterzogen, um das Risiko einer Re-Identifikation zu evaluieren. Sofern erforderlich, werden weitere Maßnahmen zur Minimierung des Risikos durchgeführt, bevor die Ergebnisse an die Nutzungsberechtigten übermittelt werden (Ihle, et al., 2022, S. 42). Die Forschenden können ihr Skript über das Portal des FDZ einreichen und daraufhin Zwischenergebnisse oder ein finales Ergebnis anfordern (Ludwig, et al., 2024, S. 135f.). In dem virtuellen Analyseraum können die Tools Python, R oder SQL zur Erstellung von Analyse-Skripten verwendet werden, so teilt es das FDZ-Gesundheit über LinkedIn mit (Forschungsdatenzentrum Gesundheit, 2024c).

Die zentrale DK-Stelle für Gesundheitsdaten konzentriert sich darauf, als Vermittler zwischen Datennutzenden und Datenhaltern zu agieren. Sie unterstützt Datennutzende umfassend bei der Beantragung von Gesundheitsdatennutzung und -verknüpfung (Henke & Schneider, 2024, S. 219). Laut § 3 Abs. 1 GDNG wird die DK-Stelle beim BfArM eingerichtet und übernimmt Aufgaben wie die Bereitstellung von Informationen, die Beratung der Antragstellenden sowie die Entgegennahme von Anträgen auf Gesundheitsdatennutzung, die sie an die entsprechenden datenhaltenden und zuständigen Stellen weiterleitet. Darüber hinaus ist sie laut § 3 Abs. 2 GDNG befugt, Anträge zur Verknüpfung von FDZ-Daten mit Krebsregistern zu genehmigen. Sie erstellt zudem Konzepte zur Nutzung sicherer Umgebungen und entwickelt Konzepte zur Weiterentwicklung der DK-Stelle als eigenständige Institution (Weichert, 2024, S. 68).

Gemäß § 4 Abs. 5 GDNG erfolgt bei Erteilung einer Genehmigung zur Verknüpfung mit dem klinischen Krebsregister eine Verknüpfung der Daten der Krebsregister in einer gesicherten Verarbeitungsumgebung mit den Daten des FDZ-Gesundheit. Für die Erstellung einer Forschungskennziffer wird die Krankenversichertennummer

herangezogen, wobei die Vertrauensstelle gemäß § 303c SGB V involviert ist. Die konzeptionelle Ausarbeitung des genauen Ortes der Verknüpfung durch die DK-Stelle steht noch aus. Ebenso soll die DK-Stelle Konzepte zur „Verknüpfung und gemeinsamen Verarbeitung von pseudonymisierten Gesundheitsdaten verschiedener datenhaltender Stellen“ entwickeln. Nach der Verknüpfung stehen die pseudonymisierten Einzeldatensätze den Forschenden zur Verfügung (Straub & Kowalski, 2024, S. 740).

In Übereinstimmung mit der in § 8 GDNG festgelegten Publikationspflicht sind Forschungsvorhaben, bei denen personenbezogene Daten ohne Einwilligung der betroffenen Personen verarbeitet wurden, verpflichtet, die Forschungsergebnisse innerhalb von 24 Monaten nach Projektabschluss in einer wissenschaftlichen Publikation zu veröffentlichen. Der hier beschriebene Ansatz gewährleistet die erforderliche Transparenz und erfüllt zugleich die Informationspflicht gegenüber der Öffentlichkeit gemäß Artikel 13 und 14 der Datenschutzgrundverordnung (Henke & Schneider, 2024, S. 222).

Vorrausichtlich ab Anfang 2025 stehen die Abrechnungsdaten der gesetzlich Versicherten auf Antrag zu Verfügung. Ab dem Sommer 2025 sollen auch die Daten der elektronischen Patientenakte zur Verfügung stehen (Dannenberg, 2024, S. 17). Das FDZ-Gesundheit ist dazu angehalten, die Reihenfolge der Antragsbearbeitung anhand objektiver Kriterien festzulegen. Die entsprechenden Kriterien werden bis zum 31. Januar 2025 auf der Website des FDZ-Gesundheit veröffentlicht. Inhaltsgleiche Folgeanträge werden einem beschleunigten und vereinfachten Antragsverfahren unterzogen (Bundesministerium für Gesundheit, 2024b, S. 18).

Im Rahmen der jüngsten Anpassung der Datentransparenz-Gebührenverordnung wurde eine Neuregelung bezüglich der Kosten für die Bearbeitung eines Antrags auf Datenzugang beim FDZ-Gesundheit implementiert. Die Gebühren für den Zugang zu Gesundheitsdaten umfassen eine Grundgebühr für die Bearbeitung von Anträgen, eine tägliche Gebühr für die Nutzung gesicherter virtueller Umgebungen sowie individuell berechnete Kosten für zusätzliche Leistungen wie Beratungen oder Auswertungen. Die Höhe dieser Gebühren variiert je nach Umfang und Komplexität der angefragten Leistungen (Bundesministerium für Gesundheit, 2024b, S. 23f.) Auf Antrag kann das FDZ-Gesundheit die Gebühren für festgelgte Institutionen auf ein Zehntel ermäßigen. Zu den Institutionen gehören zentrale Organisationen der Gesundheitsversorgung, -forschung und -politik, darunter kassenärztliche Vereinigungen, Hochschulen, Forschungsinstitute, Interessenvertretungen der

Patientinnen und Patienten, Spitzenorganisationen der Leistungserbringer sowie relevante Bundes- und Landesinstitutionen im Gesundheitswesen (Bundesministerium für Gesundheit, 2024b, S. 24f.). Der gesamte Prozess des Zugangs- und Nutzungsprozess gemäß des GDNG ist in Abbildung 2 abgebildet.

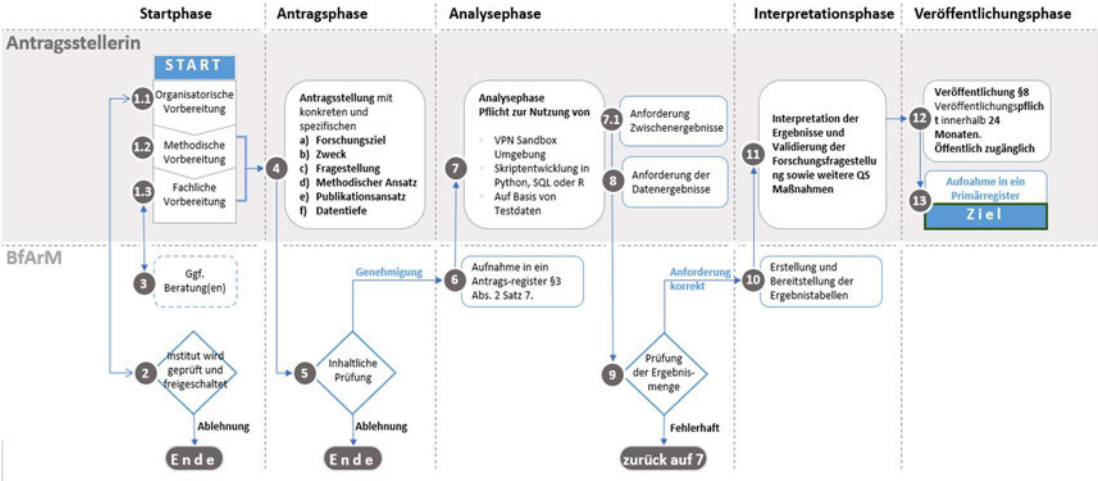

Abbildung 2 Zugangs- und Nutzungsprozess gemäß des GDNG (eigene Darstellung, 2024)

5.3 Möglichkeiten der Datennutzung durch das Gesundheitsdatennutzungsgesetz

Das FDZ-Gesundheit hat über ein LinkedIn Post die erlaubten Nutzungszwecke nach dem GDNG bekanntgegeben, diese umfassen:

„1. Wahrnehmung von Steuerungsaufgaben
2. Verbesserung der Qualität der Versorgung und der Sicherheitsstandards der Prävention, Versorgung und Pflege
3. Planung von Leistungsressourcen, z. B. Krankenhausplanung
4. Wissenschaftliche Forschung
5. Unterstützung politischer Entscheidungsprozesse
6. Analysen zur Wirksamkeit der Kranken- und Pflegekassen
7. Wahrnehmung von Aufgaben der Gesundheitsberichterstattung
8. Wahrnehmung gesetzlicher Aufgaben
9. Entwicklung, Weiterentwicklung und Überwachung der Sicherheit von Arzneimitteln und ähnlichen Produkten
10. Nutzenbewertung von Arzneimitteln und ähnlichen Produkten“

(Forschungsdatenzentrum Gesundheit, 2024a).

Ebenfalls hat das FDZ-Gesundheit über ein LinkedIn Post die verbotenen Nutzungszwecke nach dem GDNG bekanntgegeben, diese umfassen

„1. Entscheidungen hinsichtlich des Abschlusses oder der Ausgestaltung eines Versicherungsvertrags mit Bezug auf eine natürliche Person oder eine Gruppe natürlicher Personen.

2. Treffen von Entscheidungen zum Nachteil einer natürlichen Person auf der Grundlage ihrer elektronischen Gesundheitsdaten.

3. Entwicklung von Produkten oder Dienstleistungen, die Einzelpersonen oder der Gesellschaft insgesamt schaden können, insbesondere illegale Drogen, alkoholische Getränke und Tabakerzeugnisse.

4. Nutzung der Daten für Marktforschung, Werbung und Vertriebstätigkeiten für Arzneimittel, Medizinprodukte und sonstige Produkte.“

In Bezugnahme auf, die in Punkt 4 festgelegte Regelung ist es untersagt, die am FDZ-Gesundheit gespeicherten Daten zu Zwecken der kommerziellen Werbung zu nutzen (Forschungsdatenzentrum Gesundheit, 2024b).

Von 73 Millionen gesetzlich Versicherten werden die pseudonymisierten Gesundheitsdaten mit einer großen Breite an Variablen in dem FDZ-Gesundheit beantragbar sein (Ihle, et al., 2022, S. 39f.). Die erfassten Daten umfassen diagnostische und therapeutische Informationen, Verordnungen von Arzneimitteln sowie Heil- und Hilfsmitteln. Darüber hinaus werden die Art der Leistungserbringung, der Ort und Zeitpunkt der Leistungserbringung, die mit der Krankenkasse abgerechneten Kosten, die Anzahl der Arbeitsunfähigkeitstage sowie der Wohnort der Versicherten berücksichtigt (Günster & Schmuker, 2024, S. 159). Die Pharmazentralnummer ermöglicht eine eindeutige Identifikation von Wirkstoff, Darreichungsform, Wirkstoffstärke und Packungsgröße, was für die Arzneimittelforschung von großer Bedeutung ist, da auf diese Weise eine präzise und nachvollziehbare Dokumentation gewährleistet wird. Des Weiteren erlaubt die Pharmazentralnummer die Analyse zeitlicher Abläufe, welche durch die Verordnungs- und Abgabedaten dokumentiert sind (Schneider & Hess, 2023, S. 24). Die Abrechnungsdaten der gesetzlichen Krankenversicherungen bilden das Versorgungsgeschehen nahezu lückenlos ab. Sie erhalten sektorübergreifend die Daten aller Leistungserbringer und auch die Schnittstellen (Ihle, et al., 2022, S. 33). Gemäß §363 SGB V werden die Abrechnungsdaten in Zukunft zudem durch Daten der elektronischen Patientenakte ergänzt (Ihle, et al., 2022, S. 1).

Die Daten eigenen sich für Longitudinalanalysen und für raumbezogene Analysen, zudem erlauben Sie spezifische Auswertungen zur Mortalität, zur Prävalenz und Inzidenz von Erkrankungen, zur Inanspruchnahme von Gesundheitsleistungen und

Gesundheitsausgaben. Die Erfassung der gesamten Versichertenpopulation ermöglicht die Berücksichtigung von Personengruppen, die in Primärdatenerhebungen selten oder nur schwer zu erfassen sind. Dazu zählen beispielsweise Personen mit seltenen Erkrankungen (Günster & Schmuker, 2024, S. 159). Eine weitere Eigenschaft der Daten ist das weitgehende Fehlen von Selektionseffekten. Die Erfassung der Abrechnungsdaten erfolgt unabhängig von den individuellen Voraussetzungen der Versicherten. In der Konsequenz manifestieren sich keine Erinnerungsbias, sprachlichen oder sozioökonomischen Verzerrungen, wie sie bei einer direkten Datenerhebung bei den Versicherten zu erwarten wären (Ihle, et al., 2022, S. 34).

Da die Daten bereits vorliegen und für die statistische Auswertung nutzbar sind, erweisen sich Analysen mit Sekundärdaten hinsichtlich des Kosten- und Zeitaufwandes als ressourcenschonendere Vorgehensweise im Vergleich zu Primärdatenerhebungen (Slagmann, et al., 2023, S. 89).

Die groß angelegten Datensätze weisen methodische Vorteile auf. Aufgrund der hohen Fallzahl sind kleinere Effekte besser erkennbar. Zudem stehen für seltene Erkrankungen erst genügend Daten für eine fundierte statistische Analyse zur Verfügung (Ihle, et al., 2022, S. 40).

Einzelne Limitationen der GKV-Routinedaten können durch die Kombination verschiedener Datenquellen überwunden werden. Ein exemplarisches Beispiel hierfür ist das Datenlinkage zwischen Routinedaten der gesetzlichen Krankenversicherung (GKV) und Daten der Krebsregister. Infolgedessen können Verlaufs- und Inanspruchnahmedaten aus der Krankenversicherung mit zusätzlichen Informationen aus den Krebsregistern, wie beispielsweise Histologie und Tumorstadium, die in den GKV-Daten nicht enthalten sind, kombiniert werden. (Günster & Schmuker, 2024, S. 160). In Übereinstimmung mit § 4 des GDNG kann die neu etablierte Koordinierungsstelle auf Antrag die Verknüpfung der Daten des FDZ-Gesundheit mit gesetzlich regulierten medizinischen Registern bewilligen. Diese Option ist zunächst auf die KKR gemäß § 65c des Sozialgesetzbuches (SGB) V beschränkt. Diese gesetzlich verankerte Möglichkeit zur Datenverknüpfung markiert einen Wendepunkt für die Forschung mit praxisnahen Versorgungsdaten (Straub & Kowalski, 2024, S. 737).

Patientinnen und Patienten mit seltenen Erkrankungen sind für ihre überdurchschnittliche Bereitschaft ihre Daten für die Forschung zur Verfügung zu stellen bekannt (EURORDIS, 2022).

Die Ergebnisse künftiger Studien, welche auf den Daten des FDZ-Gesundheit basieren, können Versorgerinnen und Versorger wie Ärztinnen und Ärzte bei der Bewältigung ihrer täglichen Praxis unterstützen. Der Zugang zu neuen Erkenntnissen und evidenzbasierten Informationen ermöglicht eine fundierte Entscheidungsfindung bei Diagnosen, Behandlungen und Therapieplänen. Insgesamt können Behandlungsstrategien und Interventionen präzise analysiert und optimiert werden (Ludwig, et al., 2024, S. 132).

Die Bereitstellung von Daten sowie die kontinuierliche Nutzung einer Datenquelle in einem Forschungsdatenzentrum führen zu einer signifikanten Steigerung der Datenqualität. Durch eigene Qualitäts- und Plausibilitätsprüfungen sowie durch das wertvolle Feedback der Nutzenden können wesentliche Verbesserungen realisiert werden. Daher ist ein kontinuierlicher und regelmäßiger Austausch mit der Gemeinschaft der Nutzungsberechtigten von essenzieller Bedeutung (Ihle, et al., 2022, S. 36).

Die Neufassung des § 25b Abs. 1 SGB V erweitert die Möglichkeiten der Kranken- und jetzt auch Pflegekassen zur datengestützten Erkennung individueller Gesundheitsrisiken. Im Rahmen dessen erfolgt eine Präzisierung der Datenverarbeitungszwecke, insbesondere hinsichtlich der Erkennung seltener Erkrankungen, der Erkennung von Krebserkrankungen sowie der Erkennung schwerwiegender Gesundheitsgefährdungen, welche unter anderem durch Arzneimittel verursacht werden können (Moreno, 2024, S. 94). Gemäß § 25b Abs. 4 SGB V sind Kranken- und Pflegekassen dazu berechtigt, die Versicherten zu informieren, sofern eine Auswertung gemäß Abs. 2 auf eine konkrete Gesundheitsgefährdung hinweist. Die Betroffenen sind unverzüglich über das identifizierte Gesundheitsrisiko in einer verständlichen und präzisen Weise zu informieren, wobei eine klare und einfache Sprache zu verwenden ist. Der Hinweis auf eine identifizierte Gefährdung ist dabei mit der Empfehlung zu verbinden, eine ärztliche, zahnärztliche, psychotherapeutische oder pflegerische Beratung in Anspruch zu nehmen. Die Empfehlung ist durch die Kostenträger entsprechend den Vorgaben des Gesetzgebers zu begründen (Moreno, 2024, S. 94f.).

Der neu eingeführte § 303b Abs. 1 Satz 2 SGB V nimmt das Anliegen auf, zeitaktuelle Untersuchungen durchführen zu können. Ein neues Übermittlungsschema sieht statt einer halbjährlichen, eine quartalsweise rollierende Datenlieferung vor, sodass eine Abwägung zwischen Datenvollständigkeit (= Qualität) und Aktualität (= zeitkritische Komponente) möglich wird (Straub & Kowalski, 2024, S. 739).

Der § 6 regelt zum einen die Möglichkeiten der Eigenforschung, was für viele datenhaltende Versorgungsorganisationen Sicherheit schafft. Hierdurch soll eine Annäherung von Forschung und Versorgung erreicht werden (Ortmann, et al., 2024, S. 198f.). § 6 des GDNG erlaubt unter bestimmten Voraussetzungen, wie zum Beispiel öffentlicher Förderung und Zustimmung der Datenschutzaufsichtsbehörde, die Weitergabe von Versorgungsdaten sowie deren gemeinsame Auswertung für Forschungszwecke. Dadurch entstehen umfassende Möglichkeiten zur Nutzung der Daten zertifizierter Zentren (Ortmann, et al., 2024, S. 199).

5.4 Fallbeispiel: Zugangs- und Nutzungsprozess für Forschung zu „Orphan Diseases“ bei Kindern

Im Folgenden wird ein praxisnaher Einblick in die Antragsstellung gegeben. Dazu wird exemplarisch dargestellt, wie eine Antragsstellung gemäß dem Antragsportal des FDZ-Gesundheit aussehen kann. Aufgrund der begrenzten Möglichkeiten einer Bachelorarbeit wurde dabei auf eine ausführliche Darstellung verzichtet.

- **Projekttitel:** "Seltene Erkrankungen im Fokus: Diagnosehäufigkeit im Vergleich zwischen DAK-Gesundheit und Gesamt-GKV"
- **Forschungsvorhaben:** Das Projekt untersucht die Diagnosehäufigkeit seltener Erkrankungen bei verschiedenen Altersgruppen (Kinder, Jugendliche und Erwachsene). Der Schwerpunkt liegt auf einem Vergleich zwischen DAK-Gesundheit und der Gesamt-GKV

- **Forschungsziel:** Das Ziel dieser Arbeit ist es, eine detaillierte Analyse zur Prävalenz seltener Erkrankungen in unterschiedlichen Altersgruppen durchzuführen. Besonderer Fokus liegt auf der Gegenüberstellung der DAK-Gesundheit und der Gesamt-GKV, um spezifische Unterschiede in den Daten verschiedener Krankenversicherungen und deren möglichen Einfluss auf die Versorgungsqualität zu evaluieren

- **Zulässige Zwecke:** Verbesserung der Qualität der Versorgung und der Sicherheitsstandards der Prävention, Versorgung und Pflege,
  Wissenschaftliche Forschung,
  Wahrnehmung von Aufgaben der Gesundheitsberichterstattung
- **Wissenschaftlicher Hintergrund:** siehe Kapitel „Seltene Krankheiten bei Kindern“

- **Fragestellungen:** Wie häufig werden seltene Erkrankungen bei Kindern, Jugendlichen und Erwachsenen in Deutschland diagnostiziert?

  Welche Unterschiede bestehen zwischen den Daten der DAK-G und der Gesamt-GKV?

- **Methodischer Ansatz:** Retrospektive Kohortenanalyse, Abrechnungs- und Diagnosedaten aus der DAK-G und der Gesamt-GKV, Statistische Analyse der Diagnosehäufigkeit seltener Erkrankungen, Subgruppenanalyse nach Altersgruppen
- **Art der Testdaten:** Pseudonymisierte Einzeldatensätze der Originaldaten
- **Datenauswahl:** Alter, Geschlecht und Wohnort, ICD- Codes
- **Beschreibung Ergebnismenge:** Individuelle Ergebnistabellen

5.5 Limitationen und Herausforderungen der Datennutzung durch das Gesundheitsdatennutzungsgesetz

Die Forschung mit den Daten des FDZ-Gesundheit ist mit zwei zentralen Herausforderungen verbunden. Einerseits sind die Daten primär Abrechnungsdaten und somit nicht originär für die Forschung erhoben worden, woraus sich die Interpretation der Ergebnisse erschwert. Andererseits können Verzerrungen durch die Datenlieferzeitpunkte auftreten. Bei der Verwendung der Daten sind, wie bereits teilweise angemerkt, einige Einschränkungen zu berücksichtigen. Die Daten wurden primär für die Abrechnung von Leistungen gegenüber den Krankenkassen erstellt und stellen somit administrative Daten dar. Infolgedessen sind Angaben zu Untersuchungsergebnissen wie Laborwerten sowie patientenbezogene Informationen wie Lebensqualität oder Zufriedenheit nicht verfügbar. Ebenso sind soziodemografische Daten lediglich in begrenztem Umfang verfügbar, was die Berücksichtigung von Confoundern in analytischen Studien erheblich erschwert. Diagnosen sind in der Regel nach ICD-10 kodiert, wodurch Informationen zum Schweregrad oder Krankheitsstadium lediglich in diesem Rahmen verfügbar sind, was die Aussagekraft für bestimmte Studien begrenzen kann. Des Weiteren ist zu berücksichtigen, dass für pharmakoepidemiologische Untersuchungen keine Dosierungsangaben vorliegen. Diese sind jedoch von zentraler Bedeutung, um die Adhärenz und Exposition adäquat bewerten zu können. Zudem ist der Verordnungsgrund, ICD-10-Diagnose, nicht direkt mit der Verordnung verknüpft, was die Aussagekraft der Ergebnisse weiter einschränkt (Schubert, 2022, S. 8).

Die Datenqualität und -validität ist im Hinblick auf den Erhebungskontext zu betrachten. Bei der Pauschalabrechnung ambulanter ärztlicher Leistungen kann kein

taggenauer Leistungsbezug hergestellt werden, diese liegen nur auf Quartalsebene vor (Günster & Schmuker, 2024, S. 159).
Die Analyse von Routinedaten birgt die Herausforderung, alle für die Forschungsfrage relevanten Informationen zu operationalisieren und zu validieren. Oftmals sind bestimmte Informationen in den Daten nicht vorhanden, lediglich suboptimal operationalisiert oder lediglich als Ersatzparameter verfügbar. Der Anteil an fehlenden oder unplausiblen Werten variiert zudem in Abhängigkeit von der Datenquelle und muss vor Beginn einer Sekundärdatenanalyse identifiziert und berücksichtigt werden. Zusätzlich sollten mögliche Einflussfaktoren auf die Datendokumentation sorgfältig überprüft werden. Im Bereich der klinischen Dokumentation, im Gegensatz zu Abrechnungsdaten, ist die Datenqualität oft gering und die Dokumentation weist eine geringe Standardisierung auf (Slagmann, et al., 2023, S. 89).

Gemäß der aktuellen Regelung unterliegt jede Krankenkasse ab dem Jahr 2025 der Pflicht, etwa sechs Wochen nach Quartalsende die Daten der letzten vier Quartale zu übermitteln. Die Übermittlungspflicht beginnt mit dem ersten Quartal des Jahres 2025. Je nach Art der Leistung und Abstand zum Zeitpunkt der Datenabfrage wird die Fülle der Datenlieferungen jedoch variieren. Die Datenstände weisen hinsichtlich ihrer Vollständigkeit und langfristigen Gültigkeit unterschiedliche Eigenschaften auf (Schulte, 2024, S. 140).
Etwa sechs Wochen nach Quartalsende liegen den Krankenkassen ca. 96 % der Krankenhausdaten des Vorquartals vor, jedoch nur ca. 11 % der Heilmitteldaten, wobei ca. 6 % der Krankenhausdaten nachträglich korrigiert werden. Die Heilmitteldaten sind meist unvollständig, da der letzte Monat des Quartals fehlt und etwa ein Drittel der Datensätze der ersten beiden Monate nachträglich durch Abrechnungsprüfungen korrigiert werden. Ambulante Abrechnungsdaten liegen erst drei bis sechs Monate nach Quartalsende vor (Schulte, 2024, S. 140). Um diesen Verzögerungen entgegenzuwirken, fordert § 295b GDNG ab dem ersten Quartal 2025 eine frühzeitige Lieferung ungeprüfter Arztdaten durch die Kassenärztlichen Vereinigungen, die die Krankenkassen zunächst verwenden, bis geprüfte Daten vorliegen. Darüber hinaus sind veraltete Daten eines Quartals zu löschen, sobald neue Meldungen eingehen, wobei Nutzungsberechtigte, die diese Daten bereits erhalten haben, damit weiterarbeiten können. Für die Datentransparenzstelle bedeutet dies, dass sie zukünftig unterschiedliche Datenstände für laufende Forschungsprojekte vorhalten muss. Forschende wiederum müssen entscheiden, ob vorläufige Datenstände für die Beantwortung ihrer Forschungsfragen geeignet sind, welches eine fundierte Kenntnis der Datenflüsse voraussetzt. Insbesondere bei

Leistungsdaten wie Arzneimitteln sind Lücken noch relativ einfach zu erkennen, während bei anderen Leistungsarten komplexere Verzerrungen auftreten können (Schulte, 2024, S. 141). Eine Lösung dieses Problems ist derzeit nicht in Sicht, da die Verzögerungen bei der Datenverfügbarkeit direkt aus den Abrechnungsregeln der einzelnen Leistungserbringer resultieren. So können Heilmittelerbringer ihre Leistungen bis zu neun Monate nach Behandlungsende abrechnen und Hebammen haben die Möglichkeit, ihre Abrechnungen bis zum 30. Juni des Folgejahres einzureichen (Schulte, 2024, S. 142). Bei den Routinedaten der GKV gibt es aufgrund der Turnuslängen eine Verzögerung, so dass hochaktuelle Daten nicht zur Verfügung stehen (Ihle, et al., 2022, S. 36).
Ziel durch das GDNG war es, repräsentative und qualitativ hochwertige Daten zeitnah zur Verfügung zu stellen. Zwischen den genannten Zielen besteht ein Zielkonflikt (Schulte, 2024, S. 143).

Das GDNG schafft eine zeitliche Lücke in der Datenbereitstellung, die das gesamte Jahr 2023 sowie das erste Quartal 2024 betrifft. In dieser Übergangszeit sind die Daten nicht in den jährlichen Datenlieferungen enthalten und stehen gleichzeitig noch nicht in den rollierenden Quartalslieferungen zur Verfügung (Schulte, 2024, S. 140).

Die Datenbasis des FDZ-Gesundheit basiert auf den Abrechnungsdaten der GKV, wodurch sich eine Beschränkung auf Informationen ergibt, die im Rahmen der GKV abgerechnet werden. Einige Leistungsbereiche der Gesundheitsversorgung sowie Teile der Bevölkerung werden in den FDZ-Gesundheit Daten nicht abgebildet. Beispielsweise die Rehabilitation zu Lasten der gesetzlichen Rentenversicherung und Privatversicherte (March, et al., 2023, S. 147). Ebenfalls fehlen Informationen zu sogenannten individuellen Gesundheitsleistungen und freiverkäuflichen „Over the Counter“ Präparate (Ihle, et al., 2022, S. 35).

Aufgrund von „Upcoding“ müssen Verzerrungen bei der Routinedatenanalyse und Interpretation als Möglichkeit bedacht werden (Ihle, et al., 2022, S. 33). Upcoding tritt auf, wenn ein Gesundheitsdienstleister Codes für schwerere als diagnostizierte Erkrankungen eingereicht hat, damit die Patientin oder der Patient eine höhere Erstattung erhält (Phillips & Cohen, 2018).

Bei der Aufbereitung von Routinedaten werden „Verrauschungstechniken“ eingesetzt. Die Anwendung von Techniken wie Differential Privacy, k-Anonymisierung und Anderen kann die statistische Unsicherheit der Analyseergebnisse erhöhen und das Risiko systematischer Verzerrungen steigern. Spezifische Patientengruppen wie Patientinnen und Patienten mit seltenen Erkrankungen wären von diesen

Herausforderungen besonders betroffen (Schmitt, et al., 2023, S. 217). Schon einzelne Widersprüche können bei Auswertungen mit geringen Fallzahlen eine Auswirkung haben (Schmitt, et al., 2023, S. 217). Haushalte mit Kindern sind bei der Zustimmung der Registerspeicherung zurückhaltender (Haug, et al., 2024, S. 3).

Einrichtungsübergreifende Analysen von Patientenakten sind zumindest dann eingeschränkt, wenn personenbezogene Daten außerhalb der jeweiligen Einrichtung genutzt oder zusammengeführt werden sollen. Diese Beschränkung auf die Forschung innerhalb der eigenen Institution stellt ein erhebliches Hindernis für Versorgungsanalysen dar, da gerade die Verknüpfung von Gesundheitsdaten aus unterschiedlichen Quellen als besonders wertvoll angesehen wird (Straub & Kowalski, 2024, S. 738).

6 Diskussion

Im folgenden Kapitel werden die in der Literatur gefunden Ergebnisse zusammengefasst und interpretiert, zudem werden die Fragestellungen beantwortet und die Stärken und Schwächen des Scoping Reviews aufgezeigt.

6.1 Zusammenfassung der Ergebnisse

Im Rahmen der vorliegenden Arbeit wurde ein Scoping Review zum Thema des Zugangs- und Nutzungsprozesses im Rahmen des GDNG durchgeführt. Darüber hinaus wurden die spezifischen Möglichkeiten und Limitationen, die sich aus dem GDNG für die Forschung zu Ophan Diseases bei Kindern ergeben, untersucht.

Der Zugang und die Nutzung von Gesundheitsdaten gemäß dem GDNG werden über zwei zentrale Institutionen organisiert: das FDZ-Gesundheit und die DK-Stelle beim BfArM (Henke & Schneider, 2024, S. 221). Die Nutzung der Daten erfolgt zweckgebunden und umfasst unter anderem die Versorgungsforschung, Prävention, KI-Entwicklung sowie sektorenübergreifende Analysen (Henke & Schneider, 2024, S. 221). Die Nutzung der Daten durch interessierte Institutionen erfolgt nach Stellung eines detaillierten Antrags im FDZ-Portal, in welchem die Nutzungszwecke, Methoden und benötigten Daten spezifiziert werden müssen (Ludwig, et al., 2024, S. 134). Nach einer formalen und inhaltlichen Prüfung, die innerhalb eines Zeitraums von maximal vier Monaten abgeschlossen sein muss, wird den genehmigten Projekten der Zugriff auf einen virtuellen Analyseraum gewährt (Ihle, et al., 2022, S. 41f.; Datentransparenzverordnung, 2020, S. 7). Im Rahmen dessen werden Daten in sicherer Weise bereitgestellt, wobei ein Export von Resultaten erst nach einer Prüfung durch das FDZ-Gesundheit erfolgt, um das Risiko einer Re-Identifikation zu

minimieren (Ihle, et al., 2022, S. 42). Im Rahmen der Nutzung des Analyseraums stehen den Forschenden Tools wie Python, R und SQL zur Verfügung (Forschungsdatenzentrum Gesundheit, 2024c).

Die Funktion der DK-Stelle besteht in der Vermittlung zwischen Datennutzenden und datenhaltenden Stellen (Henke & Schneider, 2024, S. 219). Zu den Aufgaben der DK-Stelle zählt die Beratung, die Genehmigung von Anträgen auf Datenverknüpfung sowie die Entwicklung von Konzepten für sichere Verarbeitungsumgebungen (Weichert, 2024, S. 68). Es sei insbesondere auf die Möglichkeit hingewiesen, Daten des FDZ-Gesundheit mit denen von Krebsregistern zu verknüpfen. Nach erfolgter Verknüpfung werden die hier verwendeten pseudonymisierten Einzeldatensätze in sicheren Umgebungen bereitgestellt (Straub & Kowalski, 2024, S. 740). Die Transparenz der Datennutzung wird durch ein öffentliches Register des FDZ-Gesundheit sowie die Publikationspflicht für Forschungsergebnisse gewährleistet. Projekte, die personenbezogene Daten ohne Einwilligung nutzen, sind dazu verpflichtet, ihre Ergebnisse innerhalb von 24 Monaten zu veröffentlichen (Henke & Schneider, 2024, S. 222). Ab dem Jahr 2025 stehen sowohl Abrechnungsdaten der gesetzlich Versicherten als auch Daten der elektronischen Patientenakte zur Verfügung (Dannenberg, 2024, S. 17). Eine neue Gebührenordnung legt für die Antragsbearbeitung und die Bereitstellung der Daten klare Kosten fest, die von der Dauer und Komplexität des jeweiligen Falls abhängig sind (Bundesministerium für Gesundheit, 2024b, S. 23f.).

Das GDNG eröffnet vielfältige Möglichkeiten zur Nutzung von Gesundheitsdaten. Zu den erlaubten Nutzungszwecken gehören unter anderem die wissenschaftliche Forschung, die Qualitätssicherung in der Versorgung, die Planung von Leistungsressourcen sowie die Unterstützung politischer Entscheidungsprozesse (Forschungsdatenzentrum Gesundheit, 2024a). Gleichzeitig grenzt das GDNG die Nutzung klar ab, indem es kommerzielle Zwecke wie Werbung oder die Nutzung zum Nachteil Einzelner verbietet. Diese Abgrenzung schafft einen rechtlichen Rahmen, der eine datenschutzkonforme und zielgerichtete Nutzung sicherstellt (Forschungsdatenzentrum Gesundheit, 2024b). Die im FDZ-Gesundheit verfügbaren Daten von rund 73 Millionen gesetzlich Versicherten umfassen ein breites Spektrum an Variablen wie diagnostische und therapeutische Informationen, Arzneimittelverordnungen sowie Abrechnungsdaten (Ihle, et al., 2022, S. 39f.; Günster & Schmuker, 2024, S. 159). Besonders hervorzuheben ist, dass diese Daten das Versorgungsgeschehen nahezu vollständig abbilden und keine Selektionseffekte aufweisen, so dass eine weitgehend unverzerrte Analyse möglich ist (Ihle, et al.,

2022, S. 33f.). Aufgrund der Größe und Tiefe des Datensatzes sind die Daten für Analysen seltener Erkrankungen geeignet, da ausreichend große Stichproben generiert werden können. Auch Längsschnitt- und räumliche Analysen sowie die Betrachtung von Prävalenzen, Inzidenzen und Mortalitätsraten sind möglich (Günster & Schmuker, 2024, S. 159). Ein wesentlicher Vorteil der Nutzung dieser Sekundärdaten liegt in der Ressourcenschonung hinsichtlich Kosten und Zeit (Slagmann, et al., 2023, S. 89). Gleichzeitig können durch die hohe Datenqualität und umfangreiche Plausibilitätsprüfungen verlässliche Ergebnisse erzielt werden, die für die Versorgungsforschung und die Verbesserung der Versorgungspraxis von großem Nutzen sind (Ihle, et al., S. 36). Die Pharmazentralnummer ermöglicht dabei die präzise Erfassung und Analyse von Arzneimitteldaten, während die Verknüpfung mit medizinischen Registern zusätzliche Kontextinformationen, zum Beispiel zu Tumorstadien liefert (Schneider & Hess, 2023, S. 24; Günster & Schmuker, 2024, S. 160). Die gesetzlich geregelte Möglichkeit der Verknüpfung solcher Daten stellt einen Wendepunkt in der Forschung dar, insbesondere für praxisnahe Versorgungsstudien (Straub & Kowalski, 2024, S. 737). Auch die Neuerungen im GDNG, wie die quartalsweise Datenübermittlung und die explizit erlaubte Eigenforschung der datenhaltenden Organisationen, fördern die Aktualität und Relevanz der Forschungsergebnisse (Straub & Kowalski, 2024, S. 739). Diese Maßnahmen tragen dazu bei, Forschung und Versorgung enger zu verzahnen und die Nutzung von Gesundheitsdaten im Sinne der Patientensicherheit und Versorgungsqualität zu optimieren (Ortmann, et al., 2024, S. 198f.).

Dennoch ist die Nutzung von den FDZ-Gesundheitsdaten für Forschungszwecke ist mit zahlreichen Herausforderungen verbunden, die sowohl in der Natur der Daten als auch in organisatorischen Rahmenbedingungen begründet sind. Da die Daten primär Abrechnungszwecken dienen, enthalten sie keine klinischen Untersuchungsergebnisse oder patientenbezogene Informationen wie Lebensqualität oder Zufriedenheit. Auch soziodemographische Informationen sind nur eingeschränkt verfügbar, was die Berücksichtigung von Confoundern erschwert. Die Diagnosen sind in der Regel nach ICD-10 kodiert, geben aber keine detaillierten Informationen zum Stadium oder Schweregrad der Erkrankung. Darüber hinaus fehlen für pharmakoepidemiologische Studien wichtige Angaben wie Dosierungen und Verordnungsgründe, was die Aussagekraft dieser Analysen weiter einschränkt (Schubert, 2022, S. 8).

Die Datenqualität ist je nach Erhebungskontext sehr unterschiedlich (Günster & Schmuker, 2024, S. 159). Während Abrechnungsdaten relativ standardisiert sind, weisen klinische Dokumentationen häufig eine geringe Standardisierung auf. Routinedatenanalysen stehen vor der Herausforderung, fehlende oder unplausible Werte zu identifizieren und alle relevanten Informationen adäquat zu operationalisieren (Slagmann, et al., 2023, S. 89). Hinzu kommen zeitliche Verzögerungen in der Datenverfügbarkeit (Ihle, et al., 2022, S. 36). So sind Krankenhausdaten schneller verfügbar als Heilmitteldaten oder ambulante Abrechnungsdaten (Schulte, 2024, S. 140). Mit der Einführung des GDNG soll die Bereitstellung durch rollierende Quartalslieferungen ab 2025 beschleunigt werden, allerdings verbleiben gewisse Verzögerungen aufgrund der Abrechnungsregeln der Leistungserbringer (Ihle, et al., 2022, S. 36).

Die Verarbeitung von Routinedaten wird zusätzlich durch methodische Einschränkungen erschwert. Datenschutzmaßnahmen wie Verrauschungstechniken erhöhen die Unsicherheit der Analysen und können insbesondere bei kleinen Fallzahlen, zum Beispiel in der Forschung zu seltenen Erkrankungen, zu Verzerrungen führen (Schmitt, et al., 2023, S. 217). Zudem muss bei der Interpretation von Routinedaten das Risiko systematischer Fehler wie „Upcoding“ berücksichtigt werden (Ihle, et al., 2022, S. 33). Eine weitere Herausforderung ergibt sich aus der begrenzten Datenbasis. Bestimmte Versorgungsbereiche wie die Rehabilitation außerhalb der gesetzlichen Krankenversicherung oder die Versorgung von Privatversicherten sind in den Daten nicht abgebildet (March, et al., 2023, S. 147). Ebenso fehlen Leistungen wie individuelle Gesundheitsleistungen, kurz: IGeL oder frei verkäufliche Arzneimittel (Ihle, et al., 2022, S. 35). Darüber hinaus hat das GDNG eine Datenlücke für den Übergangszeitraum 2023/2024 geschaffen, da in dieser Phase weder die Jahreslieferungen noch die neuen rollierenden Quartalslieferungen vollständig vorliegen (Schulte, 2024, S. 140).

Forschungen durch Arbeitsgemeinschaften sind ebenfalls eingeschränkt, da personenbezogene Daten häufig nicht einrichtungsübergreifend genutzt oder verknüpft werden dürfen. Dies erschwert insbesondere Versorgungsanalysen, die auf eine umfassende Verknüpfung von Gesundheitsdaten angewiesen sind (Straub & Kowalski, 2024, S. 738).

Insgesamt zeigt sich, dass die Daten der FDZ-Gesundheit zwar eine wertvolle Grundlage für die Forschung darstellen, ihre Nutzung aber durch methodische, technische und organisatorische Restriktionen eingeschränkt ist. Für die Forschung

zu Seltenen Erkrankungen, die auf präzise und umfassende Daten angewiesen ist, bleiben diese Einschränkungen eine große Herausforderung.

6.2 Interpretation der Ergebnisse

Die Ergebnisse des Scoping Reviews zeigen ein ambivaltentes Bild. Einerseits bietet das GDNG mit seiner strukturierten und zentralisierten Verarbeitung von Gesundheitsdaten durch das FDZ-Gesundheit und die DK-Stelle neue Möglichkeiten für die datenbasierte Forschung, insbesondere zu seltenen Erkrankungen (Henke & Schneider, 2024, S. 221). Anderseits bleiben zentrale Herausforderungen, die die Nutzug der Daten einschränken.

Das GDNG erleichtert den Zugang zu Gesundheitsdaten und erweitert die Nutzungsmöglichkeiten. Mit dem FDZ-Gesundheit und der DK-Stelle werden zentrale, transparente und rechtssichere Plattformen geschaffen, die Forschenden klare Anlaufstellen bieten (Henke & Schneider, 2024, S. 221). Durch die geplante Standardisierung der Prozesse und die Veröffentlichung von Antragskriterien wird die Transparenz erhöht, und die detaillierte Antragsverfolgung sowie die beschleunigte Bearbeitung von Folgeanträgen fördern die Effizienz (Bundesministerium für Gesundheit, 2024b, S. 18). Besonders hervorzuheben ist der Paradigmenwechsel vom Akteursbezug zum Zweckbezug (Henke & Schneider, 2024, S. 221). Statt der institutionellen Zugehörigkeit rückt der Forschungszweck in den Mittelpunkt, was eine präzisere Prüfung der ethischen, rechtlichen und gesellschaftlichen Vertretbarkeit von Projekten ermöglicht. Dieser Fokus erlaubt es, Projekte, die einen nachweislichen Beitrag zur Verbesserung von Gesundheitssystemen, Versorgung oder öffentlicher Gesundheit leisten, gezielt zu priorisieren (Forschungsdatenzentrum Gesundheit, 2024a). Die Anträge werden am FDZ-Gesundheit von Wissenschaftlerinnen und Wissenschaftlern geprüft, daher ist es wichtig den Antrag wissenschaftlich und präzise zu stellen (Ludwig, et al., 2024, S. 134). Das benötigt vorab eine intensive Beschäftigung mit dem Forschungsthema und eine gründliche Vorbereitung. Besonders relevant ist dieser Ansatz für die Forschung zu seltenen Erkrankungen, da Projekte mit klarem Zweckbezug einfacher Zugang erhalten können.

Mit einer nahezu vollständigen Abbildung des Versorgungsgeschehens durch die Daten von 73 Millionen gesetzlich Versicherten können verlässliche Prävalenz- und Inzidenzschätzungen durchgeführt werden, die gerade bei seltenen Erkrankungen von großer Bedeutung sind (Ihle, et al., 2022, S. 39f.; Bundesministerium für Gesundheit, 2024a). Auch die Möglichkeit zur Datenverknüpfung, beispielsweise mit Krebsregistern, eröffnet neue Perspektiven, etwa für Langzeitbeobachtungen von

Therapien und Krankheitsverläufen (Straub & Kowalski, 2024, S. 737). Die kontinuierliche Übermittlung von Abrechnungsdaten sowie die geplante Integration der elektronischen Patientenakte ab 2025 erweitern die Datenbasis und erhöhen ihre Aktualität und Tiefe. Das stellt einen entscheidenden Vorteil für die Analyse von Versorgungspfaden, Arzneimitteltherapien und Behandlungsergebnissen bei seltenen Erkrankungen (Ihle, et al., 2022, S. 36; Dannenberg, 2024, S. 17)

Trotz dieser Potenziale bleiben methodische, technische und organisatorische Einschränkungen. Die Daten des FDZ-Gesundheit sind primär für Abrechnungszwecke ausgelegt und enthalten keine klinischen Parameter wie Laborwerte oder Angaben zur Lebensqualität (Schubert, 2022, S. 8). Diese Lücken begrenzen die Aussagekraft der Analysen, insbesondere für komplexe Fragestellungen. Zudem variieren die Datenqualitäten erheblich. Während Abrechnungsdaten standardisiert sind, fehlen bei klinischen Dokumentationen häufig einheitliche Kodierungen, was zu unvollständigen oder ungenauen Daten führen kann (Günster & Schmuker, 2024, S. 159; Slagmann, et al., 2023, S. 89). Anonymisierungsmaßnahmen und Verrauschungen können bei kleinen Fallzahlen, wie sie bei seltenen Erkrankungen typisch sind, zu Verzerrungen führen und die Validität der Ergebnisse beeinträchtigen (Schmitt, et al., 2023, S. 217). Hinzu kommen organisatorische Hürden: Die Nutzung der Daten des FDZ-Gesundheit erfordert erhebliche Ressourcen und Expertise, insbesondere bei der Datenaufbereitung und -analyse. Für kleinere Forschungsteams oder Institutionen stellen diese Anforderungen oft eine hohe Hürde dar. Auch hohe Gebühren und komplexe Antragsverfahren könnten abschreckend wirken.

Für Kinder, die die größte Gruppe bei seltenen Erkrankungen ausmachen, ist die Zurückhaltung der Daten der Familien ein zusätzlicher Aspekt. Während Erwachsene mit seltenen Erkrankungen oft eine hohe Bereitschaft zur Datenspende zeigen, sind Familien mit betroffenen Kindern skeptischer. Schon wenige Widersprüche können dazu führen, dass wichtige Daten fehlen, was die Forschung erheblich behindern kann (EURORDIS, 2022; Haug, et al., 2024, S. 3).

Zusammenfassend bietet das GDNG mit seinen überarbeiteten Regularien eine innovative Grundlage für die datenbasierte Gesundheitsforschung. Es ermöglicht nicht nur wissenschaftliche Erkenntnisgewinne, sondern schafft auch einen direkten Nutzen für die Versorgungspraxis und die Patientenversorgung. Besonders für seltene Erkrankungen eröffnen sich neue Chancen durch die Breite und Tiefe der verfügbaren Daten. Gleichzeitig bleibt der Zugang und die Nutzung dieser Daten

durch methodische, technische und organisatorische Herausforderungen begrenzt, was insbesondere für kleinere Forschungsteams und Projekte zu einer Barriere werden kann.

Die Antragsstellung wird voraussichtlich ab Anfang des zweiten Quartals 2025 möglich sein, während die Daten des ePA ab dem Herbst 2025 verfügbar sein werden. Die Implementierung der Register in das FDZ-Gesundheit wird bis zum europäischen Datenraum 2027 andauern. Trotz der als anspruchsvoll einzustufenden Antragsprozesse ist davon auszugehen, dass mit dem Start des FDZ-Gesundheits eine große Anzahl an Anträgen eingehen wird.

6.3 Beantwortung der Fragestellungen

Nachfolgend werden die beiden Fragestellungen basierend auf den Ergebnissen des Scoping Reviews beantwortet.

6.3.1 Wie erfolgt der Zugang zu Gesundheitsdaten und deren Nutzung im Rahmen des Gesundheitsdatennutzungsgesetzes?

Der gesamte Prozess ist in Abbildung 3 abgebildet. Bevor ein Institut Zugang zu den Daten des FDZ-Gesundheit erhält, muss es sich zunächst im FDZ-Gesundheit registrieren. Nach der Registrierung erfolgt eine Prüfung des Instituts. Erst nach erfolgreicher Freischaltung ist eine Antragstellung möglich. Die Antragstellung erfordert eine gründliche methodische und fachliche Vorbereitung. Methodisch wird geprüft, ob die technische Infrastruktur des Instituts mit den Anforderungen des FDZ-Gesundheit kompatibel ist und ob die notwendigen Kenntnisse in den unterstützten Programmiersprachen vorhanden sind. Fachlich muss sich das Institut intensiv mit seiner Forschungsthematik auseinandersetzen. Der Antrag muss wissenschaftlich fundiert und präzise formuliert sein. Er enthält unter anderem Angaben zu Forschungsziel, Zweck, Fragestellung, methodischem Vorgehen, Publikationsplan und erforderlicher Datentiefe (Ludwig, et al., 2024, S. 134). Der Antrag wird beim BfArM eingereicht, wo er zunächst auf Vollständigkeit geprüft und anschließend inhaltlich bewertet wird (Ihle, et al., 2022, S. 41). Bei Genehmigung wird der Antrag in ein Antragsregister aufgenommen (Datentransparenzverordnung, 2020, S. 7). In der nächsten Phase werden je nach Forschungsfrage Testdaten in einem virtuellen Analyseraum bereitgestellt. Dieser befindet sich in einer gesicherten VPN (Virtual Private Network) -Sandbox-Umgebung (Ludwig, et al., 2024, S. 134f.). Die Datenanalyse erfolgt über Skripte, die in Python, R oder SQL geschrieben werden können (Forschungsdatenzentrum Gesundheit, 2024c). Es besteht die Möglichkeit, Zwischenergebnisse anzufordern, das Skript zu testen und gegebenenfalls

anzupassen (Ludwig, et al., 2024, S. 135f.). Sobald die Datenergebnisse angefordert werden, wird das Skript auf den Originaldaten angewendet, um Ergebnistabellen zu generieren. Diese Tabellen werden auf das Risiko einer Re-Identifikation geprüft, bevor sie an die Antragstellenden weitergegeben werden (Ihle, et al., 2022, S. 42). Schließlich müssen die Ergebnisse innerhalb von 24 Monaten interpretiert, validiert und veröffentlicht werden. Darüber hinaus ist die Registrierung der Forschung in einem Primärregister obligatorisch (Henke & Schneider, 2024, S. 222).

6.3.2 Welche spezifischen Möglichkeiten und Limitationen ergeben sich daraus für die Forschung zu seltenen Erkrankungen bei Kindern?

Die Nutzung der im FDZ-Gesundheit bereitgestellten pseudonymisierten Daten von 73 Millionen gesetzlich Versicherten eröffnet bedeutende neue Möglichkeiten für die Forschung zu seltenen Erkrankungen bei Kindern (Ihle, et al., 2022, S. 39f.). Die hohe Anzahl an Fällen erlaubt die Durchführung fundierter statistischer Analysen, die bei Primärdatenerhebungen oft nicht möglich wären. Die Daten erlauben sowohl Längsschnitt- als auch räumliche Analysen und bieten detaillierte Informationen zu Mortalität, Prävalenz und Inzidenz von Krankheiten, zur Inanspruchnahme von Gesundheitsleistungen sowie zu Gesundheitsausgaben. Von besonderer Relevanz ist, dass die gesamte gesetzliche Versichertenpopulation einbezogen wird, wodurch auch schwer erreichbare Personengruppen, wie beispielsweise Kinder mit seltenen Erkrankungen, berücksichtigt werden können (Günster & Schmuker, 2024, S. 159). Die umfassende Datenbasis eröffnet die Möglichkeit, die Forschung zur frühen Erkennung, Verlaufskontrolle und Therapieoptimierung dieser Erkrankungen entscheidend voranzutreiben.

Obgleich diese Vorteile evident sind, sind dennoch wesentliche Limitationen zu konstatieren, welche die Forschung vor Herausforderungen stellen. Die Daten des FDZ-Gesundheit stammen primär aus Abrechnungsprozessen und sind administrative Routinedaten. In der Konsequenz sind in den Daten keine klinischen Informationen wie Laborwerte oder patientenspezifische Ergebnisse enthalten (Schubert, 2022, S. 8). Des Weiteren ist zu berücksichtigen, dass die Datenqualität und -vollständigkeit in Abhängigkeit von der jeweiligen Quelle und dem Zeitpunkt der Datenerfassung variieren (Günster & Schmuker, 2024, S. 159). Verzerrungen wie "Upcoding" oder die Anwendung von Techniken zur Anonymisierung, können die statistische Unsicherheit erhöhen und insbesondere bei geringen Fallzahlen, wie sie für seltene Erkrankungen typisch sind, gravierende Auswirkungen haben (Ihle, et al., 2022, S. 33; Schmitt, et al., 2023, S. 217). Ein weiteres Problem stellt die

zurückhaltende Zustimmung von Haushalten mit Kindern zur Speicherung in Registern dar (Haug, et al., 2024, S. 3). Dabei zeigen Patientinnen und Patienten mit seltenen Erkrankungen allgemein eine hohe Bereitschaft, ihre Daten für Forschungszwecke bereitzustellen (EURORDIS, 2022). Da die Mehrzahl seltener Erkrankungen bereits im Kindes- und Jugendalter manifestiert wird, können selbst geringe Widersprüche durch Nichtteilnahme die Aussagekraft der Forschungsergebnisse erheblich beeinträchtigen (Schmitt, et al., 2023, S. 217). Die Nichtberücksichtigung bestimmter Leistungsbereiche sowie anderer Sozialgesetzbücher, die nicht über die gesetzliche Krankenversicherung abgerechnet werden und mögliche Verzögerungen bei der Datenlieferung können die Generalisierbarkeit der Ergebnisse einschränken (March, et al., 2023, S. 147).
Zusammenfassend ermöglichen die FDZ-Gesundheitsdaten einen ressourcenschonenden Zugang zu einer umfangreichen Datenbasis, die neue Erkenntnisse zu seltenen Erkrankungen bei Kindern erlaubt. Gleichzeitig erfordern die genannten Limitationen eine sorgfältige Planung und Validierung der Analysen, um aussagekräftige und belastbare Ergebnisse zu generieren.

6.4 Stärken und Einschränkungen des Scoping Reviews

Die im Rahmen des Scoping Reviews untersuchten Fragestellungen wurden erfolgreich beantwortet und die damit verbundenen Ziele wurden erreicht. Der Ansatz des Scoping Reviews ermöglichte eine umfassende Betrachtung eines breiten Themengebiets unter Einbezug vielfältiger Perspektiven, deren Auswertung in einem zusammenfassenden Bericht dargestellt wurde. Dies ermöglichte die Identifikation der wesentlichen Elemente und die Generierung von Empfehlungen für die Forschung mit dem FDZ-Gesundheit-Datensatz, die sich durch eine hohe Anwendungsrelevanz und Realitätsnähe auszeichnen.

Die Durchführung eines solchen umfassenden Reviews wäre mit den üblichen systematischen Methoden kaum möglich gewesen. Nach aktuellem Forschungsstand stellt dieser Scoping-Review die erste umfassende Untersuchung des GDNG dar.

Gleichzeitig ist zu berücksichtigen, dass die vorliegende Arbeit einige Limitationen aufweist, die im Folgenden näher erläutert werden. Im Rahmen eines Scoping Reviews erfolgt keine Bewertung der Qualität der einbezogenen Artikel. Darüber hinaus werden Literaturquellen unabhängig von ihrem Design einbezogen, was die Validität der Ergebnisse beeinflusst. Es besteht die Möglichkeit, dass eine Ausweitung der Suche auf weitere Datenbanken zu einer größeren Anzahl passender Artikel führt und somit relevante Informationen übersehen wurden. Zudem mussten mehrere Artikel ausgeschlossen werden da kein Zugriff vorhanden war. Eine weitere

Einschränkung kann der Reader Bias darstellen, der besagt, dass die subjektive Wahrnehmung, das Interesse und die Vorurteile des Autors einen Einfluss auf die Auswahl der Artikel haben können.

7 Fazit

Die vorliegende Arbeit ging der Frage nach wie der Zugangs- und Nutzungsprozess gemäß des GDNG gestaltet ist und welche Möglichkeiten und Limitationen sich draus für die Forschung zu seltenen Erkrankungen bei Kindern gegeben. Für die Beantwortung wurde ein Scoping Review durchgeführt.

Das GDNG bildet eine wegweisende Grundlage für die Gesundheitsforschung, insbesondere für seltene Erkrankungen, und bringt zentrale Neuerungen wie die Zentralisierung und Standardisierung des Datenzugangs mit sich. Die Einrichtung des FDZ-Gesundheit und der DK-Stelle schafft transparente Plattformen und standardisierte Prozesse, die Forschenden den Zugang zu umfassenden Daten erleichtern. Besonders bedeutsam ist die Nutzung nahezu vollständiger Daten von 73 Millionen gesetzlich Versicherten, die verlässliche Schätzungen und tiefgehende Analysen ermöglichen. Der Zugang zu den Gesundheitsdaten erfolgt in Übereinstimmung mit den FAIR-Prinzipien. Die Beantragung von Daten aus verschiedenen Quellen sowie deren Analyse in einem speziell dafür vorgesehenen Raum werden zukünftig an einer einzigen Stelle möglich sein. Die Antragprozesse sind transparent und digital gestaltet. Die Daten werden in einem Format zur Verfügung gestellt, das eine Integration mit anderen Datensätzen ermöglicht. Dazu werden gemeinsame Standards verwendet. Anträge auf Datennutzung können jederzeit von Interessierten Institutionen zu bestimmten Zwecken angefragt werden. Dadurch sind die Daten reproduzierbar.

Trotz einiger aufgeführten Herausforderungen überwiegt der Nutzen für die Patientinnen und Patienten, der sich aus dem neuen Gesetz ergibt. Die Forschenden sind folglich dazu angehalten, im Vorfeld der Antragsstellung zu eruieren, ob die Daten für den jeweiligen Forschungszweck geeignet sind.

7.1 Auswirkungen aus gesundheitswissenschaftlicher Sicht

Das GDNG ist ein notwendiger Schritt, um die Entwicklung von Digital Health in Deutschland voranzutreiben. Es ermöglicht die Zusammenführung und

Nutzbarmachung von Daten der GKV für die Forschung und wird so eine Vielzahl neuer Studien anstoßen. Dadurch wird der Forschungsstandort Deutschland gestärkt.

Aus gesundheitswissenschaftlicher Perspektive hat das GDNG weitreichende Auswirkungen auf die Nutzung von Gesundheitsdaten. Es schafft einen rechtlichen Rahmen, der den Zugang zu Daten für Forschung, Versorgung und politische Entscheidungsfindung erleichtert und gleichzeitig Datenschutz und Privatsphäre der Bürgerinnen und Bürger schützt. Besonders wichtig ist die Förderung von Standardisierung und Interoperabilität, die zentrale Voraussetzungen für eine effektive Gesundheitsforschung schaffen – etwa bei seltenen Erkrankungen, bei denen größere Datenpools entscheidend sind. Zudem profitieren die Versorgungsforschung und die Qualitätssicherung, da das Gesetz fundierte Analysen zu Versorgungslücken und Behandlungsqualität ermöglicht. Ein weiterer Vorteil ist der Aufbau sicherer digitaler Infrastrukturen. Trotz dieser Fortschritte stehen Herausforderungen im Raum, darunter die praktische Umsetzung, die Komplexität des Datenzugangs und der Spagat zwischen breiter Verfügbarkeit und hoher Sicherheit. Ad-hoc Analysen in pandemischen Situation sind jedoch noch nicht möglich, dafür muss eine Priorisierung der Anträge festgelegt werden. Insgesamt ist das GDNG ein wichtiger Meilenstein auf dem Weg zu einer datengetriebenen Gesundheitsforschung, dessen Erfolg jedoch von einer konsequenten Umsetzung und der Akzeptanz aller Beteiligten abhängt.

In Bezug auf den sich im Aufbau befindlichen europäischen Raum für Gesundheitsdaten ist die Interoperabilität von entscheidender Bedeutung. Die Verknüpfung von Daten innerhalb der einzelnen Länder ist eine grundlegende Voraussetzung für die Schaffung einer internationalen Dateninfrastruktur.

7.2 Handlungsempfehlungen

Im Rahmen der Untersuchung der Antragsstellung und Datennutzung wurden einige Herausforderungen identifiziert, die es zu überwinden gilt, um valide Forschungsergebnisse mit den FDZ-Gesundheit-Daten zu erzielen. Antragsstellende sollten daher vor einer Antragsstellung einige Punkte berücksichtigen, um die Grundlage für eine erfolgreiche Forschung mit den FDZ-Gesundheit-Daten zu schaffen. Forschende müssen sorgfältig abwägen, ob die verfügbaren Datenbestände für ihre Forschungsfragen geeignet sind, was fundierte Kenntnis der Datenflüsse erfordert. Während Lücken in Leistungsdaten, wie beispielsweise Arzneimitteldaten, relativ einfach zu identifizieren sind, können bei anderen

Leistungsarten komplexere Verzerrungen auftreten. Für die Antragstellenden ist es entscheidend, ihre Forschungsfrage präzise zu formulieren. Andernfalls können zwei zentrale Probleme entstehen: Erstens kann der Antrag als nicht relevant eingestuft und abgelehnt werden. Weitaus problematischer ist jedoch, wenn der Antrag bewilligt wird, sich aber im Analyseprozess herausstellt, dass die Daten für eine fundierte Beantwortung der Forschungsfrage nicht ausreichen. In diesem Fall besteht das Risiko, Ergebnisse von geringer Qualität veröffentlichen zu müssen, was negative Auswirkungen auf die Außenwahrnehmung der Wissenschaft sowie die Reputation der Forschungsinstitute haben kann.

Ein weiteres Risiko bei der Arbeit mit den Daten des FDZ-Gesundheit besteht darin, dass die sogenannten verfügbaren Zeitscheiben im FDZ-Gesundheit Datensatz für einige Fragestellungen zu grob angelegt sind. Um dies zu vermeiden, sollte die Forschungsfrage an die Verfügbarkeit der Daten angepasst werden, wobei ein besonderer Fokus auf zeitliche Trends gelegt werden kann. Bereits vor Beginn der Sekundärdatenanalyse sollte berücksichtigt werden, dass die Aussagekraft der Ergebnisse durch Datenlücken eingeschränkt sein kann. Ebenso ist es wichtig, frühzeitig zu entscheiden, ob vorläufige Datenstände für die Analyse ausreichen, da diese während der Forschungsdauer im FDZ-Gesundheit korrigiert wird und die eigene Forschung damit obsolet ist.

Darüber hinaus stellen die Datenanalyse selbst und die damit verbundenen technischen Anforderungen weitere Herausforderungen dar. Zum einen erfordert die Nutzung der Daten spezifische Fähigkeiten, insbesondere in der Programmierung von Skripten. Hier können die Schulungsangebote des FDZ-Gesundheit hilfreich sein. Zum anderen kann eine Forschungskooperation mit einem externen Forschungsinstitut von Vorteil sein, insbesondere wenn die technischen Anforderungen, wie die vom BfArM bereitgestellten Arbeitsmittel, nicht mit der eigenen technischen Infrastruktur kompatibel sind, etwa bei Inkompatibilitäten zwischen VPN-Lösungen.

Eine nutzwertorientierte Gestaltung des Zugangs- und Nutzungsprozesses von Gesundheitsdaten erfordert verschiedene Ansätze, die über technische und methodische Optimierungen hinausgehen. Neben der Bereitstellung von Programmiersprachen sollten auch benutzerfreundliche, klickbasierte Lösungen und vordefinierte Standartauswertungen angeboten werden, um den Zugang für eine breitere Nutzerschaft zu erleichtern. Außerdem kann der Einsatz von künstlicher Intelligenz die schnelle Auswertung großer Datenmenge unterstützen. Die

Optimierung der Datenstruktur spielt ebenfalls eine zentrale Rolle. Eine stärkere Integration klinischer Parameter, beispielsweise durch die Verknüpfung von Gesundheitsdaten mit Krankenhausinformationssystemen, könnte die Aussagekraft und den praktischen Nutzen der Daten erheblich steigern.

Zur Abbildung von Versorgungs- und Krankheitsgeschehen ist ein sukzessiver Aufbau umfassender Datenbestände notwendig. Dabei sollten fall- oder personenbezogene Verknüpfungen innerhalb des FDZ-Gesundheit sowie mit externen Datenquellen von Anfang an berücksichtigt werden. Wichtige Datenquellen umfassen:

1. Pflegeversicherungsdaten,
2. Rentenversicherungsdaten
3. Unfallversicherungsdaten,
4. amtliche Statistiken,
5. Daten privater Kranken- und Pflegeversicherungen.

Durch das GDNG wird eine Grundlage für den Datenzugang zu Gesundheitsdaten geschaffen. Um jedoch besser auf pandemische Situationen wie in der Einleitung beschrieben reagieren zu können, müssen einige Regularien weiterentwickelt werden. Es sollte eine fast-line eingerichtet werden, in der die Prüfungsdauer verkürzt ist, um ad-hoc Analysen erstellen zu können. Gleichzeitig ist die Weiterentwicklung von Datenschutzmaßnahmen essenziell. Technologische Innovationen wie Differential Privacy bieten vielversprechende Ansätze, um den Schutz sensibler Daten zu gewährleisten und gleichzeitig deren Qualität für Forschungszwecke zu bewahren.

Um neue Kenntnisse zu formulieren, sollte weitere Forschung bezüglich des Zugangs- und Nutzungsprozess betrieben werden. Sobald der Datensatz des FDZ-Gesundheit für die Forschung nutzbar ist und erste Anträge bearbeitet wurden können die theoretischen Ansätze in der Praxis auf Machbarkeit geprüft werden. Zudem sollte das Scoping Review wiederholt und mit diesem abgeglichen werden. Zudem sollten dann Expertinnen- und Experteninterviews mit Antragsstellern in die Arbeit einbezogen werden.

8 Literaturverzeichnis

Arksey, H., & O'Malley, L. (2005). Scoping Studies: Towards a Methodological Framework. *International journal of social research methodology 8.1*, 19-32. doi:https://doi.org/10.1080/1364557032000119616

Bundesministerium für Bildung und Forschung. (o.J.). *Forschen am Seltenen*. Abgerufen am 18. Dezember 2024 von https://www.gesundheitsforschung-bmbf.de/de/forschen-am-seltenen-6439.php

Bundesministerium für Gesundheit. (2024a). *Seltene Erkrankungen*. Abgerufen am 18. Dezember 2024 von https://www.bundesgesundheitsministerium.de/themen/praevention/gesundheitsgefahren/seltene-erkrankungen/

Bundesministerium für Gesundheit. (2024b). *Verordnung zur näheren Regelung von Verfahren nach dem Gesetz zur verbesserten Nutzung von Gesundheitsdaten.* Abgerufen am 18. Dezember 2024 von https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/G/RefE_GDNGUmsV.pdf

Bundesministerium für Gesundheit. (o.J.). *Digitalisierung im Gesundheitswesen: Ein Überblick*. Abgerufen am 18. Dezember 2024 von https://gesund.bund.de/digitalisierung-im-gesundheitswesen#forschung

Busch, F. (2022). *Zu wenig Daten: Sachverständige kritisieren Corona-Politik.* Abgerufen am 18. Dezember 2024 von https://web.de/magazine/politik/daten-sachverstaendige-kritisieren-corona-politik-37069072

Dannenberg, P. A. (2024). Eine für alle - alle für eine? *Der freie Zahnarzt*, 14-29. Von https://doi.org/10.1007/s12614-024-1765-z abgerufen

Datentransparenzverordnung. (2020). *Verordnung zur Umsetzung der Vorschriften über die Datentransparenz.* Abgerufen am 18. Dezember 2024 von https://www.gesetze-im-internet.de/datrav_2020/DaTraV.pdf

Denis, A., Mergaert, L., Fostier, C., Cleemput, I., & Simoens, S. (2010). Issues surrounding orphan disease and orphan drug policies in Europe. *Appl Health Econ Health Policy 8*, 343–350. doi:https://doi.org/10.2165/11536990-000000000-00000

Deutsche Forschungsgemeinschaft. (2022). *Wissenschaften in der Coronavirus-Pandemie.* DFG. doi:10.5281/zenodo.7043356

DGP, & GPP. (2022). *LUNGENERKRANKUNG BEI MUKOVISZIDOSE: PSEUDOMONAS AERUGINOSA.* Abgerufen am 18. Dezember 2024 von https://register.awmf.org/de/leitlinien/detail/026-022

Digital Chiefs. (2022). *Deutschland hinkt bei Digitalisierung im Gesundheitswesen noch immer hinterher.* Abgerufen am 18. Dezember 2024 von https://www.digital-chiefs.de/deutschland-digitalisierung-gesundheitswesen/

Digital-Gesetz. (2024). *Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesen*. Abgerufen am 18. Dezember 2024 von https://www.recht.bund.de/bgbl/1/2024/101/VO.html

Europäisches Parlament und Rat. (1999). Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel und seltene Leiden. *Amtsblatt der Europäischen Gemeinschaft, L 18*, 1-5. Abgerufen am 18. Dezember 2024 von https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32000R0141

European Commision. (2014). *Implementation report on the commission communication on rare diseases: Europe's challenges.* Abgerufen am 18. Dezember 2024 von https://health.ec.europa.eu/document/download/d9c988a0-d295-47c8-a61b-9b6dd9a6460c_en

EURORDIS. (2022). *European Health Data Space: Unlocking the potential of health data for rare diseases*. Abgerufen am 18. Dezember 2024 von https://www.eurordis.org/european-health-data-space/

Fehr, A., & Prütz, F. (2023). Seltene Erkrankungen: Herausforderung für Medizin und Public Health. *J Health Monit 8(4)*, 3-7. Abgerufen am 18. Dezember 2024 von https://edoc.rki.de/bitstream/handle/176904/11420/JHealthMonit_2023_04_Editorial.pdf?sequence=1

Fischer, M., Jürgens, H., Mangelsdorf, S., Reif, S., Ullrich, H., & Wuppermann, A. (2023). Für einen besseren Datenzugang im Bereich Gesundheit. Stellungsnahme aus dem Verein für Socialpolitik. *RatSWD Working Paper 284/2023*. doi:https://doi.org/10.17620/02671.88

Forschungsdatenzentrum Gesundheit. (2024a). *Eine häufig gestellte Frage, die uns erreicht: Wofür werden die Daten am FDZ Gesundheit eigentlich genutzt?* LinkedIn. Abgerufen am 18. Dezember 2024 von https://www.linkedin.com/posts/forschungsdatenzentrum-gesundheit_wof%C3%BCr-darf-man-die-daten-am-fdz-gesundheit-activity-7173664097337495552-ic4V?utm_source=share&utm_medium=member_desktop

Forschungsdatenzentrum Gesundheit. (2024b). *Letzten Monat haben wir über die erlaubten Hashtag#Nutzungszwecke gesprochen - diesen Monat stellen wir die verbotenen vor!* LinkedIn. Abgerufen am 18. Dezember 2024 von https://www.linkedin.com/posts/forschungsdatenzentrum-gesundheit_die-verbotenen-nutzungszwecke-am-fdz-gesundheit-activity-7183795940292501507-vt4R?utm_source=share&utm_medium=member_desktop

Forschungsdatenzentrum Gesundheit. (2024c). *Was ist eigentlich der virtuelle Analyseraum? Und warum nutzen wir ihn?* LinkedIn. Abgerufen am 18. Dezember 2024 von https://www.linkedin.com/posts/forschungsdatenzentrum-gesundheit_der-virtuelle-analyseraum-activity-7216773171029438464-lvUt?utm_source=share&utm_medium=member_desktop

Gassner, U. M. (2022). Forschung und Innovation im europäischen Gesundheitsdatenraum. *Datenschutz Datensich 46*, 739-746. doi:https://doi.org/10.1007/s11623-022-1696-2

Gesundheitsdatennutzungsgesetz. (2024). *Gesetz zur verbesserten Nutzung von Gesundheitsdaten.* Abgerufen am 18. Dezember 2024 von https://www.recht.bund.de/bgbl/1/2024/102/VO.html

Günster, C., & Schmuker, C. (2024). Gesundheit und Klimawandel – welche Potenziale haben versorgungsnahe Daten? *Bundesgesundheitsblatt 67*, 155-163. doi:https://doi.org/10.1007/s00103-023-03828-8

Haug, S., Altenbuchner, A., & Weber, K. (2024). Elektronische Patientenakte und Freigabe von Gesundheitsdaten. *Paediatr. Paedolog*, 1-5. doi:https://doi.org/10.1007/s00608-024-01237-x

Heisig, J. P., Li, J., & Allmendinger, J. (2021). Gesundheitsdaten als öffentliches Gut. *Was ist Gesundheit*, 363-375. doi:https://doi.org/10.1515/9783110713336-025

Henke, V., & Schneider, H. (2024). Der European Health Data Space und das Gesundheitsdatennutzungsgesetz. In V. Henke, G. Hülsken, H. Schneider, & J. Varghese, *Health Data Management- Schlüsselfaktor für erfolgreiche Krankenhaus* (S. 207-224). Wiesbaden: SpringerGabler. Von https://doi.org/10.1007/978-3-658-43236-2_22 abgerufen

Hoffmann, G. F., & Hebestreit, H. (2022). Seltene Erkrankungen. *Monatsschr Kinderheilkd 170*, S. 10-12. doi:https://doi.org/10.1007/s00112-021-01371-x

Ihle, P., Schneider, K., & Heß, S. (2022). *Das Forschungsdatenzentrum Gesundheit - Routinedaten der Gesetzlichen Krankenversicherung für die Gesundheits- und Versorgungsforschung.* Berlin: BARMER Institut für Gesundheitsforschung. doi:10.30433/GWA2022-32

Klein, M. (2018). Braindrain. *Gabler Wirtschaftslexikon*. Springer Gabler. Abgerufen am 14. Januar 2025 von https://wirtschaftslexikon.gabler.de/definition/braindrain-30595/version-254172

Kole, A., & Faurisson, F. (2009). *The voice of 12,000 patients. experiences and expectations of rare disease patients on diagnosis and care in Europe.* EURORDIS-Rare Diseases Eu. Abgerufen am 18. Dezember 2024 von https://books.google.de/books?hl=de&lr=&id=13g4hrVlBnoC&oi=fnd&pg=PA323&dq=The+voice+of+12,000+patients.+experiences+and+expectations+of+rare+disease+patients+on+diagnosis+and+care+in+Europe&ots=sml1-qnCps&sig=2L1Q9wNn-BVBebcx89sM6M-5Qpg&redir_esc=y#v=one

Kostera, T., & Thranberend, T. (2018). #SmartHealthSystems: Digitalisierung braucht effektive Strategie, politische Führung und eine koordinierende nationale Institution. *Spotlight Gesundheit: Daten, Analyse, Perspektiven Nr. 5*. Abgerufen am 18. Dezember 2024 von https://rsm-bst-update.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/VV_SG_SHS_dt.pdf

Lampl, B., Edenharter, B., Leitzmann, M., & Salzberger, B. (2024). Analyse zur Datenqualität/-vollständigkeit bei COVID-19-Fällen: Warum eine digitale integrierte Datenerfassung auch zur Pandemiebekämpfung notwendig ist. *Das Gesundheitswesen, 86(06)*, 442-446. doi:10.1055/a-2234-6894

Ludwig, M., Schneider, K., Heß, S., & Broich, K. (2024). Aufbau des neuen "Forschungsdatenzentrums Gesundheit" zur Datenbereitstellung für die Wissenschaft. *Bundesgesundheitsbl 2024 · 67*, 131–138. doi:https://doi.org/10.1007/s00103-023-03831-z

March, S., Hoffmann, F., Andrich, S., Gothe, H., Icks, A., Meyer, I., . . . Ihle, P. (2023). Forschungsdatenzentrum Gesundheit – Vision für eine Weiterentwicklung aus Sicht der Forschung. *Das Gesundheitswesen 85.S 02*, 145-153. doi:10.1055/a-1999-7436

Moreno, A. (2024). Public Health hält Einzug in ein neues Gesundheitsdatennutzungsgesetz (GDNG). *Ethik Med 36*, 91-99. doi:https://doi.org/10.1007/s00481-023-00791-5

Ohlmeier, C., Frick, J., Prütz, F., Lampert, T., Ziese, T., Mikolajczk, R., & Garbe, E. (2014). Nutzungsmöglichkeiten von Routinedaten der Gesetzlichen Krankenversicherung in der Gesundheitsberichterstattung des Bundes. *Bundesgesundheitsbl. 57*, 464-472. doi:https://doi.org/10.1007/s00103-013-1912-1

Ortmann, O., Wesselmann, S., & Kowalski, C. (2024). Perspektiven für die registerbasierte Krebsforschung in DKG-zertifizierten Zentren. *Forum 39*, 196-200. doi:https://doi.org/10.1007/s12312-024-01321-0

Phillips & Cohen. (2018). *What is Upcoding and Unbundling Fraud? Healthcare Whistleblower Guide*. Abgerufen am 18. Dezember 2024 von Fighting forWhistleblowers for 30 Years: https://www.phillipsandcohen.com/upcoding-unbundling-fragmentation/

Prasser, F., Riedel, N., Wolter, S., & Ludwig, M. (2024). Künstliche Intelligenz und sichere Gesundheitsdatennutzung im Projekt KI-FDZ: Anonymisierung, Synthetisierung und sichere Verarbeitung für Real-World-Daten. *Bundesgesundheitsbl 67*, 171-179. doi:https://doi.org/10.1007/s00103-023-03823-z

Schmitt, J., Bierbaum, T., Geraedts, M., Gothe, H., Härter, M., Hoffmann, F., . . . Hoffmann, W. (2023). Das Gesundheitsdatennutzungsgesetz – Potenzial für eine bessere Forschung und Gesundheitsversorgung. *Gesundheitswesen, 85(04)*, 215-222. doi:10.1055/a-2050-0429

Schneider, K., & Hess, S. (2023). Forschungsdatenzentrum Gesundheit, European Health Data Space und DARWIN EU - neue Chancen für die Arzneimittelforschung mit Real World Data. *Bulletin zur Arzneimittelsicherheit*, 23-25. Abgerufen am 18. Dezember 2024 von https://www.pei.de/SharedDocs/Downloads/DE/newsroom/bulletin-arzneimittelsicherheit/2023/1-2023.pdf?__blob=publicationFile&v=5

Schubert, I. (2022). Arzneimittelepidemiologie und Arzneimittelverbrauchsforschung auf der Basis von Krankenkassendaten: zentrale Untersuchungsfragen und methodische Hinweise. *PrävGesundheitsf*. doi:https://doi.org/10.1007/s11553-022-00968-8

Schulte, C. (2024). Warum das Gesundheitsdatennutzungsgesetz der Datentransparenzstelle und damit der Forschung mit Routinedaten schadet.

In *Gesundheitswesen aktuell 2024* (S. 134-145). doi:10.30433/GWA2024-134

Slagmann, A., Hoffmann, F., Horenkamp-Sonntag, D., Swart, E., Vogt, V., & Herrmann, W. J. (2023). Analyse von Routinedaten in der Gesundheitsforschung: Validität, Generalisierbarkeit und Herausforderungen. *Z Allg Med 99*, 86-92. doi:https://doi.org/10.1007/s44266-022-00004-0

SPD, Bündnis 90/ Die Grünen, & FDP. (2021). *Mehr Fortschritt wagen: Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit.* Abgerufen am 18. Dezember 2024 von https://www.spd.de/koalitionsvertrag2021

Straub, N., & Kowalski, C. (2024). Gesundheitsdatennutzungsgesetz und Digital-Gesetz unter der Lupe. *Onkologie*, 735-743. doi:https://doi.org/10.1007/s00761-024-01546-7

Veit, K., Wessels, M., & Deiters, W. (2019). Gesundheitsdaten und Digitalisierung – Neue Anforderungen an den Umgang mit Daten im Gesundheitswesen. In M. A. Pfannstiel, P. Da-Cruz, & H. Mehlich, *Digitale Transformation von Dienstleistungen im Gesundheitswesen VI: Impulse für die Forschung* (S. 19-33). Wiesbaden: Springer Gabler. doi:https://doi.org/10.1007/978-3-658-25461-2_2

Von Elm, E., Schreiber, G., & Haupt, C. C. (2019). Methodische Anleitung für Scoping Reviews (JBI-Methodologie). *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen, 143*, 1-7. doi:https://doi.org/10.1016/j.zefq.2019.05.004

Weber, S., & Kuhn, J. (2023). Leitplanken statt Sperrbalken. *Bundesgesundheitsblatt 66*, 103-104. doi:https://doi.org/10.1007/s00103-022-03646-4

Weichert, T. (2024). Gesundheitsdatennutzung ohne Datenschutz? *Datenschutz Nachichten 2/2024*, 66-73. Abgerufen am 18. Dezember 2024 von https://www.datenschutzverein.de/wp-content/uploads/2024/06/DANA_24_2_Web.pdf

Wiarda, J.-M. (2021). Die Datenerhebungskatastrophe. *Wissenschaftspoltik*. Abgerufen am 18. Dezember 2024 von https://www.jmwiarda.de/2021/04/28/die-datenerhebungskatastrophe/

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data 3, 160018*, 1-9. doi:10.1038/sdata.2016.18

Witte, J., Zeitler, A., Hasemann, L., & Diekmannshemke, J. (2023). *DAK-Kinder- und Jugendreport 2023: Gesundheit und Gesundheitsversorgung während und nach der Pandemie.* Abgerufen am 18. Dezember 2024 von https://www.dak.de/dak/unternehmen/reporte-forschung/dak-kinder-und-jugendreport-2023_45524

9 Anhang

9.1 Anhang A: Suchstring

Suchstring für die Suche in GoogleScholar und LIVIVO (Gesundheitsdatennutzungsgesetz OR GDNG OR "Forschungsdatenzentrum Gesundheit" OR "FDZ-Gesundheit" OR "Versorgungsnahe Daten" OR Sekundärdatennutzung OR "Datenzugangs- und Koordinierungsstelle" OR "Forschungsdaten im Gesundheitswesen" OR (Datenzugang AND Gesundheit) OR Routinedaten)

9.2 Anhang B: Datenextraktion in Form einer Tabelle

Tabelle 2 Datenextraktion in Form einer Tabelle

Nr.	Autorinnen und Autoren, (Jahr)	Titel	Ziel des Artikels
1.	Henke & Schneider, (2024)	Der European Health Data Space und das Gesundheitsdatennutzungsgesetz.	Der Beitrag beschreibt die geplante Infrastruktur des Europäischen Datenraums für elektronische Gesundheitsdaten, die sogenannte "Umsetzungsbrücke" im deutschen Gesundheitssystem durch das GDNG und die daraus resultierenden Implikationen für Stakeholder und das Health Data Management.
2.	Dannenberg, (2024)	Eine für alle - alle für eine?	Der Artikel thematisiert die Implementierung der elektronischen Patientenakte für alle gesetzlich Versicherten in Deutschland ab dem Jahr 2025. Die Bereitstellung dieser Akte erfolgt automatisch durch die jeweiligen Krankenkassen und ihre Daten sind für verschiedene Zwecke, wie beispielsweise Forschung, Versorgungsverbesserung und die Entwicklung von Gesundheitssystemen, vorgesehen.
3.	Ludwig, et al., (2024)	Aufbau des neuen "Forschungsdatenzentrums Gesundheit" zur Datenbereitstellung für die Wissenschaft.	Der Beitrag analysiert die Vorzüge des FDZ-Gesundheit und dessen Potenzial für Forschung und Bevölkerung, erläutert die Struktur und zentrale Aspekte, präsentiert einen Ausblick auf die Chancen der Datenverknüpfung und illustriert die Antrags- und Datennutzungsprozesse anhand eines Beispiels zur Analyse von LongCOVID.
4.	Ihle, et al., (2022)	*Das Forschungsdatenzentrum Gesundheit - Routinedaten der Gesetzlichen Krankenversicherung für die Gesundheits- und Versorgungsforschung*	Der vorliegende Beitrag widmet sich der Erfolgsgeschichte der Routinedaten und deren Bereitstellung im FDZ-Gesundheit. Der historische Kontext bis zur Gründung des FDZ-Gesundheit wird beleuchtet, die Vorzüge und Limitationen von GKV-Routinedaten analysiert, die geplante Arbeitsstruktur des FDZ-Gesundheit erörtert und dessen Beteiligung an der

			Konzeption des europäischen Gesundheitsdatenraums diskutiert.
5.	Haug, et al., (2024)	Elektronische Patientenakte und Freigabe von Gesundheitsdaten.	Im Rahmen der vorliegenden Untersuchung wird die Nutzungsbereitschaft der elektronischen Patientenakte sowie die Einstellung zu Gesundheitsdaten und deren Übermittlung bei Personen mit und ohne Kinder analysiert.
6.	Weichert, (2024)	Gesundheitsdatennutzung ohne Datenschutz?	Der Artikel von Thilo Weichert äußert eine kritische Haltung gegenüber der geplanten Nutzung von Gesundheitsdaten im FDZ-Gesundheit und betont die Notwendigkeit eines umfassenden Datenschutzes. Weichert warnt vor den potenziellen Auswirkungen der neuen Regelungen auf den Datenschutz und das Patientengeheimnis.
7.	Straub & Kowalski, (2024)	Gesundheitsdatennutzungsgesetz und Digital-Gesetz unter der Lupe	Im Rahmen der vorliegenden Untersuchung wird das GDNG und das Digital-Gesetz hinsichtlich ihrer Bedeutung für die Forschung mit versorgungsnahen Daten analysiert. Zu diesem Zweck werden die Gesetze im Lichte bestehender Regelungen zur Zulässigkeit der Datenspeicherung und -verarbeitung analysiert und mit anderen Modellen der Datennutzung der Versorgungsforschung ins Verhältnis gesetzt.
8.	Günster & Schmuker, (2024)	Gesundheit und Klimawandel – welche Potenziale haben versorgungsnahe Daten?	In diesem Beitrag wird die Frage erörtert, welche potenziellen Auswirkungen der Klimawandel auf die Gesundheit haben könnte und inwiefern versorgungsnahe Daten zur Forschung in diesem Bereich beitragen können.
9.	Schneider & Hess, (2023)	Forschungsdatenzentrum Gesundheit, European Health Data Space und DARWIN EU - neue Chancen für die Arzneimittelforschung mit Real World Data.	Der Artikel zielt darauf ab, die bisher ungenutzten Potenziale von Gesundheitsdaten aus dem Versorgungskontext aufzuzeigen und zu erläutern, wie der Aufbau des FDZ-Gesundheit am BfArM und des European Health Data Space künftig einen sicheren und zeitnahen Zugang zu diesen Daten für die Forschung ermöglichen soll. Darüber hinaus werden neue Möglichkeiten für wissenschaftliche Analysen im Bereich der Arzneimittelregulation hervorgehoben.
10.	Slagmann, (2023)	Analyse von Routinedaten in der Gesundheitsforschung: Validität, Generalisierbarkeit und Herausforderungen.	In diesem Artikel wird dargelegt, dass Routinedaten aus der Gesundheitsversorgung für wissenschaftliche Analysen, insbesondere in der Allgemeinmedizin, eine signifikante Bedeutung haben. Darüber hinaus wird die Notwendigkeit einer kritischen Reflexion der Stärken und Schwächen von Routinedatenanalysen in der Gesundheitsforschung betont.
11.	EURORDIS, (2022)	*European Health Data Space: Unlocking the potential of health data for rare diseases*.	Der Artikel zielt darauf ab, die Relevanz des Austauschs von Gesundheitsdaten zur Förderung der wissenschaftlichen Forschung und

			zur Optimierung der klinischen Praxis für seltene Krankheiten zu betonen. Dabei wird betont, dass dieser Austausch von entscheidender Bedeutung ist, da Wissen und Expertise begrenzt sind, die Patientenpopulationen geografisch verstreut sind und ihre Gesundheitsdaten verstreut vorliegen.
12.	Moreno, (2024)	Public Health hält Einzug in ein neues Gesundheitsdatennutzungsgesetz (GDNG).	Der vorliegende Beitrag fokussiert sich auf den durch den Gesetzesentwurf des GDNG initiierten Diskurs über Nachhaltigkeit, Gemeinwohlorientierung und die Notwendigkeit einer begleitenden Verantwortungsethik bei der Nutzung von Gesundheitsdaten.
13.	Ortmann, et al., (2024)	Perspektiven für die registerbasierte Krebsforschung in DKG-zertifizierten Zentren	Der Artikel zielt darauf ab, die Relevanz der von der Deutschen Krebsgesellschaft zertifizierten Zentren und deren qualitativ hochwertigen Registerdaten für die Forschung zu betonen. Es wird dargelegt, dass sich die Daten für die Implementierung neuer Handlungsempfehlungen und die Beantwortung von Forschungsfragen, die nicht durch randomisierte kontrollierte Studien abgedeckt sind, eignen. Darüber hinaus wird die Notwendigkeit einer Vereinheitlichung der Tumordokumentationssysteme sowie die Rolle der klinischen Krebsregister der Länder für spezifische Forschungsprojekte betont.
14.	Schubert, (2022)	Arzneimittelepidemiologie und Arzneimittelverbrauchsforschung auf der Basis von Krankenkassendaten: zentrale Untersuchungsfragen und methodische Hinweise.	Seit den späten 1970er-Jahren werden Arzneimitteldaten der Krankenkassen genutzt, um die Entwicklung von Verbrauchsmengen und Kosten darzustellen. Der vorliegende Beitrag hat die Intention, das breite Spektrum an Fragestellungen in der Arzneimittelverbrauchsforschung und Pharmakoepidemiologie aufzuzeigen, die mithilfe von Krankenkassendaten untersucht wurden.
15.	Schulte, (2024)	Warum das Gesundheitsdatennutzungsgesetz der Datentransparenzstelle und damit der Forschung mit Routinedaten schadet.	Der Artikel zielt darauf ab, die negativen Auswirkungen des Gesundheitsdatennutzungsgesetzes auf die Datentransparenzstelle und die Forschung mit Routinedaten aufzuzeigen. Die Argumentation des Artikels fußt auf der Prämisse, dass die neuen Regelungen die Lieferung von Abrechnungsdaten unnötig verkomplizieren und zu weiteren Verzögerungen führen, was einen Rückschritt in der Nutzbarmachung dieser Daten darstellt.
16.	March, et al., (2023)	Forschungsdatenzentrum Gesundheit – Vision für eine Weiterentwicklung aus Sicht der Forschung	Aus der Perspektive der Forschung wurden zehn Thesen für eine Weiterentwicklung formuliert, die das Potenzial eines Forschungsdatenzentrums aufzeigen und Ideen für eine zukunftsfähige und

			beständige Ausgestaltung und Entwicklung präsentieren.
17.	Schmitt, et al., (2023)	Das Gesundheitsdatennutzungsgesetz – Potenzial für eine bessere Forschung und Gesundheitsversorgung	Der vorliegende Artikel verfolgt das Ziel, aus der Perspektive von Versorgungsforscher*innen die Voraussetzungen und Wünsche für eine optimale Gestaltung des Gesetzes zu formulieren. Das Positionspapier wurde vom Deutschen Netzwerk Versorgungsforschung und der Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten der Deutschen Gesellschaft für Sozialmedizin und Prävention sowie der Deutschen Gesellschaft für Epidemiologie erstellt. Die unterzeichnenden Fachgesellschaften unterstützen das Positionspapier, welches vor der Veröffentlichung des Referentenentwurfs zum GDNG verfasst wurde.

10 Eidesstattliche Erklärung

Ich erkläre hiermit, dass die vorliegende Arbeit selbständig und ausschließlich unter Benutzung der angegeben Hilfsmittel ohne Nutzung einer gKI-Anwendung (wie z.B. ChatGPT) angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche durch Anführungszeichen kenntlich gemacht und die Arbeit war in gleicher oder ähnlicher Fassung noch nicht Bestandteil einer Studien- oder Prüfungsleistung.

[redacted]

Hamburg, den 17.01.2025