



# Sekundärnutzung klinischer Daten aus der Patientenversorgung für Forschungszwecke – Eine qualitative Interviewstudie zu Nutzen- und Risikopotenzialen aus Sicht von Expertinnen und Experten für den deutschen Forschungskontext

Anja Köngeter, Martin Jungkunz, Eva C. Winkler, Christoph Schickhardt und Katja Mehlis

---

A. Köngeter (✉)

Nationales Centrum für Tumorerkrankungen (NCT), Sektion für Translationale Medizinethik, Universitätsklinikum Heidelberg, Heidelberg, Deutschland

E-Mail: [anja.koengeter@med.uni-heidelberg.de](mailto:anja.koengeter@med.uni-heidelberg.de)

M. Jungkunz

Nationales Centrum für Tumorerkrankungen (NCT), Sektion für Translationale Medizinethik, Universitätsklinikum Heidelberg, Heidelberg, Deutschland

E-Mail: [martin.jungkunz@med.uni-heidelberg.de](mailto:martin.jungkunz@med.uni-heidelberg.de)

E. C. Winkler

Nationales Centrum für Tumorerkrankungen (NCT), Sektion für Translationale Medizinethik, Universitätsklinikum Heidelberg, Heidelberg, Deutschland

E-Mail: [eva.winkler@med.uni-heidelberg.de](mailto:eva.winkler@med.uni-heidelberg.de)

C. Schickhardt

Nationales Centrum für Tumorerkrankungen (NCT), Sektion für Translationale Medizinethik, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Deutschland

E-Mail: [Christoph.Schickhardt@med.uni-heidelberg.de](mailto:Christoph.Schickhardt@med.uni-heidelberg.de)

K. Mehlis

Nationales Centrum für Tumorerkrankungen (NCT), Sektion für Translationale Medizinethik, Universitätsklinikum Heidelberg, Heidelberg, Deutschland

E-Mail: [katja.mehlis@med.uni-heidelberg.de](mailto:katja.mehlis@med.uni-heidelberg.de)

© Der/die Autor(en) 2022

G. Richter et al. (Hrsg.), *Datenreiche Medizin und das Problem der Einwilligung*,

[https://doi.org/10.1007/978-3-662-62987-1\\_10](https://doi.org/10.1007/978-3-662-62987-1_10)

## 1 Einleitung

Der Sekundärnutzung klinischer Daten zu Forschungszwecken wird großes Potenzial für verschiedene Arten von nicht-interventionellen, datengetriebenen Studien zugeschrieben. Unter der Sekundärnutzung klinischer Daten soll im Folgenden die Sammlung und Wiederverwendung von klinischen Versorgungsdaten für Forschungs- und Lernaktivitäten verstanden werden (siehe Jungkunz et al. in diesem Sammelband). Diese Sekundärnutzung beinhaltet keine zusätzlichen physischen Eingriffe oder Interventionen zur Datengenerierung. Eine solche Nutzungsart wird zu Forschungs- oder Lernzwecken durchgeführt, um das biomedizinische Wissen und somit die medizinische Versorgung zu verbessern; jedoch bietet diese Forschung üblicherweise keinen unmittelbaren Eigenutzen für die Patient\*innen, welche ihre klinischen Daten bereitstellen.

In Deutschland stellt die Sekundärnutzung klinischer Daten für Forschungs- und Lernaktivitäten (im Folgenden: Verwendung klinischer Daten zu Forschungszwecken) bisher die Ausnahme dar. Aktuell werden an der Schnittstelle der Datennutzung zwischen Forschung und Versorgung insbesondere ethische Problemstellungen im Spannungsfeld von Autonomie/informationeller Selbstbestimmung und Solidarität diskutiert (Deutscher Ethikrat 2017). Auch geltendes Recht, regulatorische Institutionen und technische Infrastrukturen scheinen den aktuellen Anforderungen einer datengetriebenen medizinischen Forschung nicht gerecht zu werden (Faden et al. 2013; Kass et al. 2013; Fiscella et al. 2015; Winkler 2017; Whicher et al. 2015; Schilsky et al. 2014). Für den nationalen Kontext kann entsprechend ein praktischer Handlungsdruck festgestellt werden, da Deutschland als Forschungsstandort „den Anschluss zu verlieren“ droht (Deutsche Hochschulmedizin e. V. 2019). Die vom Bundesministerium für Bildung und Forschung geförderte Medizininformatikinitiative (Semler et al. 2018) zielt deshalb darauf ab, die Daten aus der Krankenversorgung (zunächst mit Schwerpunkt auf den Universitätskliniken) für die Forschung nutzbar zu machen, indem sie die Voraussetzungen für eine standortübergreifende Verknüpfung von Versorgungsdaten schafft.

Um den nationalen ethischen, rechtlichen und regulatorischen Rahmen entsprechend anpassen zu können, gilt es Nutzen und Risiken der Verwendung klinischer Daten zu Forschungszwecken abzuschätzen. Tatsächlich mangelt es in dieser Hinsicht aktuell an empirischer Evidenz (Budrionis und Bellika 2016; Ford et al. 2019). Zwar wurden Nutzen und Risiken in projektbezogenen und teils hypothetischen Berichten und Studien für verschiedene nationale (The Nuffield Council on Bioethics 2015; Laurie et al. 2014; Myers et al. 2008; Arbeitsgruppe Personalisierte Medizin des BAG 2017) und internationale Kontexte (Thorogood 2020) erörtert und teilweise anhand konkreter (Vor-) Fälle empirisch belegt. In dieser Form zeigen für Deutschland jedoch lediglich das vom Bundesministerium für Bildung und Forschung geförderte Gutachten „Big Data im Gesundheitsbereich“ (Weichert 2018) und das vom Bundesministerium für Gesundheit geförderte Gutachten „,Datenspende‘ – Bedarf für die Forschung, ethische Bewertung, rechtliche, informationstechnologische und organisatorische Rahmenbedingungen“ (Strech et al. 2020) den möglichen Nutzen sowie individuelle, institutionelle und

gesellschaftliche Risiken von Big Data Anwendungen auf, welche auch einen Teil der untersuchten Sekundärnutzung einschließt. Das Gutachten bezieht zur Generierung dieser Einschätzungen keine weiteren Akteure systematisch ein.

Für eine fundierte und gesellschaftlich akzeptierte Abwägung von Nutzen und Risiken in diesem Handlungsfeld sind allerdings die *Expertise und die Erfahrungen relevanter nationaler Akteursgruppen* aus den Bereichen Forschung, Versorgung, Medizin-informatik, Patientenvertretung und Politik erforderlich, um anhand deren Handlungspraxis die oftmals auf Hypothesen und Einzelfällen beruhende Literatur zu bewerten und zu komplementieren. Nach unserer Kenntnis liegt allerdings *keine sozialempirische Studie* vor, welche die Perspektiven *verschiedener Akteure* systematisch und umfassend beleuchtet. Internationale Forschungsarbeiten informieren über relevante Themenfelder der Sekundärnutzung lediglich aus der Perspektive *einzelner Akteursgruppen*: In qualitativen und quantitativen Studien mit *Ärzt\*innen* wurden vor allem Risikopotenziale und Befürchtungen in Bezug auf die kommerzielle Datennutzung, die missbräuchliche Verwendung genetischer Daten und die Re-Identifikation beschrieben (Mayo et al. 2017; Perera et al. 2011; Ipos MORI 2016; Vezyridis und Timmons 2019). *Mitglieder von Ethikkommissionen* (Salerno et al. 2017) zeigten sich bei der Verwendung klinischer Daten für Forschungszwecke insbesondere besorgt im Hinblick auf die Datenintegrität, die Umsetzung der informierten Einwilligung und den Schutz der Privatsphäre von Patient\*innen. Zur Einstellung von *Patient\*innen/Bürger\*innen* bezüglich der Nutzung ihrer klinischen Daten für Forschungszwecke veröffentlichten Aitken et al. (2016) einen umfassenden Literaturüberblick. Kernthemen der darin untersuchten qualitativen Studien sind von Patient\*innen/Bürger\*innen wahrgenommene Befürchtungen wie der Verlust über die Kontrolle persönlicher Daten und möglicher Datenmissbrauch, sowie die Bedingungen für die öffentliche Akzeptanz einer Datennutzung, insbesondere in Hinblick auf den Einwilligungsprozess. Diese Überblicksarbeit weist auf ein geringes Bewusstsein der Bürger\*innen bezüglich der aktuellen Nutzungspraktiken dieser Daten für Forschungszwecke hin und betont die Relevanz des wahrgenommenen gesellschaftlichen Nutzens für die Patient\*innen/Bürger\*innen bei deren Abwägungsprozessen bezüglich einer Datenfreigabe. *Ländervergleichende Studien* verweisen auf nationale Unterschiede und eine *relativ skeptische Haltung* in der deutschen Bevölkerung gegenüber Datennutzung im Allgemeinen und im speziellen Kontext der Biobankenforschung (European Commission 2014, 2015; Hobbs et al. 2012; Gaskell et al. 2012). Lediglich *zwei sozialempirische Forschungsarbeiten mit Erhebungen in Deutschland* untersuchen die Sichtweisen von Akteuren in Bezug auf *klinische Daten* (Voigt et al. 2020; Richter et al. 2019). Zwar sind die Studien aufgrund unterschiedlicher Studiendesigns nur eingeschränkt vergleichbar, doch zeigen diese für den deutschen sozialempirischen Forschungskontext zusammenfassend ein *ambivalentes Bild* bezüglich der *Bereitschaft von Patient\*innen bzw. Bürger\*innen*, die Verwendung klinischer Daten für Forschungszwecke zu unterstützen. Mit Blick auf die *Abschätzung von Nutzen und Risiken* aus der Perspektive relevanter Akteursgruppen finden sich nach unserer Kenntnis *keine sozialempirischen Studien mit Erhebungen in Deutschland*.

Das **Ziel der vorliegenden Studie** ist es daher, diese ersten und zum Teil heterogenen Befunde durch einen Blick auf weitere *relevante Akteursgruppen* aus den Bereichen Forschung, Versorgung, Medizininformatik, Patientenvertretung und Politik durch eine sozialemprirische Analyse zu ergänzen. Auf diese Weise soll das *breite Spektrum an Nutzen- und Risikopotenzialen* abgebildet werden. Konkret bedeutet dies, die Wahrnehmungen und Erwartungen dieser Akteursgruppen hinsichtlich der Nutzung klinischer Daten aus der Patientenversorgung für Forschungszwecke in Deutschland in Erfahrung zu bringen. Den analytischen Rahmen bilden *vier Anwendungsfelder* für die Verwendung klinischer Daten: klinische (nicht-interventionelle) Forschung, Public Health Forschung, Forschung zur Qualitätsverbesserung und explorative Verwendung (siehe Jungkuntz et al. in diesem Sammelband). Durch die Analyse qualitativer Interviews beabsichtigen wir, dem Mangel an sozialempririschen Studien zur Abschätzung von Nutzen und Risiken in einem ersten Schritt relevante *Themenkomplexe* entgegenzusetzen, die von den Akteursgruppen geäußerte, *spezifische Bedarfe* im Kontext wahrgenommener Nutzen- und Risikopotenziale der Verwendung klinischer Daten zu Forschungszwecken widerspiegeln. Eine tiefergehende Diskussion dieser Punkte soll auf dringliche *Forschungsdiesiderate* hinweisen.

---

## 2 Methodisches Vorgehen

Im Zentrum des Forschungsinteresses standen das explizite Fachwissen der Expert\*innen, ihr implizites Erfahrungswissen sowie ihre Erwartungen gegenüber der Verwendung klinischer Daten zu Forschungszwecken. Deshalb wurde ein explorativer Zugang zum Forschungsgegenstand gewählt. Die Erhebungsmethode der qualitativen Interviews ist einerseits durch Offenheit gekennzeichnet, lässt unter der Verwendung eines Leitfadens aber auch eine gewisse Strukturierung der zu behandelnden Themenfelder zu (Flick 2017). So können die aktuelle Handlungspraxis anhand von Beschreibungen unterschiedlicher Akteure illustriert, neue Aspekte dieser Datennutzung in Erfahrung gebracht und übergeordnete Handlungsmuster identifiziert werden. Hierfür wurden *qualitative leitfadengestützte Interviews* (Bogner et al. 2014; Meuser und Nagel 2009) mit Expert\*innen aus Deutschland (n=20) und Österreich (n=1) geführt. Um den Forschungsgegenstand adäquat abzubilden und der komplexen Akteurslandschaft gerecht zu werden, wurden zunächst auf Basis nationaler Literatur (Weichert 2018; Thiel et al. 2018; Strech 2018; Deutscher Ethikrat 2017; Deutsche Hochschulmedizin e. V. 2019; Blachetta et al. 2016) relevante Akteure aus Forschung, Versorgung, Medizininformatik, regulierenden Institutionen (Ethikkommission), Patientenvertretung und Politik identifiziert. Um die Wahrnehmungen und Erwartungen der unterschiedlichen Stakeholdergruppen differenziert darzustellen, wurde als Sampling-Strategie ein Purposive Sample<sup>1</sup> gewählt (Flick 2017), welche eine gezielte Auswahl von Fällen erlaubt.

---

<sup>1</sup>In dieser Studie wurden „typische Fälle“ gewählt und auf eine maximale Variation dieser Fälle geachtet (Flick 2017).

**Tab. 1.** Samplestruktur: Zuordnung der interviewten Stakeholder zu Akteursgruppen

Erste Akteursgruppe (n = 13) <i>direkt beteiligt</i>	Zweite Akteursgruppe (n = 8) <i>mittelbar beteiligt und/oder betroffen</i>
Forschende Ärzt*innen/Forscher*innen (n = 5)	Nicht-forschende Ärzt*innen (n = 2)
Expert*innen für IT-Infrastrukturen (n = 8)	Vertreter*innen für Patienteninteressen (n = 2)
	Politikakteure (n = 2)
	Vertreter*innen des regulatorischen Tätigkeitsfelds (n = 1)
	Vertreter einer Institution des nationalen Gesundheitssystems (n = 1)

Die identifizierten Akteure lassen sich in zwei Gruppen unterteilen (siehe Tab. 1): Die *erste Akteursgruppe* ist aktuell oder perspektivisch an der Umsetzung der Verwendung klinischer Daten für Forschungszwecke in Deutschland *direkt beteiligt bzw. betroffen*. Im Mittelpunkt dieser ersten Akteursgruppe stehen forschende Ärzt\*innen/Forscher\*innen und Expert\*innen für IT-Infrastrukturen (n = 13). Das Sample deckt zudem alle vier vorgestellten Anwendungsfelder sowie verschiedene medizinische Fachgebiete (Onkologie, Kardiologie, Infektiologie, Allgemeinmedizin, Gerontologie, Epidemiologie) ab. Die beruflichen Positionen der Befragten sind sowohl auf der Leitungsebene von Instituten oder Abteilungen als auch auf der mittleren Hierarchieebene angesiedelt. Interviewpartner der *zweiten Akteursgruppe* sind an dieser Datennutzung aktuell oder perspektivisch *mittelbar beteiligte und/oder hiervon betroffene* Akteure. Aussagen dieser Gruppe sollen die Wahrnehmungen der ersten Akteursgruppe kontextualisieren und komplementieren. Diese zweite Akteursgruppe umfasst nicht-forschende Ärzt\*innen, Expert\*innen des regulatorischen Tätigkeitsfelds, Vertreter\*innen von Institutionen des Gesundheitssystems, Patientenvertreter\*innen sowie Politikakteure (n = 8). Die erste Erhebungswelle (n = 15) des Purposive Sampling orientierte sich an festgelegten Kriterien wie die Zuordnung zu Akteursgruppen, wissenschaftlichen Fachgebieten und Anwendungsfeldern. Nach einer ersten Auswertung der Interviews und Beratung in der Arbeitsgruppe kam im Rahmen eines Theoretical Sampling<sup>2</sup> (Flick 2017) eine zweite Erhebungsphase (n = 6) hinzu, die es erlaubte, sowohl die Komplementarität des Samples als auch die inhaltliche Sättigung der als besonders relevant eingestuften Themenkomplexe gezielt zu erhöhen.

Der Interviewleitfaden wurde basierend auf Literaturrecherchen und Diskussionen in der interdisziplinären Arbeitsgruppe entwickelt, welche Wissenschaftler\*innen aus den Disziplinen der empirischen Sozialwissenschaften, Medizinethik, Rechtswissenschaften und Medizin umfasst. Die im Interviewleitfaden erfassten Themenfelder speisten sich

<sup>2</sup>Bei dem theoretischen Sampling steht die Entwicklung einer Theorie auf Grundlage der erhobenen Daten im Vordergrund. Die Auswahl der Fälle und Untersuchungsgruppen erfolgt hierbei im Prozess der Datenauswertung (ibid.).

vornehmlich aus den vorgestellten internationalen sozialempirischen Studien und nationaler nicht-sozialempirischer Literatur.

Die Interviews fanden zwischen September und Dezember 2019 statt. Nach einer mündlichen und schriftlichen Aufklärung unterschrieben alle Interviewteilnehmer\*innen eine Einwilligungserklärung zur Teilnahme an der Studie. Die Interviews dauerten durchschnittlich eine Stunde und fünf Minuten. Alle Interviews wurden von AK geführt und digital aufgezeichnet, anschließend im Wortlaut nach festen Transkriptionsregeln verschriftlicht und unter Verwendung der Software MAXQDA 2020 für qualitative Datenanalyse ausgewertet. Personenbezogene Daten und Details, die Rückschlüsse auf die Identität von einzelnen Studienteilnehmenden erlaubten, wurden im Zuge der Transkription entfernt. Die Studie wurde von der Ethikkommission der medizinischen Fakultät der Universität Heidelberg berufsrechtlich beraten, ohne dass diesbezüglich Bedenken geäußert wurden.

Im Zentrum der *qualitativen Inhaltsanalyse* nach Mayring (2008) stand die Entwicklung eines deduktiv-induktiven Kategoriensystems (Kuckartz 2018): Das Kategoriensystem wurde auf Basis der in der Literatur dargestellten Themenkomplexe deduktiv angelegt und im Zuge der Auswertung des Interviewmaterials induktiv erweitert. Im Auswertungsprozess fand zuerst eine einzelfallbezogene Analyse statt, um detailliert die Einstellungen, Meinungen, Deutungsmuster und Erfahrungen in einem spezifischen (Anwendungs-) Kontext zu rekonstruieren. Anschließend wurden bei der fallübergreifenden Betrachtung die Kategorien verdichtet und unter zentralen, teilweise neuen, Kategorien angeordnet. Die Kodierungen der Interviews erfolgten durch AK und HK. Drei der Interviews wurden doppelt kodiert, um die Plausibilität des Kategoriensystems zu prüfen. Die Ergebnisse der Auswertung wurden mehrfach in der Arbeitsgruppe diskutiert (AK, HK, KM).

---

## 3 Ergebnisse

### 3.1 Handlungspraktiken der Sekundärdatennutzung im Spannungsfeld zwischen Forschung und Versorgung

Bei der Rekonstruktion der aktuellen Handlungspraxis anhand des Interviewmaterials wurde deutlich, dass die *Durchführung von datenintensiven multizentrischen Verbundprojekten* in Deutschland gegenwärtig *erschwert* ist. Hierzu äußerten sich hauptsächlich Interviewpartner\*innen der ersten Akteursgruppe, welche als Forscher\*innen/forschende Ärzt\*innen und als Expert\*innen für IT-Infrastrukturen direkt an der Umsetzung der Verwendung klinischer Daten für Forschungszwecke beteiligt sind. Die meisten wiesen in diesem Zusammenhang auf *hohe administrative Aufwände durch regulatorische Fragmentierung, föderale Strukturen sowie auf daraus resultierende rechtliche Unsicherheiten* bei der Initiierung von Studien hin<sup>3</sup>:

---

<sup>3</sup>Die im Folgenden angeführten Zitate entsprechen weitestgehend dem Wortlaut der Interviewten; sie sind lediglich zugunsten der besseren Lesbarkeit angepasst worden: Pausen, inhaltsleere Füllwörter, nonverbale Gesprächssignale und Betonungen werden hier nicht mit angegeben.

„[Diese Studien] umspannen auch immer mehr Ethikkommissionen. [...] Das dauert Jahre, und das ist nicht übertrieben, bis man bei einer Studie alle Zentren an Bord hat.“ (ID12, Experte IT-Infrastrukturen)

„[D]as ist das Gravierende, einen Flickenteppich. [...] Wenn Sie also übergreifende Verbundstrukturen einrichten [...], die länderübergreifend sind, haben die immer mehrere, am Ende bis zu 17 Datenschutzrechtsordnungen im Spiel. [...] Das macht es natürlich außerordentlich schwierig, weil das heißt, man muss für jedes Land eine eigene Rechtslage berücksichtigen.“ (ID11, Vertreter des regulatorischen Tätigkeitsfelds)

„Mit diesem Bescheid [der Ethikkommissionen] fällt es viel einfacher, wenn es oft gar nicht nötig wäre zu fragen [...]. Denn [die klinischen Partner] fühlen sich stets in einer Grauzonenfalle, ob die nun kooperieren dürfen mit den Daten oder nicht. Das heißt, lieber lassen sie die Daten ohne Nutzung liegen, zum Schaden des Patienten, anstatt selber noch in Haftungsfallen zu geraten.“ (ID17, Experte IT-Infrastrukturen)

Ein Grund für eine Behinderung – und teilweise Verunmöglichung – von Forschung mit bereits vorhandenen Daten aus der klinischen Routine sind laut der Expert\*innen häufig die aktuellen *Einwilligungspraktiken*. So müsse für die Nutzung von klinischen Daten üblicherweise eine Einwilligung für einen spezifischen Forschungszweck vorliegen bzw. eingeholt werden. Die wissenschaftliche Fragestellung und somit der Zweck könne zum Erhebungszeitpunkt jedoch häufig noch nicht spezifiziert werden, beispielsweise weil technologische Entwicklungen sich nur schwer vorhersagen lassen:

„Wenn ein Patient sagt, vor zehn Jahren, er gibt sein Tumorgewebe frei, [...] ist es dann auch in Ordnung damit Single-Cells-Sequencing zu machen, was es noch nicht gab? [...] So ähnlich denke ich auch, dass so ein Einverständnis [...] zukünftige Fragestellungen nicht komplett antizipieren kann.“ (ID1, Forschender Arzt/Forscher)

Vor diesem Hintergrund thematisieren einige Interviewpartner\*innen implizit und explizit das Verschwimmen der Grenzen von Forschung und Versorgung und die Schwierigkeit, das eigene Forschungsfeld zu verorten. Unsicherheiten regulatorischer und rechtlicher Art – und das Anliegen diese Unsicherheiten zu klären – scheinen in jenen Forschungsgebieten ausgeprägter zu sein, die sich dem „*Graubereich*“ zwischen *Forschung und Versorgung* zuordnen lassen. Dies trifft insbesondere auf die Anwendungsfelder der nicht-interventionellen klinischen Forschung und der Forschung zur Qualitätsverbesserung zu:

„Also wir machen ein molekulares Tumorboard und wenn [forschende Ärzte] ähnliche Patienten suchen... was ist mit denen passiert, wie war das Outcome, haben die noch ein zwei Jahre überlebt? Kann man da dies oder jedes Medikament nehmen? Ist das Versorgung? Ist das Forschung? Unsere Ärzte sagen, das ist Forschung. Ich würde sagen, aus meiner naiven Perspektive, das ist Versorgung.“ (ID2, Experte IT-Infrastruktur)

„Es ist eigentlich immer so, dass die Fragestellungen ganz klar aus der unmittelbaren Krankenversorgung kommen [...]. Natürlich gehen die dann unterschiedlich weit in die Forschung. [...] Bei uns ist das wirklich sehr eng verzahnt. Das ist dann wieder dieser Graubereich, wo ich mich immer wieder unsicher fühle.“ (ID18, Forschender Arzt/Forscher)

### 3.2 Erwartungen und Hoffnungen auf den Nutzen durch die Forschung mit klinischen Daten

#### *Forschungs- und Lernaktivitäten*

Die Expert\*innen der ersten Akteursgruppe fokussierten zumeist auf *unmittelbare Nutzenpotenziale für die eigenen Forschungsgebiete*. Sie erkannten hohe Nutzenpotenziale in allen vier Anwendungsfeldern, d. h. klinische nicht-interventionelle Forschung, Public Health Forschung, Forschung zur Qualitätsverbesserung, und explorative Verwendung. Zudem wiesen die Expert\*innen auf die Schnittmengen und Synergien zwischen den einzelnen Anwendungsfeldern bei der Nutzengenerierung hin. Es herrscht großes Interesse diesen ‚Datenschatz‘ zu heben. Einige Expert\*innen sprachen allerdings weniger von Erwartungen als von der ‚Hoffnung‘ auf den zukünftigen Nutzen, da es an entsprechender Evidenz aktuell noch fehle:

„Das Potenzial besteht in allen vier Bereichen, das ist ganz offensichtlich. Und da jetzt einen Schwerpunkt zu setzen macht auch keinen Sinn, weil die sich gegenseitig ergänzen und insofern wirklich nebeneinander betrieben werden müssen.“ (ID6, Vertreter für Patienteninteressen)

„Eine Verknüpfung von Gesundheitsdaten ist natürlich grundsätzlich eine riesige Ressource ist, ein riesiger Schatz für die Forschung.“ (ID12, Experte IT-Infrastrukturen)

„Man hat sehr viel Hoffnung. [...] Die Evidenz dazu ist ja oft noch relativ limitiert, weil es noch so neu ist.“ (ID19, Politikakteur)

Einigen Expert\*innen fiel es allerdings schwer die Nutzenpotenziale im Anwendungsfeld der *explorativen Verwendung* zu beurteilen. Häufig setzten die Interviewpartner\*innen diese Verwendungsart mit ‚Big Data Forschung‘ gleich, welcher trotz einer gewissen Skepsis gegenüber „Traumversprechungen“ durchaus Potenzial zugeschrieben wurde:

„Den Nutzen zu beurteilen, das kann ich nicht. Gerade bei diesen Big Data Konzepten, sogar die Fachleute tun sich da schwer. [...] Ob das jetzt so weit wie die Traumversprechungen von Big Data geht, kann ich nicht beurteilen, da bin ich auch eher skeptisch. Aber das muss ja gar nicht sein. Nutzen fängt ja nicht erst bei der Revolution an, sondern schon auf kleiner Flamme.“ (ID11, Vertreter des regulatorischen Tätigkeitsfelds)

Die Interviewpartner\*innen erkannten einen forschungsökonomischen Nutzen im Sinne der *Datensparsamkeit*. Einmal erhobene Daten können so effizient für zukünftige Forschungszwecke (wieder-)verwendet werden.

„Das Potenzial ist natürlich, dass man nicht nochmal Daten erheben muss. Das ist auch ein Nutzen für den Patienten. [...] Also sozusagen ein Gebot der Datensparsamkeit und der Ökonomie.“ (ID9, Forschender Arzt/Forscher)

Die Verwendung von „Real World Data“ ermögliche aus forschungsmethodischer Sicht neue Forschungsfragen zu formulieren, indem sie bisher häufig vernachlässigte



Subpopulationen wie beispielsweise Hochaltrige (>80 Jahre) einzubeziehen vermag, und so die *gesamte Population abbilden* könne:

„Potenziale sind natürlich, dass wir hier mit Real World Data das Krankheitsgeschehen und auch das Versorgungsgeschehen, so wie es in der tatsächlichen Praxis ist, abbilden können. Und nicht im Labor, an einer klinischen Studie sind, wo wir eine Unterrepräsentation der älteren Patienten haben.“ (ID9, Forschender Arzt/Forscher).

### *Gesundheitsversorgung*

In beiden Akteursgruppen wurde regelmäßig der erhoffte mittelbare Nutzen für eine verbesserte Gesundheitsversorgung durch einen *gezielteren und schnelleren Rückfluss der Forschungsergebnisse in den Versorgungskontext* von Forscher\*innen/forschenden Ärzt\*innen thematisiert, beispielsweise durch eine verbesserte Medikation bei Komorbiditäten oder durch neue Erkenntnisse, die der Qualitätsverbesserung in Krankenhäusern zu Gute kämen:

„Also bekommt jemand, der einen Herzinfarkt hatte oder eine chronische Lungen-erkrankung, das Chemotherapeutikum oder ist man da lieber schonender oder macht weniger. [...] Also Begleiterkrankungen, die die Behandlung entscheidend beeinflussen.“ (ID9, Forschender Arzt/Forscher)

„Sie haben eine Aussage zur Qualität der klinischen Versorgung, sie können natürlich sehr schön auch Vergleiche machen zwischen den einzelnen Krankenhäusern, den Leistungserbringern.“ (ID9, Forschender Arzt/Forscher)

Eine Patientenvertreterin gab an, dass Patient\*innen durch das Verfügbarmachen der eigenen klinischen Daten für die Forschung neben Hoffnungen für die eigene Gesundheit und die der Angehörigen auch eine Gelegenheit zu altruistischem und solidarischem Handeln sähen. Diese könne wiederum einen *positiven Einfluss auf den Krankheitsverlauf* ausüben und ein Zusammengehörigkeitsgefühl schaffen:

„Damit meine Erkrankung, meine Daten, meine Therapie anderen fremden Menschen oder der Entwicklung der Therapie helfen kann.“ (ID10, Vertreter für Patienteninteressen)

„[Der] Patient denkt: Vielleicht ist meine Krankheit für was gut [...]. Ich glaube das trägt viele: einfach in der Bewältigung der Krankheit.“ (ID10, Vertreter für Patienteninteressen)

## **3.3 Wahrgenommene Risikopotenziale und aktuelle Hindernisse der Sekundärdatennutzung**

### *Neuartige Bedarfe bei der Risiko-Abwägung*

Bei der Frage, wie schwerwiegend die Risiken bei der Verwendung klinischer Daten für Forschungszwecke insgesamt zu bewerten seien, divergierten die Ansichten der Interviewpartner\*innen stark: So erkannten einige Expert\*innen sehr große Risikopotenziale und andere schätzten diese als gering ein. Bei Letzteren gründete sich diese Ein-

schätzung häufig auf Abwägungsprozesse, bei denen der erwartete Nutzen einbezogen wurde. So wiesen im Zuge dieser Einschätzung insbesondere die Expert\*innen der direkt an Sekundärnutzung beteiligten Akteursgruppe darauf hin, dass die Risiken nicht allein zu betrachten seien: Der erwartete *Nutzen – sowohl für Forschung als auch für die Versorgung* – müsse stärker einbezogen werden. Für einen optimalen Abwägungsprozess müsse zudem das Fachwissen mehrerer Disziplinen, insbesondere von technischer Seite, berücksichtigt werden. Einige Interviewpartner\*innen kritisierten in diesem Zusammenhang an der aktuellen Praxis der Risikobewertung in Deutschland, dass Datenschützer\*innen die Risiken systematisch überbewerten und dieses Praxisfeld dominieren würden. Gleichzeitig wiesen einige Expert\*innen auf neue, bisher nicht bedachte Risikopotenziale bei der Verwendung klinischer Daten für Forschungszwecke hin:

„[Eine], aus meiner Sicht, viel zu konservative Interpretation dieses Gesetzes, die eigentlich eher von Datenschützern dominiert wird, als von den Anwendern. [...] Da das Gesetz – die Interpretation des Gesetzes – es immer nur darum geht, Risikominimierung zu betreiben, aber nicht Patientennutzen im Vordergrund stehen zu haben.“ (ID1, Forschender Arzt/ Forscher)

„Die [Risiken] einzuschätzen ist auch sehr schwierig, man braucht eigentlich auch technische Expertise.“ (ID11 Vertreter des regulatorischen Tätigkeitsfelds)

Die aktuellen Instrumente zur Risikofolgenabschätzung wurden hinterfragt; ebenso die erforderlichen professionellen Kompetenzen der mit einer ethischen Bewertung von Technologien beauftragten Gremien:

„[...] Wenn man sich die Besetzung dieser Ethikkommission ansieht, [...] sind das vielleicht nicht die richtigen Leute, [...] die die ethischen Aspekte dieser Technologie beurteilen können. Also weder vom Datenschutz her, noch von anderen ethischen Aspekten.“ (ID19 Politikakteur)

Die Risikoprofile der drei Anwendungsfelder nicht-interventionelle klinische Forschung, Public Health Forschung und Forschung zur Qualitätsverbesserung wurden nur marginal unterschiedlich betrachtet. Unsicherheit herrschte bezüglich dem Anwendungsfeld der *explorativen Verwendung*: Viele Expert\*innen drückten eine allgemeine Skepsis gegenüber eines möglichen Paradigmenwechsels hin zu nicht-hypothesengetriebener Forschung aus. Oftmals standen im Mittelpunkt der Überlegungen Applikationen des Maschinellen Lernens. Durch den aktuellen Mangel an Erklärbarkeit der Ergebnisse vermuteten einige Expert\*innen Risiken durch die Überinterpretation von Korrelationen nicht-hypothesengetriebener Forschung. Da deren Nutzen- und Risikopotenziale im Vorfeld momentan schwer abzuschätzen seien, gelte es für das neuartige Anwendungsfeld der explorativen Nutzung traditionelle Nutzen-Risiko-Abwägungen zu überdenken:

„Aus dem Bottom-Up ‚Datenklamüseln‘ kriegst du immer irgendwelche Korrelationen raus. Und vergisst auf Theorie und Kausalität zu schauen. Das macht was mit der medizinischen Forschung, das nicht in Ordnung ist.“ (ID19 Politikakteur)

„[W]enn da kein sicherer Nutzen ist, würde ich die Risikobewertung deutlich strenger machen. Das ist aber, glaube ich, etwas, wo wir alle gemeinsam interprofessionell die Regeln machen müssen, bevor es zu spät ist.“ (ID18 Forschender Arzt/Forscher)

### *Re-identifikation und Datenverlust*

Alle Experten\*innen stimmten darin überein, dass eine unerwünschte Re-identifizierung bei personenbezogenen Daten trotz Pseudonymisierung oder im Rahmen datenschutzrechtlicher Anstrengungen prinzipiell möglich sei, sofern dazu ein sehr hoher Aufwand betrieben wird (siehe Spitz & Cornelius in diesem Sammelband). Das Verständnis von Risiko war unter den Interviewpartnern allgemein divers. Ein Teil der Expert\*innen äußerte deutliche Skepsis aufgrund der erhöhten *Re-identifikationsrisiken einzelner Datentypen sowie gegenüber Datentransfer in Drittländer*, deren datenschutzrechtliche Standards nicht dem Niveau der EU-Länder entsprechen. Unterschiedliche Haltungen fanden sich insbesondere in Hinblick auf das Teilen sensibler Daten mit *privatwirtschaftlichen Akteuren*. Deshalb seien kooperative Modelle wünschenswert.

„Wir reden viel über anonymisierte Daten und da fehlt so ein bisschen die Ehrlichkeit, einfach zuzugeben, dass man medizinische Daten, insbesondere wenn man darauf medizinisch forschen will, nicht anonymisieren kann.“ (ID21, Vertreter von Institutionen des Gesundheitssystems)

„[W]enn zum Beispiel eine amerikanische Gruppe die Daten analysieren will, sollte das nicht über Clouds gehen, wo es kritische Zugriffe geben kann. Sondern die Datenanalyse sollte in einem System stattfinden, das aus europäischer Sicht adäquat ist.“ (ID1, Forschender Arzt/Forscher)

„Wenn es eine kommerzielle Auswertungsmöglichkeit gibt, dann wird die auch in Anspruch genommen. Also insofern denke ich ist die Skepsis, die in Deutschland und Europa sehr weit verbreitet ist, auch absolut berechtigt.“ (ID6, Vertreter für Patienteninteressen)

„Letztendlich braucht man kooperative Modelle mit den Firmen.“ (ID1, Forschender Arzt/Forscher)

Generell hielten sich die Expert\*innen mit Konkretisierungen und Eintrittswahrscheinlichkeiten möglicher Risiken zurück: In Anbetracht vieler Unsicherheiten könne man ebenso gut in „*Glaskugeln*“ schauen. Szenarien wurden als heuristisches Konzept genutzt, um theoretische Risikopotenziale zu illustrieren. Diese bezogen sich meist auf externe Risiken wie beispielsweise das eines ‚Hackerangriffs‘. Solche externen Risiken wurden regelmäßig als realistisch, aber wenig wahrscheinlich eingeschätzt:

„Ich halte das Risiko von Hackerangriffen auf Gesundheitsdaten eher für gering. Vielleicht mag ich falsch liegen, man guckt ja immer in Glaskugeln.“ (ID21, Vertreter von Institutionen des Gesundheitssystems)

Nur punktuell wurde auf interne Risiken wie Datenverlust durch technische Pannen, Datenlecks oder interne Zugriffe ohne Berechtigung eingegangen. Schilderungen konkreter Fälle entstammten dem Versorgungskontext:

„[D]iese Software hat eine Cloud-Lösung, vor allem die USA waren betroffen, dann ist das auf irgendeinem Server gelandet und ein Whistleblower hat offenbar davon erzählt [...]. Dann waren ein paar Millionen Datensätze einfach so, ohne Login, ohne irgendwas, ohne Schutz im Netz verfügbar.“ (ID2, Experte IT-Infrastrukturen)

### *Forschungsimmanente Risiken*

Direkt an der Umsetzung der Sekundärnutzung beteiligte Forscher\*innen und forschende Ärzt\*innen erkannten regelmäßig Risikopotenziale in der Verwendung von Daten mit womöglich nicht ausreichender Datenqualität und/oder einer nicht adäquaten Interpretation der Ergebnisse:

„[A]lso ich glaube, das ist noch ein weiter Weg diese klinischen Versorgungsdaten so qualitativ aufzuarbeiten, dass die wirklich für die Forschung nutzbar sind und keine Fehlschlüsse daraus gezogen werden dann.“ (ID12, Experte IT-Infrastrukturen)

Langfristig wurde zudem die Gefahr von *gesellschaftlichen Exklusionsmechanismen* durch benachteiligende oder diskriminatorische Effekte mittels datengetriebener medizinischer Forschung und Versorgung hervorgehoben.

„Ich glaube die Leute haben einfach Angst um den Schutz ihrer Privatsphäre. Dass sie sich preisgeben und daraus einen Nachteil bekommen. Man gebrandmarkt oder stigmatisiert wird, dass es doch offiziell wird, welche Krankheit man hat.“ (ID10, Vertreter für Patient\*inneninteressen)

[Dass] eben die Hautfarbe bei bestimmten Dingen eine Rolle spielen kann, zumindest bei einem Datensatz.“ (ID2, Experte IT-Infrastrukturen)

Eine forschende Ärztin betonte, dass durch die Verwendung klinischer Daten für Forschungszwecke den Patient\*innen keine Nachteile entstehen dürften: Sie illustrierte dies anhand des Beispiels einer mit multiresistenten Krankheitserregern infizierten Patientenpopulation, die aus diesem Grund mit einer geringeren Wahrscheinlichkeit in Rehakliniken aufgenommen worden sei. Nicht-hypothesengetriebene Forschung könne ihres Erachtens eine solche Benachteiligung vulnerabler Gruppen systematisch verstärken.

„Wir wissen jetzt schon, [...] dass Patienten, die einen multiresistenten Erreger haben, später und schlechter in eine Reha aufgenommen werden. [...] Deshalb denke ich, dass wir da Kollateralschäden zu befürchten haben, wahrscheinlich kennen wir noch gar nicht alle, die da auftreten [...]“ (ID18 Forschender Arzt/Forscher)

„Also gerade im Machine Learning bekommen Sie ja ‚böse Sachen‘ im Zweifelsfall raus. [...]

*Bereitschaft der Ärzt\*innen und Forscher\*innen klinische Daten zu sammeln und Daten zu teilen*

Eine grundsätzliche Befürchtung praktizierender Ärzt\*innen war ein möglicher *Mehraufwand bei der Dokumentation*, wobei ein Mehraufwand in der Akteursgruppe der nicht-forschenden Ärzt\*innen als nicht akzeptabel galt. Außerdem werden organisationsübergreifenden *Benchmarking Tools* sehr kritisch betrachtet, sofern diese als externe Qualitätskontrolle für den Vergleich von Ärzt\*innen oder Krankenhäusern eingesetzt würden:

„Ich muss auch dazu sagen, dass viele skeptisch sind – nicht weil sie Forschung nicht unterstützen würden – das hat einen ganz einfachen Grund: Weil das [Dokumentieren für Forschungszwecke] die Arbeit verändert, wie zum Beispiel eine andere Dateneingabe zu machen. [M]an darf nicht mehr das Programm verwenden, sondern muss das verwenden – damit sind dann Kosten verbunden.“ (ID14, nicht-forschender Arzt)

„Oder wir stochern in den Datenpools und können dann nachweisen, dass der Kollege, bei dem stirbt jeder fünfte Patient, der macht bestimmt schlechte Versorgung. Solche Dinge, das sind so die Ängste.“ (ID4, Experte IT-Infrastrukturen)

Ein Grundkonflikt bezog sich auf die fachspezifische *Kultur des Datenteilens*: So war die Sorge unter Forscher\*innen und forschenden Ärzt\*innen sehr groß, dass die Bereitschaft ihrer Kolleg\*innen zu gering sei ‚ihre‘ aus der klinischen Routine erhobenen und aufbereiteten Daten zu teilen. Diese Daten würden von Ärzt\*innen teilweise als „geistiges Eigentum“ betrachtet. So wurde insbesondere von forschenden Ärzt\*innen ein „Interessenausgleich“ gefordert für jene, die die Daten generieren, aufbereiten und anderen Forschern zur Verfügung stellen. Des Weiteren sei je nach Fachgebiet der Forscher\*innen/forschenden Ärzt\*innen die Befürchtung unterschiedlich stark ausgeprägt, dass andere Wissenschaftler\*innen die selbst erhobenen Daten falsch interpretieren oder den eigenen Forschungsarbeiten zuvorkommen könnten:

„Grundkonflikt ist ja immer der, dass diese Daten als Kapital angesehen werden [...]. Und wie man Interessenausgleich schafft zwischen den Personen, die direkt am Patienten arbeiten und diese Daten erheben, und denen, die einen wissenschaftlichen Benefit davon haben.“ (ID1, Forschender Arzt/Forscher)

„Patientendaten [werden] von vielen vielleicht noch als eine Art geistiges Eigentum betrachtet werden, was man nicht gerne teilt. Was man klassischerweise, als es noch auf Papierform war, auch aus Gründen des Berufsethos, schützen musste.“ (ID16, Politikakteur)

„[Die] Gefahr ist halt, dass wenn ich größere Datenpools habe, Fremde – und sei es nur der Kollege aus der Nachbarklinik – mit meinen Daten von meinen Patienten irgendeine Paper publiziert. Und wenn möglich dann noch falsche Aussagen trifft oder zu Dingen publiziert, wo ich eigentlich dazu forsche, ohne mich vorher zu fragen. [...] Das ist auch je nach Fachgebiet sehr, sehr unterschiedlich.“ (ID4, Experte IT-Infrastrukturen)

### *Einwilligungskontext und informationelle Selbstbestimmung der Patient\*innen*

Die meisten Expert\*innen gingen bei der Verwendung klinischer Daten für Forschungszwecke davon aus, dass Patient\*innen hierfür in eine prospektive Datennutzung einwilligen und dies im Kontext einer Behandlung in einer Klinik geschehe. Unter dieser Annahme befürchteten einige Interviewpartner\*innen einen *Verlust an informationeller Selbstbestimmung* der Patient\*innen. Grund hierfür sei beispielsweise ein nicht optimaler Einwilligungs- und Aufklärungsprozess im Versorgungskontext. Viele Expert\*innen, insbesondere praktizierende Ärzt\*innen und eine Vertreterin für Patient\*inneninteressen, verwiesen auf ihre Erfahrungen aus Studienkontexten. So seien Patient\*innen nicht immer in der Lage informiert einzuwilligen, beispielsweise aufgrund von Disstress ausgelöst durch eine schwere Diagnose. Auch eine gefühlte „Zwangslage“ im Arzt-Patienten-Verhältnis und damit zusammenhängende falsche Hoffnungen auf einen persönlichen Nutzen könne die Entscheidung beeinflussen:

„Patienten-...die sind eh gebrannt und traumatisiert von ihrer Geschichte [...]. Wenn man denen dann einen Zettel gibt, wo draufsteht ‚ich gebe meine Daten frei‘. Ich glaube, die unterschreiben erst mal einfach alles, ohne dass sie wirklich aufgeklärt sind.“ (ID10, Vertreter für Patient\*inneninteressen)

„Wenn ich krank bin, habe ich ein berechtigtes Interesse daran, dass mir geholfen wird [...]. Jetzt könnte man fragen, ist das dann eine Zwangslage, in der ich mich befinde?“ (ID4, Experte IT-Infrastrukturen)

### *Risikopotenziale auf gesellschaftlicher Ebene*

Viele Expert\*innen nannten Risikopotenziale auf gesellschaftlicher Ebene. So schien eine latente Befürchtung zu sein, dass die *gesellschaftliche Akzeptanz* für die angestrebte Datennutzung nicht vorhanden sei und man es nicht schaffe, die Bevölkerung mit „ins Boot zu holen“. Wiederholt wurde zudem die *Angst vor dem* „gläsernen Menschen“ angeführt. Expert\*innen machten in diesem Zusammenhang auf den Mangel um konkretes Wissen über Wünsche und Bedürfnisse der Patient\*innen aufmerksam. Dieses Wissen könne unter anderem die Konzeption von Informationsmaterialien im Einwilligungskontext unterstützen.

„Risiken sehe ich insbesondere in der öffentlichen Akzeptanz. [...] Und dafür bedürfte es eben auch eines öffentlichen Diskurses, der nicht von oben runter stattfindet, sondern auf einer horizontalen Ebene abläuft.“ (ID6, Vertreter für Patient\*inneninteressen)

„Der ‚Gläserne Mensch‘. Das ist, glaube ich, die große Sorge.“ (ID10, Vertreter für Patient\*inneninteressen)

„Also grundsätzlich habe ich ganz große Bedenken, dass man es nicht schafft, die Bevölkerung so ins Boot zu holen, dass Benefit und Risiken vernünftig abgewogen werden können, auch von dem Einzelnen.“ (ID12, Experte IT-Infrastrukturen)

„Also das ist sicherlich bei der Konzeption von Projekten [...] oder von Einwilligungsformularen eine große Hürde, an einen Patienten und dessen Wünsche und Bedürfnisse ranzukommen.“ (ID16, Politikakteur)

Laut Vertreter\*innen für Patienteninteressen könnte das *Vertrauen der Bevölkerung in Institutionen des Gesundheitssystems* im Falle eines Datenverlusts oder eines Datenmissbrauchs erodieren. Laut Expert\*innen aus unterschiedlichen Akteursgruppen seien höchstmögliche Sicherheitsstandards der Dateninfrastruktur eine Mindestvoraussetzung für gesellschaftliche Akzeptanz und Vertrauen.

„[Auf] der anderen Seite würde [das Bekanntwerden eines Datenmissbrauchs] wahrscheinlich ganz massive Empörungen in der Öffentlichkeit auslösen.“ (ID6, Vertreter für Patient\*inneninteressen)

„Und sehe aber auch wirklich die Notwendigkeit, dass die Leute, die dann wirklich mit den Daten und dem Datenmanagement zu tun haben, auch eine vernünftige Struktur aufbauen, dass es überhaupt keine Ansatzpunkte gibt, dass man da Ängste aufbauen müsste.“ (ID12, Experte IT-Infrastrukturen)

---

## 4 Diskussion und Ausblick

Das Ziel der vorliegenden empirischen Studie ist es, erstmalig die Wahrnehmungen und Erwartungen unterschiedlicher Akteursgruppen im deutschen Forschungskontext im Hinblick auf Nutzen- und Risikopotenziale der Verwendung klinischer Versorgungsdaten für Forschungszwecke darzustellen. Durch die Analyse qualitativer Interviews beabsichtigen wir, dem Mangel an sozialem empirischen Studien in einem ersten Schritt Themenkomplexe zur Abschätzung von Nutzen und Risiken aus der Perspektive relevanter Akteursgruppen entgegenzusetzen. Insgesamt *sechs Themenkomplexe* bündeln thematisch die von den Akteursgruppen geäußerten spezifischen Bedarfe im Kontext wahrgenommener Nutzen- und Risikopotenziale der Verwendung klinischer Daten für Forschungszwecke: i) Regulatorische und rechtliche Hindernisse im Kontext der Sekundärnutzung, ii) fachliche und gesellschaftliche Nutzenpotenziale durch wissenschaftlichen Fortschritt und bessere Gesundheitsversorgung, iii) Risikopotenziale und Herausforderungen in Hinblick auf Datenqualität und -verarbeitung sowie diskriminatorischer Verwendung, iv) Implikationen für Praktiken der Risiko- und Nutzenabwägung, v) Informationelle Selbstbestimmung und Konditionen des Einwilligungsprozesses und vi) die Relevanz von gesellschaftlicher Akzeptanz und Vertrauen in Institutionen des Gesundheitswesens. Innerhalb dieser Themenkomplexe ermöglichte das methodische Vorgehen auf Basis explorativer Interviews mit Expert\*innen bisher nicht beforschte Aspekte der Datennutzung zu identifizieren. Durch die *erstmalige systematische Analyse von vier Anwendungsfeldern* (i.e. klinische [nicht-interventionelle] Forschung, Public Health Forschung, Forschung zur Qualitätsverbesserung und explorative Verwendung) für die Verwendung klinischer Daten konnten Ähnlichkeiten und Divergenzen der Nutzen- und Risikoprofile aufgezeigt werden.

Die Diskussion setzt die vorliegenden Ergebnisse aus den Interviews mit Expert\*innen in den Kontext bereits vorhandener Literatur und weist auf *Forschungsdesiderate* hin. So komplementieren diese Themenkomplexe die Gegenstandsbereiche der zentralen nicht-sozialem empirischen Literatur (Weichert 2018; Laurie et al. 2014;

Myers et al. 2008; Arbeitsgruppe Personalisierte Medizin des BAG 2017; Thorogood 2020). Sie vermögen zudem zwei sozioempirische Studien bezüglich der *Bereitschaft von deutschen Patient\*innen* (Voigt et al. 2020; Richter et al. 2019) durch einen Blick auf weitere Akteursgruppen zu kontextualisieren.

#### **4.1 Regulatorische Fragmentierung, föderale Strukturen und rechtliche Unsicherheiten als Hindernis: Sekundärnutzung im Kontext einer stärkeren Integration von Forschung und Versorgung**

Die Expert\*innen sahen insbesondere Handlungsbedarf in den Sphären des geltenden Rechts, regulatorischer Institutionen und technischer Infrastrukturen. Dabei wurde als gemeinsames Deutungsmuster identifiziert, dass in Deutschland gegenwärtig hohe administrative Aufwände die Durchführung von datenintensiven multizentrischen Verbundprojekten erschweren. Grund hierfür seien insbesondere *regulatorische Fragmentierung, föderale Strukturen sowie daraus resultierende rechtliche Unsicherheiten* (vgl. S. 190). Hierbei verunsicherte die Forschenden häufig die als unklar eingestufte Rechtslage (siehe Schrader sowie Spitz & Cornelius in diesem Sammelband). Außerdem beanstandeten sie die aktuelle *Praxis der Risikobewertung* in Deutschland, welche laut der Expert\*innen von Datenschützer\*innen dominiert sei, die *Risiken systematisch und unverhältnismäßig überbewerten* würden. Diese in der vorliegenden Studie von den Expert\*innen beschriebene Handlungspraxis untermauert den in der Literatur angezeigten praktischen Handlungsdruck bei der Verwendung von klinischen Daten für Forschungszwecke in Deutschland (Faden et al. 2013; Kass et al. 2013; Fiscella et al. 2015; Winkler 2017; Whicher et al. 2015; Schilsky et al. 2014; Deutscher Ethikrat 2017; Deutsche Hochschulmedizin e. V. 2019).

Fallspezifische Analysen der vorliegenden Interviews legen nahe, dass mit Blick auf *regulatorische Unsicherheiten insbesondere Forschungsgebiete mit starkem Bezug zum Versorgungskontext* betroffen sind. Konkret tritt diese Beobachtung verstärkt in den Anwendungsfeldern der nicht-interventionellen klinischen Forschung und der Forschung zur Qualitätsverbesserung zu Tage (vgl. S. 191). Betroffene Expert\*innen forderten in diesem Zusammenhang Klärung, wie mit der kategorialen und regulatorischen Unterscheidung von Forschung und Versorgung angesichts in der Praxis existierender Graubereiche zukünftig umgegangen werden soll. Diese Ergebnisse korrespondieren mit der qualitativen Studie von Whicher et al. (2015), die einen Handlungsdruck im Bereich der Forschung zu Qualitätsverbesserung und vergleichender Wirksamkeitsstudien empirisch untermauert. *Zukünftige Studien* sollten den aktuellen Stand der fachspezifischen Integration von Forschung und Versorgung in Deutschland wissenschaftlich eruieren, um auf dieser Basis weiterführende Diskussionen in Hinblick auf aktuelle regulatorische und rechtliche Strukturen anzuregen.



## 4.2 Ausschöpfung fachlicher und gesellschaftlicher Nutzenpotenziale als legitimierender Faktor

Alle Akteursgruppen erwarteten oder hofften auf einen mittelbaren Nutzen für Patient\*innen durch eine *verbesserte Gesundheitsversorgung* aufgrund eines gezielteren und schnelleren Rückflusses der Forschungsergebnisse in die Versorgung sowie gänzlich neue wissenschaftliche Fragestellungen (vgl. S. 192 f.). Direkt an der Verwendung klinischer Daten beteiligte Expert\*innen nannten abhängig von ihrem Forschungsgebiet konkrete Erwartungen oder vage Hoffnungen auf die unmittelbaren *Nutzenpotenziale für die eigenen Forschungsgebiete*. Sie erkannten zudem forschungsökonomische Nutzenpotenziale bei dieser Datennutzung, sofern diese im Sinne der *Datensparsamkeit* effizient und effektiv von der Forschung genutzt werde. Einen forschungsmethodischen Aspekt sahen einige Interviewpartner\*innen unterschiedlicher Akteursgruppen in der Verwendung von „Real World Data“: Das Einbeziehen der gesamten Patientenpopulation weckte die Hoffnung auf eine *bessere Versorgung selten untersuchter Studienpopulationen*, wie Hochaltrige und Ko-Morbide. Dieser Aspekt kann nicht nur ein für die Gesundheitsversorgung großes Potenzial bergen, sondern auch gesellschaftlich legitimierend wirken (Aitken et al. 2018).

*Nutzenpotenziale* wurden in der sozialempririschen Forschung nach unserer Kenntnis bislang *nicht als zentraler Gegenstandsbereich systematisch unter Einbeziehung verschiedener Akteursgruppen* untersucht. Um die genannten forschungsökonomischen und forschungsmethodischen Nutzenpotenziale gezielt zu fördern, sollten Studien in einem ersten Schritt die direkten und indirekten Wirkungen der Verwendung klinischer Daten für Forschungszwecke abschätzen. Insbesondere die Wirkung durch das Einbeziehen selten untersuchter Studienpopulationen kann einen wertvollen Beitrag zur Verringerung gesundheitlicher und sozio-ökonomischer Ungleichheit leisten.

## 4.3 Risikopotenziale und Herausforderungen in Hinblick auf Datenqualität und -verarbeitung sowie diskriminatorische Verwendung

Unter Expert\*innen galt die theoretische Möglichkeit mit *personenbezogenen Daten Individuen zu re-identifizieren als unstrittig* (vgl. S. 195). Generell hielten sich die Expert\*innen mit Konkretisierungen und Eintrittswahrscheinlichkeiten möglicher Risiken zurück und verwendeten Szenarien als heuristisches Konzept, um theoretische Risikopotenziale zu illustrieren. Diese konnten unterteilt werden in *externe Risiken* wie ‚Hackerangriffe‘, die zwar als realisierbar, aber wenig wahrscheinlich eingeschätzt wurden. Außerdem wurden *interne Risiken* wie Datenverluste durch technische Pannen genannt, wobei sich konkrete Fälle auf den Versorgungskontext bezogen. Die geschilderten Risikopotenziale bildeten das Themenspektrum sozialempririscher Studien ab, die mit einzelnen Akteursgruppen, i.e. Ärzt\*innen, Mitgliedern von

Ethikkommissionen und Patient\*innen/Bürger\*innen, durchgeführt wurden (Aitken et al. 2016; Mayo et al. 2017; Perera et al. 2011; Ipos MORI 2016; Vezyridis und Timmons 2019; Salerno et al. 2017). Als besonders relevanter und sozialempririsch bisher kaum untersuchter Aspekt stach bei der Analyse der Interviews hervor, dass insbesondere Forscher\*innen/forschende Ärzt\*innen und Expert\*innen für IT-Infrastrukturen große Bedenken in Hinblick auf eine *nicht ausreichende Datenqualität* zeigten. Diese könne falsche Forschungsergebnisse oder Fehlinterpretationen nach sich ziehen (vgl. S. 196). Die Bedenken um eine nicht ausreichende Datenqualität wurden besonders häufig in Bezug auf das Anwendungsfeld der explorativen Verwendung geäußert, insbesondere im Rahmen von Applikationen des Maschinellen Lernens und (oft synonym verwendet) ‚Big Data‘ Anwendungen.

In der sozialempririschen Literatur wurde der Aspekt einer nicht ausreichenden Datenqualität und falscher Forschungsergebnisse bisher lediglich in Salerno et al. (2017) angeführt, welche die Risiken von ‚Big Data‘ auf Grundlage von Zusammenfassungen und verschriftlichen Diskussionen eines epidemiologischen Kongresses analysierten. Die internationalen Teilnehmer\*innen äußerten ebenfalls die Befürchtung einer nicht ausreichenden Datenintegrität und daraus resultierend falscher Studienergebnisse. Die Autor\*innen stellten mögliche Lösungsansätze vor, wie beispielsweise die Einführung klarer ethischer Rahmenbedingungen und kostenlose digitale Lehrangebote, die eine ethische Auseinandersetzung mit datengetriebener Forschung ermöglichen.

Nicht-forschende Ärzt\*innen befürchteten einen Mehraufwand durch zusätzliche Dokumentationsstätigkeiten um den Ansprüchen an die Datenqualität bei der Verwendung für Forschungszwecke gerecht zu werden (vgl. S. 197). Um diese Bedenken zu adressieren, können transdisziplinäre und sozio-technologische Ansätze zielführend sein, die Ärzt\*innen (und gegebenenfalls Pflegekräfte) einbinden, um gemeinschaftlich Lösungsstrategien zu entwickeln.

Ein weiterer Aspekt ist aus Sicht einiger Expert\*innen die Gefahr *diskriminatorischer Effekte* dieser Datennutzung. So könnten vulnerable Subpopulationen beispielsweise mit Vorerkrankungen oder mit Stigmata behafteten Krankheitsbildern durch die angestrebte Datennutzung weitere Nachteile erfahren (vgl. S. 196). Auch aus den Empfehlungen der Stellungnahme der Datenethikkommission (2018) und dem Gutachten von Weichert (2018) für das Bundesministerium für Bildung und Forschung geht für die allgemeine, also nicht spezifisch auf Sekundärnutzung bezogene, Nutzung von Applikationen des Maschinellen Lernens hervor, dass die Gefahr einer Benachteiligung einzelner Bürger\*innen oder gar Subpopulationen zu berücksichtigen und zu verhindern sei. Insbesondere bei nicht ausreichender Datenqualität warnt die Stellungnahme der Datenethikkommission vor systematischen Diskriminierungen und pluralitätsfeindlichen Ergebnissen. Da Algorithmen bereits bestehende Diskriminierungen von Subpopulationen verstärken könnten, sei eine Verstärkung gesundheitlicher Ungleichheit und die *Entwicklung gesellschaftlicher Exklusionsmechanismen* möglich.

Zukünftige Forschungsvorhaben sollten auch für die Verwendung klinischer Daten für Forschungszwecke neuartige Risikopotenziale aufgrund nicht ausreichender Datenqualität oder diskriminatorischer Effekte untersuchen und kontinuierlich bewerten. Gleichzeitig sollten neuartige Herausforderungen einer möglicherweise nicht ausreichenden Datenqualität klinischer Daten frühzeitig in Fachkreisen diskutiert werden um das Vertrauen in die Datenbasis zu gewährleisten.

#### **4.4 Bedarf an neuartigen Praktiken der Nutzen-Risiko-Abwägung**

Die von Expert\*innen geäußerten Bedenken weisen darauf hin, dass die aktuellen Praktiken der Nutzen- und Risikoabwägung zu überdenken sind: So lassen die beschriebenen regulatorischen Hürden multizentrischer Studien vermuten, dass aktuelle Praktiken der *Risikobewertung den Erfordernissen bei der Verwendung klinischer Daten für Forschungszwecke nicht gerecht werden (vgl. S. 194 f.)*. Zudem ließen Bedenken in Bezug auf das Anwendungsfeld der explorativen Verwendung auf *Unterschiede in den Risikoprofilen der Anwendungsfelder* schließen. Laut einiger Expert\*innen könnte zukünftig der *Nutzen höher bewertet werden* aufgrund einer vermuteten gezielteren und schnelleren Translation von Forschungsergebnissen in die Gesundheitsversorgung (vgl. S. 193). Regelmäßig merkten die Expert\*innen an, *dass unterschiedliche Expertisen notwendig* seien, um Nutzen- und Risikopotenziale adäquat bewerten zu können (vgl. S. 193). In diesem Zusammenhang weist auch Weichert (2018) darauf hin, dass prospektiv bei Nutzen-Risiko-Abwägungen das *Fachwissen verschiedener Disziplinen* einbezogen werden sollte. Dies könne beispielsweise im Rahmen eines sogenannten *Use and Access Committees* geschehen, das als unabhängiges, lokal agierendes Gremium handelt, und in seine Entscheidungen technisch-organisatorische, datenschutzrechtliche und ethische Betrachtungen einfließen lässt. Auch die Stellungnahme der Big Data Ethikkommission betont die Relevanz eines integralen „ethics by, in and for design“-Ansatzes für ‚Big Data‘ Applikationen (Datenethikkommission 2018). Bei einer Datenweitergabe seien in diesem Zusammenhang ebenfalls rechtliche Absicherungen sowie eine „Vielzahl interner (privater) und externer (hoheitlicher) Kontrollstellen“ (ibid.) zu diskutieren. Darüber hinaus sollten die Möglichkeiten einer Re-Identifizierung wiederkehrend evaluiert werden, wie Spitz in diesem Sammelband postuliert.

Vor dem Hintergrund dieser Bedarfe für Praktiken der Nutzen-Risiko-Abwägung können *zukünftige Forschungsvorhaben* gewinnbringend sein, die Vertreter\*innen der relevanten Akteursgruppen einbinden um beispielsweise Heuristiken einer differenzierten Nutzen-Risiko-Abwägung zu entwickeln. Hierfür können wissenschaftlich fundierte *inter- und transdisziplinäre Ansätze* hilfreich sein, deren *Methodik die Integration unterschiedlicher Expertisen* ermöglicht. Einige Expert\*innen nannten die mangelnde *Kultur des Datenteilens* der Wissenschaftler\*innen und forschenden

Ärzt\*innen als grundlegendes Hindernis in Hinblick auf die Verwendung klinischer Daten für Forschungszwecke (vgl. S. 197). Die Untersuchung von fachspezifischen Charakteristika kann dazu beitragen deren Bedürfnisse besser zu antizipieren und *adäquate Anreizsysteme* zu entwickeln.

#### 4.5 Informationelle Selbstbestimmung und Konditionen des Einwilligungsprozesses

Einige Expert\*innen berichten, dass Patient\*innen im Behandlungskontext zumeist bereitwillig Einwilligungen für eine Studienteilnahme unterschrieben (vgl. S. 198). Im deutschen Forschungskontext kann dieses Ergebnis in Bezug zu zwei sozial-empirischen Arbeiten gesetzt werden. Richter et al. (2019) ermittelten in einer quantitativen Befragung von *Patient\*innen eines Universitätsklinikums* (n=700), dass ein großer Anteil (93%) kurz zuvor in die Verwendung ihrer klinischen Daten für Forschungszwecke tatsächlich eingewilligt hatten. Voigt et al. (2020) zeigen hingegen in einer repräsentativen quantitativen Studie im Rahmen einer *Online-Befragung von Bürger\*innen* (n=1506), dass die hypothetische Bereitschaft, genetische und medizinische Daten – klinische Daten wurden also nicht separat von genetischen Daten untersucht – für die Forschung verfügbar zu machen, mit 56% relativ gering ist. Die Ergebnisse sind aufgrund unterschiedlicher Studiendesigns nur eingeschränkt vergleichbar. So wird in der Studie von Voigt et al. (2020) nach der Bereitschaft gefragt *genetische und medizinische* Daten für Forschungszwecke zur Verfügung zu stellen. Da mit genetischen Daten ein stärkerer Wunsch nach Kontrolle der Datenfreigabe assoziiert wird (Shah et al. 2019), kann dieser Aspekt zu einer geringeren Bereitschaft in dieser Studie geführt haben. Diesem Effekt zuträglich ist die in der Studie von Richter et al. (2019) direkte Betroffenheit als Patient\*in im Behandlungskontext (im Vergleich zu Bürger\*innen) und einer damit möglicherweise einhergehenden erhöhten Bereitschaft klinische Daten für Forschungszwecke freizugeben. Die Aussagen von Expert\*innen der vorliegenden Studie legen nahe, dass diese ambivalenten Ergebnisse über diese Einflüsse hinaus durch die *klinische Belastungssituation* der befragten Patient\*innen im Behandlungskontext zu erklären seien (vgl. S. 198). So wiesen einige Expert\*innen auf einen aktuell nicht optimalen *Einwilligungs- und Aufklärungsprozess im Behandlungskontext* hin: Patient\*innen seien oftmals nicht in der Lage informiert einzuwilligen aufgrund einer „Zwangslage“ im Arzt-Patienten-Verhältnis oder Disstress ausgelöst durch eine schwere Diagnose. Außerdem könnten Patient\*innen im Behandlungskontext den Eindruck gewinnen, dass die Bereitstellung ihrer Daten für die eigene Behandlung von Vorteil sei – eine (gewöhnlich) unbegründete/falsche Hoffnung, die in der Fachliteratur unter dem Schlagwort ‚Therapeutic Misconception‘ diskutiert wird (Nobile et al. 2013).

Bei Patient\*innen kann dieser nicht optimale Einwilligungs- und Aufklärungsprozess im Behandlungskontext deren Wahrnehmung des Rechts auf *informationelle Selbstbestimmung* beeinträchtigen. Auf die Relevanz der Wahrnehmung dieses Rechts

weisen Vezyridis und Timmons (2019) hin, die in einer qualitativen Untersuchung vermuten, dass der ausdrückliche Widerspruch zahlreicher Bürger\*innen im Fall des care. data Programms in Großbritannien als ein Ausdruck der fehlenden Möglichkeit ihr Recht auf informationellen Selbstbestimmung wahrzunehmen zu werten sei. Die Bürger\*innen wurden im Vorfeld weder in Entscheidungen in Hinblick auf Datenverwendung eingebunden noch wurde ihnen die Möglichkeit einer Einwilligung eingeräumt. Voigt et al. (2020) stellen auf Grundlage ihrer Untersuchungen die These auf, dass insbesondere der Wille deutscher Bürger\*innen stark ausgeprägt sei die eigene informationelle Selbstbestimmung auszuüben. Deshalb sollten *zukünftige Forschungsvorhaben* empirisch im deutschen Behandlungskontext prüfen, i) ob und wie der Einwilligungs- und Aufklärungsprozess die geschilderten Reaktionen („Zwangslage“, Disstress, ‚Therapeutic Misconception‘) bei Patient\*innen herbeiführen als auch ii) ob und wie diese Reaktionen anschließend auf die individuelle Wahrnehmung des Rechts auf informationelle Selbstbestimmung einwirken. Diese Ergebnisse können weitreichende Implikationen für die zukünftige Ausgestaltung des Einwilligungsprozesses in Deutschland nahelegen. So können zukünftige Studien untersuchen, inwiefern in Deutschland ein aktiver Akt der Einwilligung oder eine auf Kaskadenmodellen (Petersen 2018; Mertz et al. 2016) basierende Datenfreigabe akzeptabel sind. Ebenfalls sind *alternative Einwilligungskontexte außerhalb klinischer Belastungssituationen* eingehend zu prüfen, wie es beispielsweise durch das Modell der Datenspende (Krutzinna und Floridi 2019; Strech et al. 2020) ermöglicht würde.

Einen weiteren Beitrag zur informationellen Selbstbestimmung der Bürger\*innen kann die wissenschaftliche Prüfung und sich anschließende Verbesserungen der *Verständlichkeit von Informationsmaterialien* leisten. So zeigten Voigt et al. (2020), dass ein individuelles *Verständnis* von Genetik die Wahrscheinlichkeit erhöhe genetische Daten der Forschung zur Verfügung zu stellen. Es ist zu untersuchen, inwiefern im Kontext der Sekundärnutzung das *Wissen um Datenverarbeitungsprozesse sowie informationstechnologische Kompetenzen* eine Bedingung dafür darstellen, klinische Daten der Forschung zur Verfügung zu stellen und darüber hinaus das Recht auf informationelle Selbstbestimmung wahrzunehmen zu können.

#### **4.6 Die Relevanz von gesellschaftlicher Akzeptanz und Vertrauen in Institutionen des Gesundheitswesens**

Viele Expert\*innen befürchteten, dass die *gesellschaftliche Akzeptanz* unter Bürger\*innen für die Verwendung klinischer Daten nicht ausreiche und vermuteten zudem ein fragiles Vertrauen aufseiten der Bevölkerung (vgl. S. 198 f.). Öffentlichen Diskursen schrieben sie das Potenzial zu, *Vertrauen in die Institutionen des Gesundheitswesens* zu schaffen und zu stabilisieren. Vor dem Hintergrund nahezu *fehlender wissenschaftlichen Erkenntnisse um die grundlegenden Einstellungen und Präferenzen deutscher Bürger\*innen* gegenüber der Verwendung ihrer klinischen Daten für

Forschungszwecke, können *zukünftige Forschungsvorhaben* einen wertvollen Beitrag leisten, indem sie *fundamentalen Motive und Befürchtungen* der Bürger\*innen untersuchen und legitimierende Aspekte der Verwendung klinischer Daten für Forschungszwecke identifizieren. Der Literaturüberblick von Aitken et al. (2016) sowie zahlreiche weitere internationale Studien (Vezyridis und Timmons 2019; Kelley et al. 2015; Krahe et al. 2019) weisen auf die Bereitschaft von Individuen hin, ihre klinischen Daten für Forschungszwecke zur Verfügung zu stellen, sofern sie Vertrauen in *Datennutzer\*innen, Gesundheitsinstitutionen und Data Governance Modelle* besitzen. In diesem Zusammenhang hat die Untersuchung des *Vertrauens deutscher Bürger\*innen* das Potenzial *legitimierende Aspekte dieser Datennutzung und Kernpunkte ausstehender gesellschaftlicher Debatten* zu Tage fördern. Hierbei können qualitative Methoden wie Tiefeninterviews oder Fokusgruppen spezifische Sinnzusammenhänge im nationalen Kontext aufdecken, um repräsentative quantitative Studien vorzubereiten (Aitken et al. 2018).

Konsistent mit den Forderungen von Kraft et al. (2016) kann dieses grundlegende Wissen überdies Hinweise liefern, ob und auf welche Weise Patientenpräferenzen in die konkrete Ausgestaltung der Verwendung klinischer Daten zu Forschungszwecken einfließen können. So werden hierfür zwar Formen der *Bürgerbeteiligung* empfohlen (McCoy et al. 2018; European Commission's Scientific Panel for Health (SPH) 2016; Aitken et al. 2019; Strech 2018; The Nuffield Council on Bioethics 2015), doch die Umsetzung birgt einerseits praktische und konzeptuelle Herausforderungen (Strech 2018; Degelsegger und Torgersen 2011); zudem ist eine Bürger- oder Patientenbeteiligung in Hinblick auf die Forschungsnutzung von klinischen Daten in Deutschland aktuell kaum ausgeprägt. Im Bereich der Biobankforschung konnten allerdings bereits durch ‚Public and Patient Involvement‘ Forschungsdesiderate bearbeitet und Informationsmaterialien entwickelt werden (Strech et al. 2016).

Um eine nachhaltige Vertrauensbildung und die gesellschaftlich akzeptierte Verwendung klinischer Daten für Forschungszwecke zu ermöglichen, sollten deshalb auf der Grundlage empirischen Wissens um Patientenpräferenzen sowohl legitimierende Aspekte der Datennutzung als auch ausstehende gesellschaftliche Debatten identifiziert werden. Es gilt darüber hinaus Strategien zu entwickeln, welche die Partizipation von Patient\*innen/Bürger\*innen in konkreten Umsetzungsprozessen ermöglichen.

---

## 5 Limitationen

Die vorliegende Studie weist folgende Limitationen auf: Die Hälfte der befragten Expert\*innen war mit Projekten der Medizininformatikinitiative assoziiert, wodurch Sichtweisen abgebildet werden konnten, die konkret mit der Etablierung von Strukturen zur Verwendung klinischer Daten zu Forschungszwecken befasst sind. Da die Forschungsgebiete dieser Expert\*innen eine relativ etablierte Verwendung klinischer Daten und möglicherweise aktuell überproportional hohe Nutzenchancen aufwiesen, könnten die vorliegenden Ergebnisse eine Überbewertung von Nutzenaspekten abbilden.

Die gewählte Sampling Strategie mit maximaler Variation der Interviewpartner ermöglichte es, das Spektrum an Nutzen- und Risikopotenzialen sowie gemeinsame Deutungsmuster der Akteursgruppen darzustellen. Sie limitierte jedoch die Informationsdichte bezüglich spezifischer Forschungsgebiete und Anwendungsfelder der datengetriebenen medizinischen Forschung und Anwendungsfelder.

**Danksagung** Dieses Manuskript wurde im Rahmen des DFG geförderten Projekts „Learning from Clinical Data (LinCDat)“ erstellt. Ein besonderer Dank gilt allen Expert\*innen, ohne deren Teilnahme an den Interviews dieser Beitrag nicht hätte entstehen können. Wir möchten uns auch bei unseren juristischen Projektpartnern Prof. Dr. Kai Cornelius und Markus Spitz (Universität Heidelberg) bedanken. Ebenso danken wir Hellen Kachler (Nationales Centrum für Tumorerkrankungen, Heidelberg) für die Unterstützung bei der Transkription und Auswertung des Interviewmaterials.

**Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – 406103282**

---

## Literatur

- Aitken M, de St. Jorre J, Pagliari C, Jepson R, Cunningham-Burley S (2016) Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics* 17(1):73. <https://doi.org/10.1186/s12910-016-0153-x>
- Aitken M, Tully MP, Porteous C, Denegri S, Cunningham-Burley S, Banner N, Black C, Burgess M, Cross L, Van Delden J, Ford E, Fox S, Fitzpatrick N, Gallacher K, Goddard C, Hassan L, Jamieson R, Jones KH, Kaarakainen M, Willison JD (2019) Consensus Statement on Public Involvement and Engagement with Data-Intensive Health Research. *Int J Population Data Sci* 4:1–6. <https://doi.org/10.23889/ijpds.v4i1.586>
- Aitken M, Porteous C, Creamer E, Cunningham-Burley S (2018) Who benefits and how? Public expectations of public benefits from data-intensive health research. *Big Data Soc* 5(2):1–12. <https://doi.org/10.1177/2053951718816724>
- Arbeitsgruppe Personalisierte Medizin des BAG (2017) Aktuelle Entwicklungen in der datengetriebenen Medizin und die damit verbundenen Herausforderungen und Aufgaben für das BAG. In: Schweizerische Eidgenossenschaft. Bundesamt für Gesundheit BAG, Bern
- Blachetta F, Bauer M, Poerschke K, Bieber N, Solbach T, Leppert F, Greiner W, Bernnat R (2016) Weiterentwicklung der eHealth-Strategie: Studie im Auftrag des Bundesministeriums für Gesundheit. In: Strategy & PwC
- Bogner A, Littig B, Menz W (2014) Interviews mit Experten: eine praxisorientierte Einführung. Springer-Verlag, Wiesbaden
- Budrionis A, Bellika JG (2016) The learning healthcare system: where are we now? A systematic review. *J Biomed Inform* 64:87–92. <https://doi.org/10.1016/j.jbi.2016.09.018>
- Datenethikkommission (2018) Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung
- Degelsegger A, Torgersen H (2011) Participatory paternalism: citizens' conferences in Austrian technology governance. *Sci Pub Policy* 38(5):391–402
- Deutsche Hochschulmedizin e. V. (2019) Mehrwert digitaler Daten für Versorgung und Forschung sicherstellen – Politik muss jetzt handeln, Berlin

- Deutscher Ethikrat (2017) Big Data und Gesundheit – Datensouverenität als informationelle Freiheitsgestaltung
- BETTER RESEARCH FOR BETTER HEALTH – A vision for health and biomedical research from the Scientific Panel for Health (2016)
- European Commission (2014) Special Eurobarometer 419: Public perceptions of science, research and innovation. In: Directorate-General for Research and Innovation, European Commission Brussels
- European Commission (2015) Special Eurobarometer 431 “Data protection”. In: Directorate-General for Research and Innovation, European Commission Brussels
- Faden RR, Kass NE, Goodman SN, Pronovost P, Tunis S, Beauchamp TL (2013) An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hastings Cent Rep* 43(s1):16–27
- Fiscella K, Tobin JN, Carroll JK, Hua H, Ogedegbe G (2015) Ethical oversight in quality improvement and quality improvement research: new approaches to promote a learning health care system. *BMC Med Ethics* 16(1):1–6. <https://doi.org/10.1186/s12910-015-0056-2>
- Flick U (2017) *Qualitative Sozialforschung: Eine Einführung* Rowohlt enzyklopädie im Rowohlt Taschenbuch Verlag, Reinbek bei Hamburg
- Ford E, Boyd A, Bowles JKF, Havard A, Aldridge RW, Curcin V, Greiver M, Harron K, Katikireddi V, Rodgers SE, Sperrin M (2019) Our data, our society, our health: a vision for inclusive and transparent health data science in the United Kingdom and beyond. *Learn Health Syst* 1-12. <https://doi.org/10.1002/lrh2.10191>
- Gaskell G, Gottweis H, Starkbaum J, Gerber MM, Broerse J, Gottweis U, Hobbs A, Helén I, Paschou M, Snell K, Soulier A (2012) Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet* 21:14–20. <https://doi.org/10.1038/ejhg.2012.104>
- Hobbs A, Starkbaum J, Gottweis U, Wichmann H, Gottweis H (2012) The privacy-reciprocity connection in biobanking: comparing German with UK strategies. *Public Health Genomics* 15(5):272–284
- Ipos MORI (2016) *The One-Way Mirror: Public attitudes to commercial access to health data – Report prepared for the Wellcome Trust*
- Kass NE, Faden RR, Goodman SN, Pronovost P, Tunis S, Beauchamp TL (2013) The research-treatment distinction: a problematic approach for determining which activities should have ethical oversight. *The Hastings Center report Spec No:4-15*. <https://doi.org/10.1002/hast.133>
- Kelley M, James C, Alessi Kraft S, Korngiebel D, Wijangco I, Rosenthal E, Joffe S, Cho MK, Wilfond B, Lee SS (2015) Patient perspectives on the learning health system: the importance of trust and shared decision making. *Am J Bioeth* 15(9):4–17. <https://doi.org/10.1080/15265161.2015.1062163>
- Kraft SA, Cho MK, Constantine M, Lee SS, Kelley M, Korngiebel D, James C, Kuwana E, Meyer A, Porter K, Diekema D, Capron AM, Alicic R, Wilfond BS, Magnus D (2016) A comparison of institutional review board professionals’ and patients’ views on consent for research on medical practices. *Clinical trials (London, England)* 13(5):555–565. <https://doi.org/10.1177/1740774516648907>
- Krahe M, Milligan E, Reilly S (2019) Personal health information in research: perceived risk, trustworthiness and opinions from patients attending a tertiary healthcare facility. *J Biomed Inform* 95:1–10. <https://doi.org/10.1016/j.jbi.2019.103222>
- Krutzinna J, Floridi L (2019) *The ethics of medical data donation*. Springer, London
- Kuckartz U (2018) *Qualitative Inhaltsanalyse*. Beltz Juventa, Weinheim Basel
- Laurie G, Stevens L, Jones KH, Dobbs C (2014) A review of evidence relating to harm resulting from uses of health and biomedical data. In: *Bioethics NCo (Hrsg) Scoping Study*



- Mayo RM, Summey JF, Williams JE, Spence RA, Kim S, Jagsi R (2017) Qualitative study of oncologists' views on the CancerLinQ Rapid Learning System. *J Oncol Pract* 13(3):176–184. <https://doi.org/10.1200/JOP.2016.016816>
- Mayring P (2008) *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Beltz, Weinheim
- McCoy MS, Jongsma KR, Friesen P, Dunn M, Neuhaus CP, Rand L, Sheehan M (2018) National Standards for Public Involvement in Research: missing the forest for the trees. *J Med Ethics* 44(12):801–804. <https://doi.org/10.1136/medethics-2018-105088>
- Mertz M, Jannes M, Schlomann A, Manderscheid E, Rietz C, Woopen C (2016) Digitale Selbstbestimmung. Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health (ce. In: Cologne Center for Ethics R, Economics, and Social Sciences of Health (ceres) (Hrsg), Köln
- Meuser M, Nagel U (2009) Das Experteninterview—konzeptionelle Grundlagen und methodische Anlage. In: *Methoden der vergleichenden Politik- und Sozialwissenschaft*. S 465–479, Springer
- Myers J, Frieden TR, Bherwani KM, Henning KJ (2008) Ethics in public health research: privacy and public health at risk: public health confidentiality in the digital age. *Am J Public Health* 98(5):793–801. <https://doi.org/10.2105/ajph.2006.107706>
- Nobile H, Vermeulen E, Thys K, Bergmann MM, Borry P (2013) Why do participants enroll in population biobank studies? A systematic literature review. *Expert Rev Mol Diagn* 13(1):35–47. <https://doi.org/10.1586/erm.12.116>
- Perera G, Holbrook A, Thabane L, Foster G, Willison DJ (2011) Views on health information sharing and privacy from primary care practices using electronic medical records. *Int J Med Inform* 80(2):94–101. <https://doi.org/10.1016/j.ijmedinf.2010.11.005>
- Petersen C (2018) Through patients' eyes: regulation, technology, privacy, and the future. *Yearb Med Inform* 27(01):10–15. <https://doi.org/10.1055/s-0038-1641193>
- Richter G, Borzikowsky C, Lieb W, Schreiber S, Krawczak M, Buyx A (2019) Patient views on research use of clinical data without consent: legal, but also acceptable? *Eur J Hum Genet* 27(6):841–847
- Salerno J, Knoppers BM, Lee LM, Hlaing WM, Goodman KW (2017) Ethics, big data and computing in epidemiology and public health. *Ann Epidemiol* 27(5):297–301
- Schilsky RL, Michels DL, Kearbey AH, Yu PP, Hudis CA (2014) Building a rapid learning health care system for oncology: the regulatory framework of CancerLinQ. *J Clin Oncol* 32(22):2373–2379. <https://doi.org/10.1200/jco.2014.56.2124>
- Semler SC, Wissing F, Heyder R (2018) German medical informatics initiative. *Methods Inf Med* 57(S 01):e50–e56. <https://doi.org/10.3414/me18-03-0003>
- Shah N, Coathup V, Teare H, Forgie I, Giordano GN, Hansen TH, Groeneveld L, Hudson M, Pearson E, Ruetten H, Kaye J (2019) Motivations for data sharing—views of research participants from four European countries: a DIRECT study. *Eur J Human Genetics: EJHG* 27(5):721–729. <https://doi.org/10.1038/s41431-019-0344-2>
- Strech D (2018) Normative Governance der Big Data Forschung. *Forschung Politik Strategie Management* 2+3:53–58
- Strech D, Bein S, Brumhard M, Eisenmenger W, Glinicke C, Herbst T, Jahns R, von Kielmansegg S, Schmidt G, Taupitz J (2016) A template for broad consent in biobank research. Results and explanation of an evidence and consensus-based development process. *Eur J Med Genetics* 59(6–7):295–309
- Strech D, Kielsmansegg S, Zenker S, Krawczak M, Semler S (2020) Wissenschaftliches Gutachten „Datenspende“ – Bedarf für die Forschung, ethische Bewertung, rechtliche, informations-technologische und organisatorische Rahmenbedingungen. In: Bd 1.1. Bundesministerium für Gesundheit, Berlin

- The Nuffield Council on Bioethics (2015) The collection, linking and use of data in biomedical research and health care: ethical issues
- Thiel R, Deimel L, Schmidtman D, Piesche K, Hüsing T, Rennoch J, Stroetmann V, Stroetmann K (2018) #SmartHealthSystems – Digitalisierungsstrategien im internationalen Vergleich. In: Stiftung B (Hrsg)
- Thorogood A (2020) International Data Sharing and Rare Disease: The Importance of Ethics and Patient Involvement. In: He Wu Z (Hrsg) Rare Diseases. IntechOpen, S 177–193
- Vezyridis P, Timmons S (2019) Resisting big data exploitations in public healthcare: free riding or distributive justice? *Sociol Health Illn* 41(8):1585–1599
- Voigt TH, Holtz V, Niemiec E, Howard HC, Middleton A, Prainsack B (2020) Willingness to donate genomic and other medical data: results from Germany. *Eur J Human Genetics*, 1–10
- Weichert T (2018) Big Data im Gesundheitsbereich. In: Bundesministerium für Bildung und Forschung
- Whicher D, Kass N, Faden R (2015) Stakeholders' views of alternatives to prospective informed consent for minimal-risk pragmatic comparative effectiveness trials. *J Law Med Ethics J Am Soc Law Med Ethics* 43(2):397–409. <https://doi.org/10.1111/jlme.12256>
- Winkler E (2017) Big Data in Forschung und Versorgung: ethische Überlegungen und Lösungsansätze. *Frankf Forum Diskurse* 3:22–31

**Open Access** Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

