NAKSE
Gemeinsam Fortschritt erreichen
18. und 19. September 2025
Holiday Inn Berlin Airport Conference Centre
sowie online auf achse.live
#NAKSE

Auf dem Weg zu **Digitalem Nutzen** für **Menschen mit Seltenen Erkrankungen durch Real-Welt-Daten** der Universitätsmedizin

Collaboration for **Ra**re **D**isease **O**bservation and **R**esearch (LAB**RADOR**) in Kooperation mit dem MII-NUM-Forschungsdatenportal für Gesundheit (FDPG)

Miriam Hübner[1]; Steffen Sander[1]; Claudia Finis[1,2] ; Marco Schaarschmidt[1]; Josef Schepers[1]

mit Mitarbeit von Christine Mundlos[3], Simone Kesten[4], Urban Geisthoff[4,5,6], Jürgen Schäfer[5], Janbernd Kirschner[9], Elisabeth Nyoungui[10], Jana Zschüntzsch[11], Richard Röttger[12], Romina Blasini[6], Marie Gebhardt[13], Philip Kleinert[13], Timo Fuchs[8], Stefanie Ritzrow[1], Johanna Nothacker[1], Stefanie Rudolph[1], Maria Rönnefarth[1] und von vielen anderen

1 Berlin Institute of Health at Charité, 2 Deutsche Gesellschaft für Osteogenesis Imperfecta (DOIG), 3 Allianz Chronischer Seltener Erkrankungen (ACHSE e.V.), 4 Morbus Osler Selbsthilfe e. V., 5 Universitätsklinikum Gießen-Marburg, Standort Marburg, Zentrum für unerkannte und seltene Erkrankungen (ZusE), 6 Universitätsklinikum Gießen und Marburg, VASCERN HHT Reference Network, Standort Marburg, Klinik für Hals-, Nasen- und Ohrenheilkunde, Kopf- und Hals-Chirurgie, 7 Universitätsklinikum Gießen-Marburg, Standort Gießen, Datenintegrations-zentrum, 8 Universitätsklinikum Erlangen, Datenintegrationszentrum, 9 Universitätsklinikum Freiburg, Klinik für Neuropädiatrie, SMArtCARE-Register, 10 Institut für Medizinische Informatik, Universitätsmedizin Göttingen 11 Universitätsklinikum Göttingen, Klinik für Neurologie, 12 Syddansk Universitet Odense, Department of Mathematics and Computer Science, 13 Forschungsdatenportal für Gesundheit in der Medizininformatik-Initiative und im Netzwerk Universitätsmedizin

Hintergrund:

Das **Nationale Aktionsbündnis für Menschen mit Seltenen Erkrankungen (NAMSE)** fordert, dass Digitalisierungsmaßnahmen im deutschen Gesundheitswesen mit der Förderung der **einrichtungsübergreifenden gemeinsamen Datennutzung** besonders den Menschen mit Seltenen Erkrankungen (SE) zugutekommen müssen.

Drei Herausforderungen spielen eine große Rolle und finden in der Medizininformatik-Initiative (MII) des BMFTR und im Netzwerk Universitätsmedizin (NUM) Berücksichtigung.

1. Präzisierung der Diagnosen:

Für die digitale Nutzung von Diagnoseangaben wird die grobe, statistische ICD-10-GM-Klassifikation durch die exakte Kodierung der SE-Diagnosen mittels Kennziffern der Orphanet-Ontologie (ORDO), auch **ORPHAcodes** genannt, ergänzt.

- Nur wenige hundert SE wie beispielsweise die *Hereditäre hämorrhagische Teleangiektasie (HHT)* [ICD:I78.0; ORPHA:774], die *Spinale Muskelatrophie (SMA)* [ICD:G12.0; Orpha:70] und die *Osteogenesis imperfecta (OI)* [ICD:Q78.0; ORPHA:666] können durch ihren ICD-Kode in Datensammlungen gefunden werden.
- Dem stehen fast 10.000 SE gegenüber, von denen regelmäßig mehrere in einer ICD-Subkategorie zusammengeführt sind und bis zur Einführung der ICD-11 nur durch die Orpha-Kodierung, die sich im Prozess des Ausrollens befindet, findbar werden.

2. Sekundär- und Mehrfachnutzung von Erkrankungsdaten:

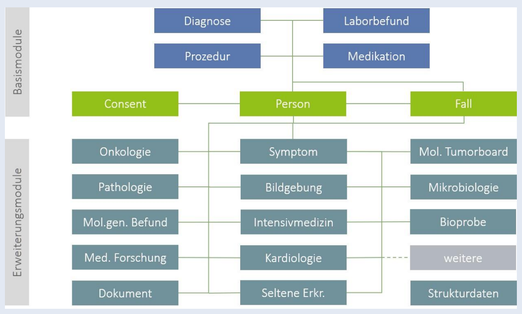
Dokumentationen sollen wegen Ressourcenknappheit und zur Qualitätsunterstützung mehrfach genutzt werden. Insbesondere werden in den Universitätsklinika (UK) Daten aus dem Versorgungskontext in Datenintegrationszentren (DIZ) so strukturiert aufbereitet, dass sie für die Beantwortung von Forschungsfragen herangezogen werden können.

Neben **Orpha-Kodierung** soll der Forschungswert der Versorgungsdaten im Bereich der SE durch **standardisierte Phänotypisierung und Genotypisierung** gesteigert werden können. Über den Abrechnungsweg gemäß § 301 SGB V gelangen die stationären ORPHA-Angaben auch zu den Krankenkassen.

3. Gemeinsame Nutzung von Erkrankungsdaten:

Die Aufbereitung der Versorgungsdaten in den DIZ verfolgt nicht nur das Ziel, die Daten jeweils hausintern für Versorgungs- (inklusive Abrechnung und Qualitätssicherung) sowie Forschungszwecke nutzen zu können, sondern auch **einrichtungsübergreifend**.

Forschende und Patientengruppen sollen die Daten vieler/aller Häuser nutzen können. Hierfür orientiert sich die Aufbereitung der Daten in allen DIZ an einem gemeinsamen Informationsmodell, dem **MII Kerndatensatz.**



Module des MII Kerndatensatzes

Technisch und organisatorisch wird die gemeinsame Nutzung der in den DIZ abgestimmten gespeicherten Erkrankungsdaten durch das **Forschungsdatenportal für Gesundheit (FDPG)**[1] unterstützt:

- Das FDPG-**Machbarkeitsportal** ermöglicht einen dynamischen, anonymen Einblick in die spezifischen Mengen der in den DIZ verfügbaren Daten.
- Über das FDPG-**Antragsportal** können an alle beteiligten UK Datennutzungsanträge für definierte Projekte gestellt werden.

Fragestellungen:

Das ***Collaboration for Rare Disease Observation and Research (LABRADOR)*** stellt eine vom BIH in der Charité koordinierte Plattform dar, auf der exploriert wird, ob und auf welchem Niveau Forschungsfragen aus dem Bereich der SE unter Nutzung der FDPG-Infrastruktur bearbeitet werden können.

Folgende Themen stehen in den verschiedenen Explorationsstudien im Vordergrund:

i. **Klinische Epidemiologie**
(Zählung der Klinischen SE-Fälle mit Alterspyramiden und Regionalverteilung)
ii. **Diagnoseunterstützung**
(z. B. HHT-Studie)
iii. **Therapiebeobachtung**
(z. B. SMA-Studie mit Zolgensma)
iv. **Machbarkeitsanalysen für Klinischen Studien**
(z. B. OI-Studie mit Losartan)

Beteiligte Standorte des NUM

Daten und Methoden:

Genutzt werden die durch das FDPG vermittelten Datenrepositorien. Umgesetzt werden Verteilte Analysen und es sollen möglichst viele UK zur Mitwirkung bewegt werden.

Es werden **drei Phasen** unterschieden:

- **Phase 1**: Studien zu SE, bei denen es einen ICD-Kode für die Diagnose gibt (HHT, SMA, OI).
- **Phase 2**: Studien mit Einbezug von ORPHA-kodierten Erkrankungen (DM1, HPP, Alström, $ATTR_{wt}$).
- **Phase 3**: Exploration des Potentials einer erweiterten SE-Basisdokumentation.

LABRADOR & FDPG treffen Klinische Epidemiologie – Exploration der Orpha-kodierten SE

Fragestellungen

Auch 16 Jahre nach dem Start des NAMSE lautet eine grundsätzliche Frage nach wie vor, *wie viele Personen in Deutschland von einer der bis 10.000 verschiedenen SE im Einzelnen und insgesamt betroffen sind.*

Es ist nur eine **Schätzung**, dass es insgesamt vier Millionen SE-betroffene Menschen sind.

Neben den Auswertungen in anderen Datenbeständen wie in verschiedenen SE-Registern sollen **Fallzählungen in den DIZ-Daten** der Universitätsklinika unterstützten, die Epidemiologie der SE zu erhellen.

- Gefragt nach **Häufigkeiten von Personen mit kodierten SE-Diagnosen** ohne Doppelzählungen.
- Differenzierend wird die Verteilung der Erkrankungen nach **Alter, Geschlecht und Region (Wohnort)** sowie die **Distanz zwischen Wohnort und Leistungsort** ermittelt.

Ergänzt werden **Datenqualitätsprüfungen** hinsichtlich des Ausrollens der Orpha-Kodierung in Form von Zeitreihen zu den ORPHAcodes und in Form von Orpha-Quoten zu den SE-eindeutigen ICD-10-Kodes.

Hintergrund

Es wird geschätzt, dass in Deutschland rund vier Millionen Personen (durchschnittlich 400 Personen je SE) von einer SE betroffen sind. Die Spanne der Betroffenen je SE reicht von einer bis zur definierten Obergrenze der Prävalenz von *5 Betroffenen in 10.000* (rund 42.000 Menschen in DE).[2]

- Von ***Ultraseltenen Erkrankungen*** spricht man unterhalb der Schwelle von *2 in 100.000* (in DE < als 1.680) Menschen mit einer konkreten SE.

Jedoch ist weder die Gesamtzahl an Betroffenen, noch die Anzahl je verschiedener Erkrankungen bisher durch eine belastbare Gesundheitsberichterstattung belegt.

Das liegt unter anderem daran, dass alle Statistiken verschiedener Quellen (z.B. RKI, Statistikämter, Kassenärzte und DRG-Kalkulation) zur Kodierung von Erkrankungen auf die 10. Revision der Internationalen Statistischen Klassifikation der Krankheiten (ICD-10-GM) zurückgreifen.

Herausforderung: ICD-10-GM & SE:

1. Der Begriff „Statistische Klassifikation“ beinhaltet, dass bei rund 12.000 ICD-Kodes für mehr als 20.000 Erkrankungsformen **nicht für jede Erkrankung eine eigene Kennziffer zur Verfügung** steht.
2. Zudem kommt es vor, dass in **einem Kode Seltene und häufige Erkrankungen zusammen kategorisiert** sind.
3. Nur **wenige hundert Kodes** wie I78.0 (HHT), G12.0 (SMA) und Q78.0 (OI) stehen **eindeutig** für Erkrankungen, weswegen wir sie bis zum fertigen Ausrollen der exakten Kodierung für alle SE mit ORPHAcode für Initialstatistiken nutzen.

Lösung:

Exakte Kodierung → jede SE hat in einem einheitlich benutzten Kodier-system weltweit **genau einen Kode**.

In Deutschland ist die Orpha-Kodierung für den stationären Sektor seit 2023 **gesetzlich vorgeschrieben**[2] und beginnt in 2025 in den DIZ anzukommen. In spezialisierten ambulanten Bereichen sollen ORPHAcodes genutzt werden, damit auch dort die diagnostizierten Fälle auffindbar und zählbar werden.

Die Frage nach der Häufigkeit einzelner SE und nach der Gesamtzahl der Betroffenen in DE kann nicht vollständig mit UK-Daten beantwortet werden. Es werden die Daten aller Krankenhäuser benötigt:

- Daher erfolgt u. a. parallel die abgestimmte Befassung mit den SE-Zahlen des Forschungsdatenzentrum für Gesundheit (FDZ)[3] des BfArM.

Da aber die UK mit ihren SE-Zentren für sehr viele SE-Betroffene mindestens episodisch eine wichtige Anlaufstelle sind, darf unter bestimmten Voraussetzungen bei Beteiligung vieler UK von einem hohen Erfassungsgrad ausgegangen werden.

Eine grundlegende **Arbeitshypothese** der LABRADOR-Studien lautet, dass es sich bei rund *15 % der 21 Mio. Personen mit Daten in den DIZ um SE-Betroffene* handelt, bzw., dass *drei Viertel aller SE-Betroffenen in DE gelegentlich Kontakt zu einem UK* haben.

Sobald die Orpha-Kodierung unter Einbeziehung der DIZ flächendeckend umgesetzt ist, kann mit deren Daten Orpha-Beobachtung und bedingt Orpha-Forschung betrieben werden – was durch die LABRADOR-Studien verifiziert oder falsifiziert werden soll.

Daten und Methoden

Es ist beim FDPG ein Datennutzungsantrag zur Realisierung einfacher Häufigkeitszählungen für ausgewählte Erkrankungen mit einem SE-eindeutigen ICD-Kode und für alle Erkrankungen mit einem ORPHAcode eingereicht worden. Für die begrenzte Liste mit ca. 300 verschiedenen SE-eindeutigen ICD-Kodes sowie für alle verfügbaren ORPHAcodes der Ebene „disorder“ werden die Ergebnisse bei hinreichenden Fallzahlen nach **Alter, Geschlecht und Region (Bundesland)** differenziert.

Das erwartete Anfluten der Orpha-Kodierung bei stationären und ambulanten Fällen im Verlauf der Jahre 2025 und 2026 wird im Zeitverlauf in absoluten Zahlen personenbezogen verfolgt – mit der Option für Feedback an die Kliniken. Bei den SE-eindeutigen ICD-Kodes wird der Anteil der Orpha-kodierten Fälle, der gegen 100 % streben müsste, im Zeitverlauf verfolgt.

Ergebnisse und Diskussion

Vorläufig kann im Sommers 2025 angezeigt werden, dass in allen UK die Orpha-Kodierung angelaufen ist und in einigen UK die ORPHAcodes auch im FDPG-Portal angekommen sind. Näheres folgt.

[1] https://forschen-fuer-gesundheit.de
[2] https://www.bundesgesundheitsministerium.de/themen/praevention/gesundheitsgefahren/seltene-erkrankungen.html
[3] https://www.forschungsdatenzentrum-gesundheit.de

Funding: Screen4Care has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 101034427. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. This material reflects only the author's view. The JU is not responsible for any use that may be made of information it contains.