

Risikoanalyse für die Sekundärnutzung klinischer Daten

Modell zur Bewertung von Risiken für Patienten im Rahmen medizinischer Forschung mit Behandlungsdaten

Martin Jungkunz¹, Anja Köngeter¹, Katja Mehlis¹, Eva C. Winkler¹, Christoph Schickhardt²

¹Nationales Centrum für Tumorerkrankungen (NCT), Sektion „Translationale Medizinethik“, Abteilung Medizinische Onkologie, Universitätsklinikum Heidelberg
²Nationales Centrum für Tumorerkrankungen (NCT), Sektion „Translationale Medizinethik“, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg

Hintergrund

Jeder Arztbesuch oder Klinikaufenthalt ist mit der Produktion von Daten verbunden. Diese Daten bergen großes Potenzial, um durch Sekundärnutzung im Rahmen medizinischer Forschung medizinisches Wissen zu generieren und somit die medizinische Versorgung zu verbessern.

Mögliche Anwendungsbereiche für die Sekundärnutzung klinischer Daten:

- Qualitätskontrolle/-verbesserung/-sicherheit
- Public Health Forschung
- Nicht-interventionelle klinische Forschung
- Explorative Verwendung (Hypothesengenerierung, Feasibility-Studien, Probandenrekrutierung)

Risiken der Sekundärnutzung klinischer Daten

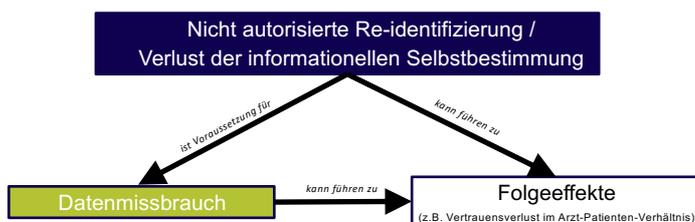
Die Sekundärnutzung klinischer Daten kann Risiken für unterschiedliche Stakeholder wie Patienten, Ärzte und Kliniken mit sich bringen. Wir beschränken unsere Risikoanalyse auf Risiken für Patienten.

Die Forschung mit klinischen Daten geschieht ohne physische Risiken für Patienten. Trotzdem bringt die Sekundärnutzung klinischer Daten Risiken anderer Art mit sich, welche im Rahmen einer ethischen Bewertung (z.B. in Ethikkommissionen) in Betracht gezogen werden müssen.

Gründe:

- Ethikkommissionen haben den Schutz des Probanden/Patienten als zentrale Aufgabe.
- Die DSGVO verlangt ein Risiko-Assessment, wenn eine Datenverarbeitung „voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge“ (DSGVO, Art. 35 Abs. 1) hat.

Zentrale Risiken für Patienten



Wir definieren **Risiko** als die Kombination von **Schweregrad** und **Wahrscheinlichkeit** eines negativen Ereignisses

- Die Re-Identifizierung ist ein Schaden für sich, aber auch eine Voraussetzung für mögliche Folgeschäden. Die **Wahrscheinlichkeit der Re-Identifizierung** kann anhand verschiedener Risikofaktoren in den Daten und ihrem Verwendungskontext abgeschätzt werden.
- Der **Schweregrad eines möglichen Folgeschadens** wird vom Informationsgehalt, welcher im Falle eines Datenmissbrauchs zu Ungunsten des Patienten verwendet werden kann.

Das Gesamtrisiko ergibt sich aus der Wahrscheinlichkeit der Re-Identifikation und der Schwere möglicher Folgeschäden. Auch wenn mögliche Folgeschäden gering sind, besteht weiterhin der primäre Schaden der Re-Identifikation.

Für eine ausführliche Darstellung, siehe:
 Jungkunz, M., et al. (in preparation): Framework for evaluating privacy risks in secondary use of clinical data.
 Jungkunz, M., et al. (in preparation): Sekundärnutzung klinischer Daten für Forschungs- und Lernaktivitäten – Begriff, Umfang, Implikationen.

Risikofaktoren von Daten in der Sekundärnutzung

Relevante Risikofaktoren

- **Einzigartigkeit:** Bis zu welchem Grad beschreibt der Datensatz eine bestimmte Person eindeutig?
- **Verfügbarkeit externer Datenbanken:** Sind Datenbanken verfügbar, mit denen die klinischen Daten abgeglichen werden können?
- **Stabilität:** Beziehen sich die Informationen in den Daten auf Zustände, die über die Zeit stabil sind?
- **Interpretierbarkeit:** Sind die Informationen in den Daten ohne technische Ressourcen und/oder Expertenwissen interpretierbar?

Wahrscheinlichkeit nicht autorisierter Re-identifikation
 datenspezifisch

- **Jurisdiktion:** Verlassen die Daten die primäre Institution/das Land/die EU?
- **Technische Möglichkeiten:** Was sind die technischen Möglichkeiten des Empfängers in Bezug auf Re-Identifikationsmethoden, Data Mining und Verknüpfung?
- **Technologische Maßnahmen:** Was wird unternommen, um nicht autorisierte Re-Identifikation zu vermeiden?
- **Aufbewahrungsdauer:** Wie lange werden die Daten gespeichert?
- **Vertrauenswürdigkeit:** Wird der Empfänger als vertrauenswürdig angesehen?
- **Data-Sharing-Modell:** Wie viele Personen werden Zugang zu dem Datensatz haben?

Schweregrad des möglichen Schadens im Falle eines Datenmissbrauchs
 kontextuell

- **Vorhersagepotenzial:** Lassen die Daten Rückschlüsse auf den zukünftigen Gesundheitszustand des Patienten zu?
- **Diskriminierungspotenzial:** Können die Informationen in den Daten zur Diskriminierung von Patienten verwendet werden oder haftet ihnen ein gesellschaftliches Stigma an?
- **Reichweite:** Sind die Informationen in den Daten auf die betroffene Person beschränkt oder betreffen sie darüber hinaus andere, z.B. Verwandte?
- **Abdeckungszeitraum:** Beschreiben die Informationen in den Daten einen einzelnen Zeitpunkt oder einen längeren Zeitraum in der Krankengeschichte des Patienten?
- **Zukünftiges Informationspotential:** Werden die Daten in den nächsten Jahren voraussichtlich an Aussagekraft gewinnen?

Anwendung der Risikoabschätzung

Unsere Risikoabschätzung kann sowohl in Ethik-Kommissionen als auch zukünftig in Data-Access-Committees zur Anwendung kommen. Hier kann sie sowohl die Identifikation und Bewertung von Risiken für Patienten als auch die Erarbeitung potenzieller Gegenmaßnahmen unterstützen.

Bewertung durch Ethikkommission oder Data Access Committee

- 1 **Identifizierung und Bewertung einzelner Risikofaktoren**
Der Schweregrad jedes Risikofaktors wird bewertet
- 2 **Umfassende Bewertung**
Zusammenfassung aller Risikofaktoren für ein Gesamtbild
Vollständige Risikobewertung
- 3 Einbeziehung weiterer Risiken anderer betroffener Akteursgruppen
Umfassende ethische Bewertung
- 4 Einbeziehung weiterer Aspekte, die für die ethische Bewertung relevant sind; insbesondere der potenziellen Nutzen der geplanten Studie
- 5 **Verminderung einzelner Risikofaktoren**
Die Identifizierung und Bewertung einzelner Risikofaktoren (Schritt 1) informiert mögliche gezielte Gegenmaßnahmen