

# Klinisch-epidemiologische Daten der Krebsregister für Forschende: Fit für Krebsregisterdaten?

Webseminar der Arbeitsgruppe Krebsepidemiologie der DGEpi und GMDS und der Plattform § 65c-Register am Donnerstag, 5. Mai 2022, 13:00 bis 17:00 Uhr (online)

## Zusammenfassung Session 1 | Sicherung der Qualität der Krebsregisterdaten

Die erste Session des Workshops hatte die Sicherung der Qualität der Krebsregisterdaten zum Thema. Dieses Thema wurde in zwei Impulsbeiträgen aus verschiedenen Perspektiven beleuchtet:

**Herr Dr. Philipp Morakis**, Leiter der Geschäftsstelle Qualitätskonferenzen des Krebsregisters Baden-Württemberg, stellte Ansatzpunkte dar, wie sich ein Krebsregister nach §65c SGB V der Frage der Datenqualität nähern kann. Am Krebsregister Baden-Württemberg werden hierfür interdisziplinäre Datenqualitätsteams zusammengestellt, die entitätsspezifische Prüfkriterien (u.a. auf Basis von Leitlinien) entwickeln, vorliegende Daten daran prüfen und fehlende bzw. unplausible Daten bei den Meldern nachfragen und ggf. eine Begehung mit dem Melder zu vereinbaren. Besondere Herausforderungen dabei sind, solche Prüfungen für möglichst viele bzw. alle Entitäten und nicht nur anlassbezogen umzusetzen sowie feldübergreifende Prüfungen bzw. Prüfungen des gesamten Therapiegeschehnisses zu implementieren.

**Herr Dr. Olaf Schoffer**, Statistiker am Zentrum für evidenzbasierte Gesundheitsversorgung (ZEGV) in Dresden, berichtete aus der Perspektive eines externen Forschenden, der Krebsregisterdaten analysiert. Im Rahmen der Analysen können Auffälligkeiten in den Daten beobachtet werden, die unbedingt mit den datenliefernden Registern geklärt werden müssen, um nicht zu falschen Interpretationen zu gelangen. Auffälligkeiten können z.B. durch den Vergleich verschiedener datenliefernder Register oder den Vergleich mit erwarteten Häufigkeiten erkannt werden. Häufig gibt es Missverständnisse darüber, wie Variablen belegt werden oder technische Besonderheiten bei der Generierung des Datenexportes. Für die Interpretation ist ebenfalls wichtig, Hintergründe der klinischen Krebsregistrierung und die Arbeitsweise des liefernden Registers zu kennen. Eine offene und vertrauensvolle Diskussion zwischen dem liefernden Register und den extern Forschenden ist essentiell.

In der anschließenden Diskussion wurden folgende Aspekte vertieft: Wie können Melder dazu motiviert werden, alle erforderlichen Meldungen zu übermitteln und diese auch in guter Qualität? Wie können fehlende Meldungen erkannt und eingeholt werden, insbesondere in den Fällen, wo die Versorgung im ambulanten Bereich stattfindet? Wie können überhaupt Kriterien entwickelt werden, um Datenqualität zu beurteilen? Es zeigte sich zum einen ein großer Bedarf in den Krebsregistern und zum anderen auch, dass bereits verschiedene Aktivitäten in Richtung Qualitätssicherung begonnen wurden. Hinsichtlich des Umgangs mit nicht gemeldeten Ereignissen im Krankheitsverlauf

wurde z.B. die Auswertung von Arztverzeichnissen sowie der Vergleich von Krebsregisterdaten mit Krankenkassendaten oder Daten der Kassenärztlichen Vereinigungen besprochen.

## Zusammenfassung Session 2 | Methodische Aspekte bei der Auswertung von Krebsregisterdaten

An die Session zur Sicherung der Qualität schloss sich der Themenbereich zu methodischen Aspekten bei der Auswertung von Krebsregisterdaten an.

**Der Beitrag von Herrn Dr. Vinzenz Voelkel**, Universität Regensburg zu modernen Auswertungsverfahren in der Krebsregisterforschung schlug den Bogen von John Snows Beobachtung von Ereignissen und möglicher Infektionsquelle, deren räumlicher Beschreibung und hypothesengestützten sampeln einer verfügbaren Kontrollgruppe im Jahr 1854, zur Vorhersage von Tumorrezidiven mit Krebsregisterdaten.

Die vom Vortragenden zur Verfügung gestellte Präsentation gibt einen Einblick in die Strukturierung einer Observationsstudie analog zu den Prinzipien eines (hypothetischen) RCTs (Target trial emulation) um zum Vergleich zweier Therapieformen genutzt werden zu können. Weiterhin wurde das INFLUENCE2.0 Modell der Risikoprädiktion von rezidiven vorgestellt.

In der sich anschließenden Diskussion wurde die Datenqualität als zentrale Voraussetzung zur Nutzung der deutschen Daten thematisiert. Im deutschen Kontext Studien auf Registerdaten als etablierte Datenpools aufzusetzen (wie Krebsregisterdaten), ist mit der aktuellen Datenqualität hinsichtlich Vollständigkeit noch stark limitiert. Ein Risikovorhersagemodell wie das vorgestellte Rezidiv- und Zweittumorrisikomodell auf Basis der niederländischen Krebsregisterdaten derzeit nicht mit deutschen Registerdaten umsetzbar. Auch Ansätze zum Datenlinkage mit weiteren Datenquellen ist in Zusammenarbeit mit einem forschungsstarken landesweiten Krebsregister in den Niederlanden bereits etabliert, in Deutschland ausbaufähig.

**Der zweite Vortrag** behandelte aktuelle Methodiken zur Untersuchung regionaler Einflussfaktoren auf unterschiedliche Outcomes von Krebserkrankungen. **Dr. Philipp Jaehn**, Institut für Sozialmedizin und Epidemiologie, Medizinische Hochschule Brandenburg Theodor Fontane zeigte methodische und inhaltliche Forschungsstände und Forschungslücken von Mehrebenenanalysen zur Untersuchung regionaler Einflussfaktoren auf. Für epidemiologische Studien in Deutschland stehen jedoch einige Datenquellen zu regionalen Einflussfaktoren zur Verfügung.

In diesem Zusammenhang wurde in der sich anschließenden Diskussion u.a. die Notwendigkeit einer Differenzierung von Konzepten wie Zusammenhang von Sozialkapital und sozioökonomischem Status auf Gesundheitsoutcomes aufgegriffen. Insbesondere wenn mehrdimensionale Indizes wie der German Index of Multiple Deprivation abgebildet werden. An einem Beispiel wurde verdeutlicht, welche spezifischen Aspekte regionaler Deprivation unabhängig von anderen bedeutsam für ein Gesundheitsoutcome sind und welchen Einfluss Kollinearität nimmt.

## Zusammenfassung Session 3 | Bedeutung der Datenstruktur für die Analyse von Krebsregisterdaten

Als dritter Themenbereich wurde die Bedeutung der Datenstruktur für die Analyse von Krebsregisterdaten genauer beleuchtet. Auch hier sind zwei Impulsvorträge gehalten worden.

Im **ersten Beitrag** von Herrn **Dr. Altmann** aus Gießen wurden **die semantischen Zusammenhänge – welche Daten kommen wie ins Krebsregister** vorgestellt. Im Mittelpunkt stand dabei das neue XML-Schema (oBDS-XML Version 3.0.0), das zum 01.03.2022 veröffentlicht wurde. Anhand eines konkreten Beispiels wurde aufgezeigt, welche Informationsobjekte wo in der realen Welt entstehen und wie diese dokumentiert und gemeldet werden. Es wurden gängige Problemfelder und

Herausforderungen vorgestellt und wie man diese mit einem verbesserten XML-Schema lösen bzw. die Ausgangslage verbessern möchte. Auch auf Herausforderungen bei der Best-of-Generierung und bei den Auswertungen wurde schon eingegangen. Als Fazit wurde herausgestellt, dass die Informationskette von der Klinik bis zur Auswertung durchaus komplex ist und es für bessere Ergebnisse wichtig ist, diese auch zu kennen.

Der zweite Beitrag von **Frau Dr. Erb** vom Krebsregister Hamburg handelte von den **Datenendpunkte der Krebsregistrierung, der Best-of-Bildung**. Frau Dr. Erb stellte die bisherigen Ergebnisse der AG Bildung Best-of-Datensatz der Plattform § 65c vor, die im November 2017 gegründet wurde. Es wurde erläutert, warum eine Best-of-Bildung überhaupt notwendig ist. Die AG hat seit der Gründung bereits eine Reihe von Best-of-Empfehlungen erarbeitet, u.a. zu Folgeereignisse und Sterbefälle, zu den Operationen und zur Strahlen- und systemischen Therapie. Trotz der guten Fortschritte auf Plattformebene durch die Arbeit der AG bleibt es eine fortwährende und herausfordernde Aufgabe, die Datenendpunkte in den Registern vergleichbar zu machen.

In der anschließenden Diskussion wurde die Bedeutung eines gemeinsamen Datenmodells des „Best-of-Datensatzes“ noch mal hochgehalten und dafür geworben, sich hier noch mehr und intensiver abzustimmen.

## Fazit

Zum Schluss der Veranstaltung wurden die drei Themengebiete und deren Verknüpfung zueinander noch mal kurz zusammengefasst. Es wurde klar, wie wichtig die Qualität der Qualität der Daten ist, um vernünftige Methode der Analyse auf diese anwenden zu können. Um richtige Schlüsse aus den Analysen zu ziehen, ist ein Wissen über die semantischen Zusammenhänge und wie diese durch die meldungsbasierte Datenerhebung entstehen, von großer Wichtigkeit.

Es wurde angeregt, einen ähnlichen Workshop auch im kommenden Jahr durch die vertretenden Arbeitsgruppen zu organisieren.