

Beschluss

des Innovationsausschusses beim Gemeinsamen Bundesausschuss gemäß § 92b Absatz 3 SGB V zum abgeschlossenen Projekt *SePaMiM* (01VSF20018)

Vom 20. Juni 2025

Der Innovationsausschuss beim Gemeinsamen Bundesausschuss hat in seiner Sitzung am 20. Juni 2025 zum Projekt *SePaMiM - Sequential Pattern Mining und Pattern Matching von Krankheits- und Behandlungsverläufen für klinische Krebsregister* (01VSF20018) folgenden Beschluss gefasst:

- I. Die im Projekt erzielten Ergebnisse werden mit Blick auf die Potentiale der Weiterentwicklung und Adaption der entwickelten Pattern Matching und Data Mining Werkzeuge zur Information an die (klinischen) Landeskrebsregister, das Zentrum für Krebsregisterdaten (ZfKD) am Robert Koch-Institut, die Arbeitsgemeinschaft Deutscher Tumorzentren (ADT) e. V., die Deutsche Krebsgesellschaft e. V. und an die Stiftung Deutsche Krebshilfe weitergeleitet.

Begründung

Das Projekt *SePaMiM* hat erfolgreich die zwei computergestützten Verfahren des Pattern Matching sowie des Data Mining genutzt, um Auswertungen von Krebsregisterdaten vorzunehmen. Übergeordnetes Ziel war es, anhand dieser zwei Ansätze zu prüfen, ob sich im onkologischen Basisdatensatz Behandlungsverläufe aufbereiten und analysieren lassen, um damit die Aufgaben von klinischen Krebsregistern zu unterstützen. Im Pattern Matching Ansatz sollten Möglichkeiten geprüft werden, auswertende Expertinnen und Experten in die Lage zu versetzen, zeitlich festgelegte Therapie- und Verlaufereignisabfolgen als Suchkriterium vorzugeben, um Patientenkohorten zu selektieren. Im Data Mining Ansatz war es Ziel, durch die Verwendung von Clustering Verfahren die automatische Identifizierung von Kohorten anhand von Verläufen untersuchen zu können. Im Rahmen des Projekts wurden dementsprechend zwei softwarebasierte Auswertungswerkzeuge (Prototypen) entwickelt und unter Verwendung von Datensätzen des Landeskrebsregisters Nordrhein-Westfalen und der Landesauswertungsstelle Niedersachsen geprüft.

Mit dem Pattern Matching Auswertewerkzeug gelang es im ersten Projektteil anhand verschiedener Fragekomplexe, die typischen Aufgaben klinischer Krebsregister abzubilden und zu demonstrieren, dass bereits vorhandene Abfragen und Auswertungen zur Beantwortung dieser Fragestellungen grundsätzlich repliziert werden können. Hierbei wurden klinisch relevante Fragestellungen aus den Bereichen ‚Qualitätsindikatoren für Leitlinien‘, ‚Plausibilitätsprüfung von Patienten- bzw. Behandlungsverläufen‘, ‚Anfragen für Abruf klinischer Krebsregisterdaten für Forschungszwecke externer Forschender‘ und ‚Routineauswertungen für die öffentliche Berichtserstattung‘ berücksichtigt. Es konnte gezeigt werden, dass das entwickelte Werkzeug geeignet ist, die Krebsregister bei der Bildung von Kohorten für unterschiedliche Fragestellungen zu unterstützen. Im zweiten Projektteil wurde mit dem Data Mining Werkzeug gezeigt, dass Krebsregisterdaten explorativ ausgewertet werden können. Dabei wurden Auswertende auch auf

unerwartete Zusammenhänge von Überlebenswahrscheinlichkeit und Therapieverlauf aufmerksam gemacht, die jedoch nicht abschließend inhaltlich geprüft wurden. Die Evaluation des Werkzeugs wurde in drei Schritten durchgeführt. Zunächst erfolgte eine technische Evaluation, die zeigte, dass die Ergebnisse des vorgenommenen Clusterings plausibel sind. Im zweiten Schritt erfolgte die Vorstellung und Diskussion der Ergebnisse mit Fachexpertinnen und -experten des Landeskrebsregisters Nordrhein-Westfalen. Im Ergebnis wurde das Clustering als wertvoller Beitrag zur Unterstützung der Mitarbeitenden der klinischen Krebsregister bewertet. Im dritten Schritt wurden die durch das Clustering ermittelten Kohorten mit Überlebenszeitanalysen weiterberechnet und insgesamt als plausibel bewertet. Beide Werkzeuge arbeiten nur mit aufbereiteten Datensätzen. Für einen bundesweiten Einsatz der beiden Anwendungen braucht es entweder eine Standardisierung innerhalb der Krebsregister oder jeweils angepasste Datenaufbereitungen. Zudem müssten aus den Prototypen nutzerfreundliche Endanwendungen entwickelt werden.

Mit den angewendeten Methoden konnten die technischen Ansätze zur Datenauswertung bearbeitet werden. Die Evaluation diente einer Pilotierung im Sinne einer Anwendungsprüfung mit dem Ziel, Anhaltspunkte für eine erste Eignung und Weiterentwicklung der technischen Ansätze im Kontext von Krebsregisterdaten zu schaffen. Die Methoden waren prinzipiell geeignet für eine erste Anwendung (Pilotierung) und Validierung. Hauptsächliche Limitation war die unspezifische Operationalisierung der Evaluations- und Gütekriterien.

Die Projektergebnisse zeigen ein substanzielles Potential, dass die beiden im Projekt entwickelten Prototypen für Pattern Matching und Data Mining zur Identifikation und Selektion komplexer Behandlungs- und Krankheitsverläufe beitragen können. Jedoch sind über das Projekt hinausgehende Entwicklungsarbeiten notwendig, um einen produktiven Einsatz der Prototypen zu ermöglichen. Vor dem Hintergrund werden die Projektergebnisse zur Information an die oben genannten Adressatinnen und Adressaten weitergeleitet.

- II. Dieser Beschluss sowie der Ergebnisbericht des Projekts *SePaMiM* werden auf der Internetseite des Innovationsausschusses beim Gemeinsamen Bundesausschuss unter www.innovationsfonds.g-ba.de veröffentlicht.
- III. Der Innovationsausschuss beauftragt seine Geschäftsstelle mit der Weiterleitung der gewonnenen Erkenntnisse des Projekts *SePaMiM* an die unter I. genannten Institutionen.

Berlin, den 20. Juni 2025

Innovationsausschuss beim Gemeinsamen Bundesausschuss
gemäß § 92b SGB V
Der Vorsitzende

Prof. Hecken