

Datenschätze gemeinsam heben

Die aktuelle BMBF-Förderrichtlinie im Rahmen der Dekade gegen Krebs unterstützt neue Ansätze der Datenanalyse, um Forschungsrelevantes besser aus existenten Daten zu filtern und zu nutzen. Auch die Kultur des Datenteilens soll damit beflügelt werden.

Die moderne Krebsbehandlung gründet genau wie die Krebsforschung auf Daten. Neue Krebstherapien zielen auf die individuellen Eigenschaften von Tumoren ab, die durch eine gezielte Analyse des Tumorgewebes aufgedeckt werden können. Die dabei anfallenden so genannten „Omics“-Daten sind sehr umfangreich und komplex.

Dazu kommen z. B. digitale Aufnahmen aus modernen Bildgebungsverfahren, die aufgrund der benötigten Qualität ebenfalls ein hohes Datenvolumen haben. Ärztinnen und Ärzte erfassen über Jahre hinweg Verlaufsdaten ihrer Patientinnen und Patienten, die u. a. an die klinischen Krebsregister übermittelt werden. Die Krebsforschung generiert ebenfalls große Datensammlungen, insbesondere in Studien.

Impuls zur Ausschöpfung der Potentiale im Rahmen der Nationalen Dekade gegen Krebs

Das Bundesministerium für Bildung und Forschung (BMBF) schreibt im Rahmen der Nationalen Dekade gegen Krebs eine „Richtlinie zur Förderung von interdisziplinären Projekten zur Entwicklung und Erprobung von neuen Ansätzen der Datenanalyse und des Datenteilens in der Krebsforschung“ aus. Ziel ist es, Wissenschaftlerinnen und Wissenschaftlern aus dem Bereich der Datenanalyse einen niederschweligen Zugang zu den hochqualitativen Daten aus der translationalen, biomedizinischen Krebsforschung und der onkologischen Routineversorgung zu ermöglichen. Wissenschaftlerinnen und Wissenschaftler aus den Bereichen der Datengewinnung und Datenanalyse sollen eng zusammenarbeiten und miteinander eine klinisch relevante onkologische Frage beantworten.

Warum ist eine Kultur des Datenteilens wichtig?

In der Praxis kommen bislang Datenanalysierende oftmals noch zu selten mit Forschenden aus den Disziplinen zusammen, die in ihrer Arbeit Daten erheben. Das führt zum einen dazu, dass erstere nach Daten suchen müssen, um ihre Modelle zu testen, und nicht wissen, wie sie Zugang dazu bekommen. Auf der anderen Seite drohen beispielsweise in Studien erhobene Datensätze nach Erstpublikation brachzuliegen, könnten jedoch noch zu einem deutlichen Erkenntnisgewinn beitragen, wenn sie mittels neuer Methodik oder neuer Fragestellung re-analysiert würden. Daher ist ein weiteres wichtiges Ziel der Förderung, Forschende des einen Forschungsbereichs mit denen des anderen zusammen zu bringen und die Kultur des Datenteilens für Forschungszwecke zu fördern.

Förderziel und Zwecksetzung

Die geförderten Vorhaben sollen ein herausforderndes Problem der Datenanalyse im Zusammenhang mit einer klinisch relevanten onkologischen Fragestellung adressieren.

Die Vorhaben sollen mindestens einem der folgenden drei Module zugeordnet werden:

Modul 1: Data-Challenge Projekte und Workshops

Hier werden Projekte gefördert, die die Durchführung mehrtägiger Workshops in fokussierter und intensiver Arbeitsatmosphäre einschließlich mehrmonatiger Vor- und Nachbereitung umfassen. Dabei soll eine Gruppe von Wissenschaftlerinnen und Wissenschaftlern eine Forschungsfrage vorab selbst definieren und mit vorhandenen und kompatibel gemachten Daten in einem Workshop analysieren. Mögliche Formate sind Datenanalyse-Wettbewerbe (Data-Challenges), bei denen eine vorgegebene klinisch relevante Datenanalyseaufgabe mittels statistischer, bioinformatischer oder KI-Methoden von konkurrierenden Teams bearbeitet und den Teams notwendige Datensätze für eine begrenzte Zeit zur Verfügung gestellt werden.

Modul 2: Projekte des förderierten Lernens

Ein oder mehrere dieser Aspekte sollen in Modul 1 und 2 adressiert werden:

- Entwicklung neuer Software oder Anpassung von vorhandener Software auf die mögliche Anwendung für die gegebene Fragestellung;
- Entwicklung von KI-Lösungen, um heterogene oder dezentrale Daten verknüpfen und bearbeiten zu können;
- Entwicklung oder Anpassung von Algorithmen für spezifische Fragestellungen, z. B. um Behandlungskonzepte oder das Design von klinischen Studien zu optimieren; auch unter Berücksichtigung von weiteren Faktoren wie der PatientReportedOutcomes (PRO).

Modul 3: Erstellung von qualitätsgesicherten Trainings-, Validierungs- und Testdatensätzen

Um die Entwicklung und Optimierung von Methoden und Algorithmen der Datenanalyse – insbesondere von KI-Methoden – voranzubringen, ist die Etablierung von qualitätsgesicherten Trainings-, Validierungs- und Testdatensätzen erforderlich. Werden diese geeignet aufbereitet und annotiert der wissenschaftlichen Gemeinschaft dauerhaft zur Verfügung gestellt, so können Datenanalysemethoden verglichen und verbesserte Methoden entwickelt werden. In Projekten des Moduls 3 sollen daher qualitätsgesicherte, annotierte und datenschutzkonforme Trainings-, Validierungs- und Testdatensätze erstellt werden, die auf realen medizinischen Daten basieren. Eine klinische Relevanz der Datenanalyseaufgabe sowie eine enge Zusammenarbeit von Datengenerierenden mit Datenanalysierenden bei der Erstellung der Datensätze ist hier ebenso wichtig wie für die Module 1 und 2. Die im Rahmen des Projekts erstellten Datensätze sollen dann der wissenschaftlichen Gemeinschaft dauerhaft zur Verfügung gestellt werden, um damit zukünftige datengetriebene Forschung und KI-Entwicklungen zu ermöglichen.

Patienteneinbindung

Ein Schwerpunkt in allen Vorhaben und Maßnahmen der Nationalen Dekade gegen Krebs ist die starke und gleichberechtigte Einbindung von Betroffenen. Das BMBF geht hier mit gutem Beispiel voran: Jede Förderung ist an die nachweisliche Einbindung von Patientinnen und Patienten gebunden, sowohl bei der Entwicklung der Ausschreibungen als auch in den geförderten Projekten selbst. Menschen, die wissen, wie es ist mit Krebs zu leben, sollen die Möglichkeit haben, sich am Forschungsprozess zu beteiligen. Durch ihre Sichtweise können sie Forschenden helfen, die entscheidenden Inhalte zum Wohle und Nutzen Betroffener zu extrahieren.

Weitere Informationen unter: <https://www.gesundheitsforschung-bmbf.de/de/13939.php>