

Monika Pobiruchin  
Dr. sc. hum.

## **Automatisierte Modellerstellung zur computergestützten Simulation von Erkrankungen unter besonderer Berücksichtigung einer gesundheitsökonomischen Verwendung am Beispiel eines klinischen Tumorregisters**

Fach/Einrichtung: Medizinische Biometrie u. Informatik  
Doktorvater: Prof. Dr. sc. hum. Meinhard Kieser

Der Einsatz von Modellen in der Gesundheitsökonomie ist gerade bei chronischen Erkrankungen, deren Gesamtverlauf selten oder gar nicht vollständig dokumentiert ist, von großer Bedeutung. Vor allem mangels Daten werden Erkrankungsmodelle in der Regel auf Basis von Ergebnissen klinischer Studien erstellt. Diese unterliegen üblicherweise strengen methodischen Vorgaben und die dort beobachteten Effekte können nicht ungeprüft auf den Routinebetrieb, heterogene Patientengruppen und die Normalbevölkerung übertragen werden. Daher ist es sinnvoll, wenn gesundheitsökonomische Experten auf sog. „Real-World-Daten“ zugreifen könnten.

Bösartige Neubildungen sind eine der Haupttodesursachen in westlichen Ländern. Die Erfassung der Krebserkrankungen geschieht u. a. in spezialisierten Dokumentationszentren: (a) den epidemiologischen bzw. bevölkerungsbezogenen und (b) den klinischen Krebsregistern. Vor allem die klinischen Register gelten aufgrund der sehr breiten und longitudinalen Dokumentation zu jedem Erkrankungsverlauf als ein wertvoller „Datenschatz“.

Im Kontext der vorliegenden Arbeit diente die Datenbank des klinischen Krebsregisters des Tumorzentrums Heilbronn-Franken als Datenbasis. Um die hierin dokumentierten Fallverläufe zur Erzeugung eines Markovmodells zu nutzen, wurde ein generischer, vierschrittiger Prozess (Definition - Selektion - Transformation - Generierung) entwickelt. Die im Rahmen der Arbeit entwickelte Analysesoftware MaMMoG (Mamma Markov-Modellierung) implementiert diesen Generierungsprozess und leitet in halb-automatisierter Weise Erkrankungsmodelle ab.

Halb-automatisiert bedeutet hierbei, dass die Software dem Anwender im ersten Schritt („Definition“) mittels Clusterverfahren Vorschläge für sinnvolle und häufig im Datensatz des Tumorzentrums existente Parameterkombinationen angezeigt. Diese können auf einen Erkrankungszustand hinweisen. Die finale Selektion der relevanten Parameter obliegt dem Benutzer. Die weiteren Schritte werden daraufhin automatisiert durchlaufen. Die erzeugte Modellstruktur besteht aus Erkrankungszuständen, Übergängen und Übergangswahrscheinlichkeiten. Der Modellvorschlag kann exportiert und anschließend bspw. in einem Tabellenkalkulationsprogramm weiterverarbeitet werden, um Kostendaten und Nutzwerte zu ergänzen oder zu modifizieren. Eine derartige semi-automatisierte Erstellung von Markovmodellen verkürzt zum einen den Entwicklungsprozess von gesundheitsökonomischen Modellen, zum anderen basieren die Modelle damit auf Routinedaten.

Der Prozess und die damit entwickelten Modelle wurden zweistufig validiert:

1. Externe Validierung: Ein Referenzmodell aus der Literatur, welches zuvor in der Evaluation von HER2-Screeningverfahren und Trastuzumab-Behandlung (eine zielgerichtete Therapie bei Brustkrebs) eingesetzt worden ist, wurde mit lokalen Daten (n=414) und den o. g. Werkzeugen nachgebildet. Die Struktur und abgeleiteten Wahrscheinlichkeiten wurden gegeneinander verifiziert.
2. Technische Validierung des Clusterverfahrens: Patienten aus einer Kohorte, die im ersten Validierungsschritt eingeschlossen war (Kolorektalkarzinom, n=3.514), wurden in Bezug auf 13 Merkmale geclustert. Die resultierenden Cluster wurden hinsichtlich der vorgenommenen Behandlung einander gegenübergestellt und mit den Empfehlungen der korrespondierenden S3-Leitlinie verglichen.

Die Struktur der in Validierungsphase 1 generierten Modelle ähnelte der des Referenzmodells aus der Literatur: Die Übergangswahrscheinlichkeiten wiesen zwar Unterschiede auf, diese waren jedoch anhand des Generierungsprozesses und der zugrunde gelegten Daten erklärbar. Teilweise waren die abgeleiteten Übergangswahrscheinlichkeiten im Vergleich zu den im Referenzmodell getroffenen Annahmen als realistischer anzusehen, beispielsweise beim Modellieren des Auftretens von Rezidiven. Insgesamt konnte jedoch der bereits in internationalen RCTs belegte positive Effekt der Trastuzumab-Behandlung bei HER2-positiven Frauen in einem Routinesetting mittels Krebsregisterdaten nachgewiesen werden. Der Autorin ist bis dato keine vergleichbare Evaluation für das deutsche Gesundheitswesen bekannt.

Das in der Software eingesetzte Clusterverfahren ermöglichte es, Patientenkohorten zu definieren, deren Staging-Eigenschaften denen der international anerkannten UICC-Stadien glichen. Ebenso konnten Leitlinienempfehlungen in Bezug auf die Therapie des Kolon- und Rektumkarzinoms in den Clustern abgebildet werden. Cluster, die sich als nicht leitlinienkonform herausstellten, wurden gesondert untersucht. In diesem Zusammenhang konnten Subkohorten identifiziert werden, deren optimale Behandlung von Experten in der S3-Leitlinie kontrovers diskutiert werden.

**Schlussfolgerung:** Die Arbeit hat gezeigt, dass Daten aus einem klinischen Krebsregister für die Modellierung in der Gesundheitsökonomie eingesetzt werden können. Für die Ableitung valider Modelle ist eine hohe Dokumentationsqualität der zugrunde gelegten Daten notwendig. Ebenso hat die Fallzahl der in das Markovmodell eingehenden Patientenkohorte einen erheblichen Einfluss auf die realitätskonforme Abbildung der errechneten Übergangswahrscheinlichkeiten.

Die Software MaMMoG kann Ärzte und Experten anderer Disziplinen unterstützen, Daten zu erschließen, Behandlungseffekte darzustellen und neue Hypothesen zu generieren. Die Verwendung von Routinedaten in der Gesundheitsökonomie ist ein vielversprechender Ansatz, dessen Potenzial im Rahmen von weiteren Forschungsaktivitäten untersucht werden sollte.