

# Big Data in der Krankenversicherung: Prozesse und Prävention optimieren

**Die Krankenversicherung – insbesondere die gesetzliche und die Vollversicherung in der PKV – ist eine Versicherungssparte mit hoher Schadenfrequenz. Mehr als 90 Prozent der deutschen Bevölkerung hat mindestens einmal im Jahr Kontakt zur ambulanten ärztlichen Versorgung, laut Barmer GEK sind es im Durchschnitt jährlich sogar 15 Arztkontakte pro Person. Im Verlauf mehrerer Jahre entstehen durch diese Versicherungsleistungen große Datenbestände. Entsprechend liegt der Gedanke nah, diesbezüglich bereits von Big Data zu sprechen. Aktuarien haben darauf einen anderen Blick.**

Mit Big Data werden üblicherweise große Datenmengen bezeichnet, die:

- ein immenses Datenvolumen umfassen,
- viele verschiedene Datentypen enthalten,
- kaum oder nicht strukturiert sind sowie
- mit sehr hoher Frequenz auftreten,

sodass sie mit traditionellen Methoden der Datenverarbeitung kaum auszuwerten sind. Im Zusammenhang mit Big Data werden deshalb oftmals neue Methoden der Datenauswertung unter Begriffen wie Data Analytics, Advanced Analytics, Predictive Modelling oder auch Maschinelles Lernen diskutiert.

Trotz der großen Datenbestände in der Krankenversicherung würden die bisherigen Krankenversicherungsdaten wohl nicht dieser allgemeinen Definition von Big Data entsprechen. Denn in der Krankenversicherung werden zum einen keine Daten in Echtzeit generiert, anders als bei Wetter- oder GPS-Daten. Zum anderen sind die Datenvolumina zwar groß, aber eben nicht immens. Des Weiteren sind Krankenversicherungsdaten strukturiert und liegen im Allgemeinen in wohldefinierten Formaten vor.

Gesundheitsdaten in großem Umfang und sehr hoher Frequenz entstehen hingegen zum Beispiel bei der Nutzung von Fitness- oder Gesundheitstrackern, sogenannten Wearables. Die dabei gesammelten Daten zum Gesundheitszustand sowie gegebenenfalls zu Medikation und Therapien könnten eher als Big Data angesehen werden.

## Strenge regulatorische Vorgaben

Für die Aktuarien stellt sich die Frage, ob und wie diese zusätzlich gewonnenen Daten in der Produkt- und Preis-

gestaltung der Privaten Krankenversicherung berücksichtigt werden können. Dabei ist aber zunächst einmal zu beachten: Die private Vollversicherung und viele Krankenzusatzversicherungen werden hierzulande nach Art der Lebensversicherung kalkuliert, sodass für diese Tarife der aufsichtsrechtliche Rahmen sehr eng gefasst ist. Preisdifferenzierungen, beispielsweise durch die Gewährung von Beitragsrabatten innerhalb eines Tarifs, führen zu getrennten Kollektiven und damit zu getrennten Tarifen.

Aber auch bei einer ausreichenden Datengrundlage, die eine preisliche Differenzierung erlauben würde, sind nachteilige Bewertungen des Versicherten im Verlauf der Versicherung zum Beispiel aufgrund einer Verschlechterung seiner Körper- oder Fitnesswerte rechtlich unzulässig. Aufgrund dieser regulatorischen Vorgaben werden in der nach Art der Lebensversicherung kalkulierten Krankenversicherung derzeit kaum Ansätze für eine preisliche Differenzierung durch die Nutzung zusätzlicher Daten gesehen.

Eine andere Möglichkeit der Differenzierung bietet sich in der unterschiedlichen Ausgestaltung der Versicherungsleistungen. Die Differenzierungen müssen sich dabei auf typische Krankenversicherungsleistungen beziehen. Als ein entsprechendes Beispiel sei das Bonusheft der Zahnabsicherung in der Gesetzlichen Krankenversicherung genannt, mit dem alle Versicherten, die über das Bonusheft eine regelmäßige Zahnvorsorge nachweisen, einen höheren Zuschuss beim Zahnersatz erhalten. Solche Leistungsdifferenzierungen wären in der PKV genauso denkbar.

## Genauere Prognosen möglich

Daneben gibt es zahlreiche andere Bereiche in der Krankenversicherung, wo Big Data und Data Analytics eingesetzt werden könnten. Hierzu zählt die Analyse von Leistungsdaten, auch zur Unterstützung im Bestands- und Gesundheitsmanagement oder bei der Prozessautomatisierung. So könnte mithilfe von Data Analytics die Betrugserkennung im Leistungsbereich verbessert werden. Auffällige untypische Muster werden durch spezielle Verfahren erkannt und mit entsprechenden Hinweisen zur detaillierten Prüfung an den Sachbearbeiter weitergegeben. Der Abgleich eingereicherter Leistungen mit statistischen „Normverläufen“ kann außerdem sinnvolle Hinweise zur Verletzung vorvertraglicher Anzeigepflichten geben.

Typischerweise werden aktuarielle Kalkulationen auf der Basis aggregierter Daten gleichartiger Risiken durchgeführt. Moderne Data-Analytics-Methoden erlauben gleichwohl Vorhersagen zukünftiger Erkrankungen und Leistungsbeträge für jeden einzelnen Versicherten auf Personenebene. Dies könnte insgesamt zu genaueren Leistungsprognosen und damit zu einer verbesserten Einschätzung der künftigen Risikolage des Versicherers führen.

### Leistungsmanagement verbessern

Durch die personenbezogene Analyse von Krankheitsverläufen auf Basis von Data Analytics könnten außerdem die Prozesse im Leistungsmanagement optimiert und der einzelne Versicherte stärker als bisher individuell betreut werden. Im Fall einer erhöhten Wahrscheinlichkeit für den Eintritt einer ernsthaften Erkrankung könnte der Versicherte seinem Bedarf entsprechend Hinweise und Angebote zur Unterstützung erhalten. Im Krankheitsfall könnte die individuelle Versorgung des Versicherten zum Beispiel durch ein besseres Zusammenspiel der Fachärzte optimiert und somit das Auftreten von Folgeerkrankungen vermieden oder hinausgezögert werden. Diese individuelle Betreuung kann über digitale Services allen Versicherten zugutekommen und würde nicht erst im fortgeschrittenen Stadium oder bei einer chronischen Erkrankung greifen.

Big Data könnte auch Präventionsmaßnahmen unterstützen. Denkbar wäre zum Beispiel die Erfassung und Bewertung verschiedener Gesundheitsdaten, sportlicher Aktivitäten und der Ernährung. Der Versicherte könnte speziell auf seine Gesundheitssituation abgestimmte Hinweise und Vorschläge, aber auch Erinnerungen an Vorsorgetermine, Medikamenteneinnahme oder andere individuelle Unterstützung erhalten.

Bei allen Anwendungsfällen und generell bei Datenanalysen sind aber grundsätzlich die datenschutzrechtlichen Regelungen zu beachten. Dies gilt in der Regel bereits vor der eigentlichen Analyse, um sicherzustellen, dass diese selbst datenschutzrechtlich unbedenklich ist. Denn das Erheben, Verarbeiten und Nutzen von personenbezogenen Daten und vor allem von Gesundheitsdaten ist nur dann zulässig, wenn der Betroffene einwilligt oder ein gesetzlich definierter Ausnahmefall vorliegt.

### Datenaufbereitung entscheidender Faktor

Eine der größten Herausforderungen beim Einsatz von Big Data wird auch künftig die Datenaufbereitung sein, die derzeit meistens 80 bis 90 Prozent der Projektzeit in Anspruch nimmt und in Anbetracht der Masse und Vielfalt der Daten immer komplexer wird. Ziel der Datenaufbe-

ereitung ist es, den Datensatz optimal für die zu verwendenden Data-Analytics-Methoden vorzubereiten. Die Datenaufbereitung spielt damit eine entscheidende Rolle in jeder Analyse. Denn Fehler, die an dieser Stelle gemacht werden, wirken sich unmittelbar auf das Resultat der Analyse aus.

Die Wahl der Data-Analytics-Methode selbst hängt stark von der Fragestellung der Analyse ab. Geht es zum Beispiel um die Vorhersage, ob eine Person mit hoher Wahrscheinlichkeit in den nächsten zwei Jahren in ein Krankenhaus eingewiesen wird, so sind Methoden aus dem Bereich der sogenannten Ereignisprognosen besonders geeignet. Dies sind beispielsweise Entscheidungsbäume, neuronale Netze oder auch die logistische Regression. Geht es hingegen um metrische Vorhersagen, zum Beispiel zu den Leistungsausgaben für eine vollversicherte Person im nächsten Jahr, so kommen dabei oft lineare Regressionsmodelle oder auch Generalisierte Lineare Modelle (GLM) zum Einsatz.

Es gibt jedoch eine Vielzahl von weiteren Methoden, die den Aktuaren zur Verfügung stehen. Neben der Frage der Zielrichtung der Prognose spielt zum Beispiel auch eine Rolle, ob häufige oder seltene Ereignisse vorhergesagt werden sollen, wie gut die Methoden mit sogenannten missing values – fehlenden Variablen – umgehen, wie stabil die Ergebnisse bei Datenaktualisierungen sind oder auch, ob die Methode eine niedrige oder hohe Rechenkapazität verlangt.

### Fazit

#### Data-Analytics-Methoden erweitern aktuariellen Werkzeugkasten

Aktuare sind geübt im Umgang mit Versicherungsdaten und arbeiten originär mit Wahrscheinlichkeiten und Vorhersagemodellen. Moderne Data-Analytics-Methoden ergänzen die klassischen aktuariellen Methoden und ermöglichen dem Aktuar, die versicherungstechnischen Risiken noch besser einzuschätzen. Aktuare können mit Data Analytics die Entwicklung neuartiger Ansätze im Versicherungsbetrieb, wie die Betrugserkennung oder individuelle Betreuung der Versicherten, unterstützen. Aktuare leisten damit einen wichtigen Beitrag beim Wandel zu einem digitalen Krankenversicherer.