

Datenwolken in der Medizin



Rechenzentren für gigantische Datenmengen: Supercomputer helfen Wissenschaftlern bei der Analyse medizinischer Daten. Mit ihrer Hilfe lassen sich bisher unbekannte Zusammenhänge offenbaren.

In der Datenflut, die Forschung und moderne Medizin generieren, verbirgt sich Wertvolles: bisher unerkannte Zusammenhänge zwischen verschiedenen Mutationen oder anderen Biomarkern und dem Therapieerfolg eines Medikaments. So könnte beispielsweise eine bestimmte Kombination solcher Genvarianten – sogenannter Allele – die Chance erhöhen, auf ein Krebsmedikament gut anzusprechen. „Wäre so eine Anlage bekannt, könnten Ärzte gezielt die richtige Therapie für den jeweiligen Patienten auswählen“, sagt Dr. Joerg Lippert, Leiter Clinical Pharmacometrics bei der Bayer-Division Pharmaceuticals. Doch dieses Wissen gilt es zuerst aus den Daten zu destillieren. „Wir müssen alle Informationen nutzen, damit wir in der modernen Medizin möglichst optimale Entscheidungen treffen können“, erklärt Lippert.

Die Wissenschaftler müssen unstrukturierte medizinische Daten zunächst nutzbar machen

Schon heute gibt es eine fast unüberschaubare Menge medizinischer Daten – aus der Routine-Diagnostik, aber auch aus medizinischen Studien. „Big Data ist in der Medizin schon lange Realität: Die drei V – volume, velocity und variety – werden zusehends den Alltag von Ärzten bestimmen“, sagt Lippert. Dabei steht „volume“ für die Datenmenge und „velocity“ für die

Geschwindigkeit, mit der die Daten generiert werden. Eine besondere Herausforderung medizinischer Daten ist deren Komplexität – die „variety“. So enthält eine Patientenakte beispielsweise Diagnosebilder zusammen mit Werten und Tabellen – also Ergebnisse unterschiedlichster Untersuchungsmethoden. „Meist sind medizinische Daten derzeit noch unstrukturiert und teilweise fehlerhaft. Daher müssen wir die Daten manuell oder mithilfe spezieller Algorithmen im Computer in eine Form überführen, in der ein Rechner mit ihnen weiterarbeiten kann“, erklärt Lippert. Dann setzen die Forscher voll auf ihre Hochleistungsrechner. „Wir lassen die Daten für sich sprechen und beginnen deshalb mit möglichst wenigen Annahmen. So schränken wir den Erwartungsraum unserer Analysen kaum ein – die Ergebnisse sind nahezu völlig offen. Dadurch kommen wir zu neuen Hypothesen“, sagt Lippert. Wegen der großen Datenmengen erfordert dieser Ansatz eine hohe Rechenleistung. „Wenn wir einen Datensatz mit 50.000 Patienten nehmen, von denen wir jeweils 5.000 Gesundheitsparameter kennen, dann gibt es eine astronomisch hohe Anzahl von Kombinationen, die wir untersuchen könnten“, erklärt der Bayer-Forscher. Deshalb setzen die Forscher auf sogenannte Heuristiken: spezielle Methoden, die helfen, Daten zu strukturieren, ohne dabei zu viele Annahmen treffen zu müssen. Ein Kompromiss, damit die komplexen Rechenoperationen in einer



Zahlen, Formeln und Muster: Dr. Joerg Lippert und sein Team suchen in großen medizinischen Datensätzen nach neuen Zusammenhängen. Diese sollen Ärzten helfen, gute Entscheidungen zu treffen.

vertretbaren Zeit abgeschlossen sind. „Wir haben riesige Datensätze, in denen wir nach statistischen Zusammenhängen suchen: Das setzt maschinelles Lernen voraus, ein Ansatz, den wir schon seit Jahren verfolgen. Die wesentliche Neuerung – heute sind die Datensätze größer, die Rechner schneller und das führt zu einer neuen Qualität“, betont Lippert.

So können die Datenspezialisten schon heute mithilfe ihrer Rechner die optimale Dosierung von Arzneimitteln abschätzen. Eine Innovation, die vor allem bei der Planung klinischer Studien sehr wertvoll ist. „Mit den richtigen Daten spart uns das mehrere Jahre Entwicklungszeit. Das hilft den Patienten, da wir ihnen eine neue Therapieoption schneller zur Verfügung stellen können“, fasst Lippert zusammen. Die neuen Ansätze, an denen er mit seinem Team arbeitet, ermöglichten beispielsweise das Überspringen eines Studienabschnitts in der Phase II der klinischen Entwicklung eines Medikaments gegen Herzinsuffizienz. Sie konnten dadurch über ein Jahr Entwicklungszeit einsparen. Die Datenexperten stehen erst am Anfang und Lippert ist überzeugt: „Wir können die Zukunft der Medizin mitgestalten und für bessere Therapien sorgen.“ ■

„Big Data: große rechtliche Herausforderungen“

„research“ sprach mit Sigrid Achenbach, Rechtsanwältin der Division Pharmaceuticals bei Bayer, über rechtliche und gesellschaftspolitische Herausforderungen von Big-Data-Analysen in der Medizin.



Sigrid Achenbach



Welche rechtlichen Fragen müssen geklärt werden?

Speziell für Unternehmen, die neue Medikamente entwickeln, sind Analysen mit krankheitsbezogenen Patienteninformationen extrem interessant. Diese kollidieren allerdings potenziell mit drei Prinzipien des Datenschutz-Gesetzes: Zunächst muss die Datenverarbeitung grundsätzlich erlaubt sein – Prinzip der Rechtmäßigkeit – und möglichst wenig persönliche Daten einschließen – Prinzip der Datenminimierung. Außerdem dürfen die persönlichen Informationen nur in der vereinbarten Studie verwendet werden – Prinzip der Zweckbindung. In der Praxis ist es sehr schwierig, alle Vorgaben in Big-Data-Studien zu erfüllen.

Eine Herausforderung. Gibt es hierzu eine Lösung?

Es gibt heute noch keine Universallösung, meist sind komplizierte Fall-zu-Fall-Entscheidungen nötig. Selbst die neue Allgemeine Datenschutzregelung der EU, die im Mai 2018 in Kraft tritt, kann diese Schwierigkeiten voraussichtlich nicht beheben. Eine mögliche Lösung wäre eine breitere Patientenzustimmung, die Big-Data-Analysen zulässt. Dafür gibt es allerdings gesetzliche Grenzen. Eine weitere Möglichkeit wäre ein unabhängiger Datenschutz-Treuhänder, der Forschungsvorhaben, die über

die ursprüngliche Vereinbarung hinausgehen, überprüfen und genehmigen könnte – bislang eine Vision. Diese Herausforderung kann nur von allen Interessensgruppen gemeinsam bewältigt werden: Das sind die verantwortlichen Ministerien und Pharmaunternehmen, aber auch akademische Institutionen und Patientenorganisationen: Ein breiter gesellschaftlicher Diskurs ist erforderlich.

Wie wird sich die Medizin in den nächsten 20 Jahren hier weiter entwickeln?

Die Forschung wird zunehmend Daten aus unterschiedlichen Quellen nutzen und komplexe Analysen durchführen. Hierdurch werden neue Erkenntnisse gewonnen werden, wodurch Patienten und der Gesellschaft als Ganzes geholfen wird. Dennoch wird es Sicherheitsmechanismen geben, die einen Missbrauch der sensiblen Daten verhindern: Damit einzelne Patienten nicht identifiziert werden können oder medizinische Informationen Dritten in die Hände fallen. Mir persönlich ist wichtig, dass die Daten nur auf freiwilliger Basis verwendet werden – jeder soll frei entscheiden dürfen, was mit seinen persönlichen Daten passiert.