gesundhyte.de

DAS MAGAZIN FÜR DIGITALE GESUNDHEIT IN DEUTSCHLAND

AUSGABE 13 DEZEMBER 2020

spezial: künstliche intelligenz

ab Seite 8

potenziale und herausforderungen von KI in der medizin

Seite 15

big data und smartphones auf der intensivstation

Seite 30

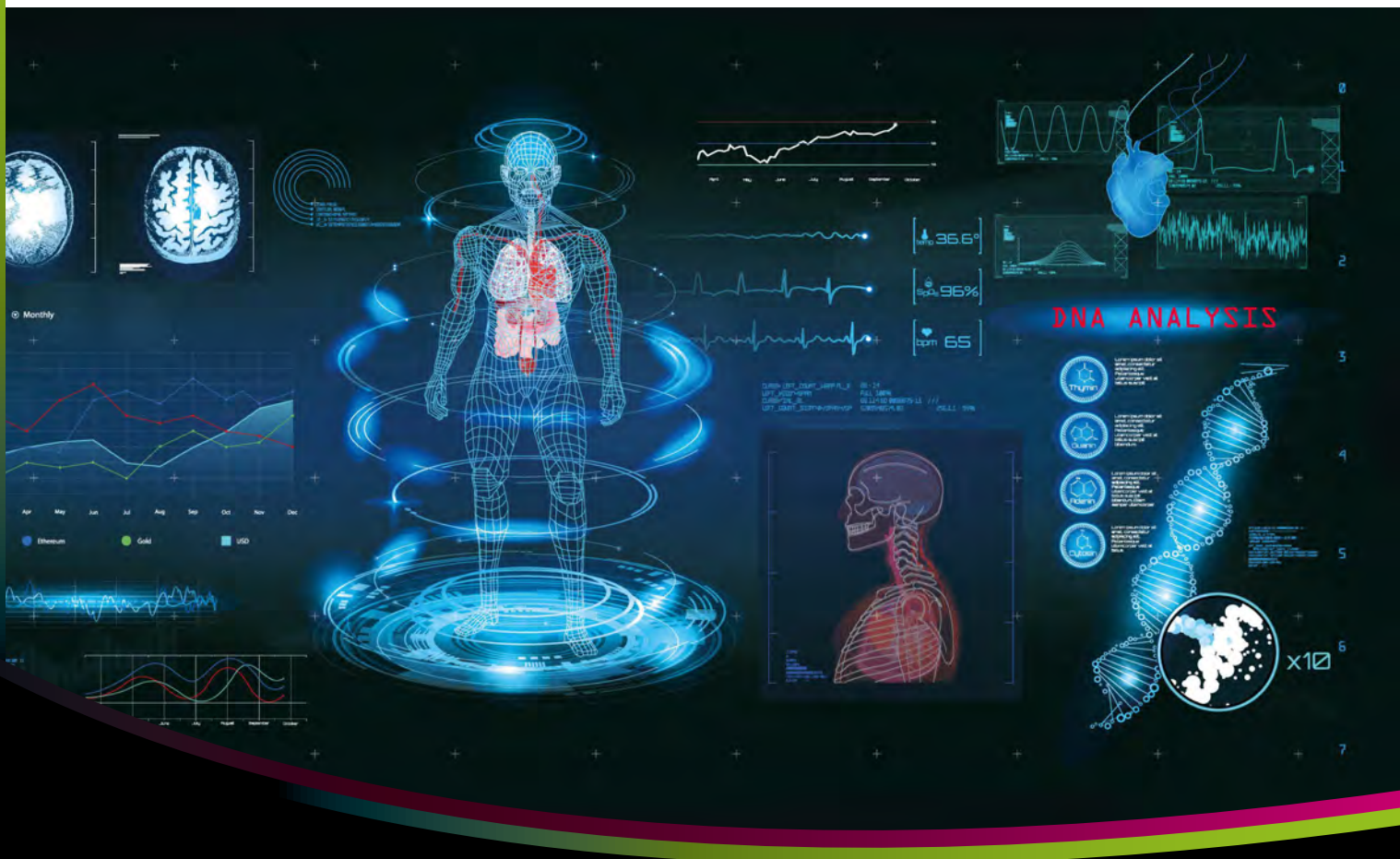
interviews mit Petra Ritter Janine Felden

Seite 52 und 72

ISSN 2702-2544

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung



gesundhyte.de

widmet sich im Zeitalter der Digitalisierung den wichtigsten Aspekten rund um die digitale Gesundheit und die Forschung im Bereich der digitalen Medizin. Sie gilt als Medizin der Zukunft: Innovative Methoden und Technologien in Kombination mit verschiedenen digitalen Daten aus dem Labor und vom Krankenbett ermöglichen präzisere Vorhersagen und personalisierte Therapiemöglichkeiten. In der Krebstherapie werden diese Methoden bereits erfolgreich eingesetzt, in Zukunft werden sie sich zusehends auch auf andere Krankheitsbilder erweitern. Hierbei spielt vor allem auch das Zusammenführen von verschiedenen Daten in der Klinik und aus der ambulanten Versorgung eine Rolle. Erfahren Sie im Magazin gesundhyte.de wie dieser aufstrebende Forschungszweig arbeitet und welche Lösungsansätze er auf die aktuellsten Herausforderungen bereitstellt.

Titelbild: iStock © peepo

Foto: AdobeStock_© SergeyBitos

grußwort

Liebe Leserinnen und Leser,

die COVID-19-Pandemie stellt die Welt vor große Herausforderungen. Sie hat uns aber auch noch einmal verdeutlicht, wie wichtig Forschung ist. Weltweit setzen Menschen ihre Hoffnung in die schnelle Entwicklung neuer Diagnose-, Therapie- und Impfverfahren. Gleichzeitig zeigen uns die enormen Anstrengungen, wie anspruchsvoll die Realisierung dieser Ziele ist. Der menschliche Körper ist keine Maschine, sondern ein komplexes Zusammenspiel von unglaublich vielen Faktoren, von denen noch lange nicht alle verstanden sind.

So reagiert jeder Mensch anders auf eine Infektion oder andere Erkrankungen. Für einige Patientinnen und Patienten stellt das neuartige Coronavirus eine lebensbedrohliche Erkrankung dar, die eine intensivmedizinische Behandlung nötig macht. Andere spüren nicht viel mehr als ein leichtes Kratzen im Hals. Was unterscheidet diese Menschen? Wie kann eine personalisierte Behandlung aussehen, die diese Faktoren berücksichtigt? Wo greift das Virus an und wie kann man es gezielt bekämpfen?

Wer Antworten auf diese Fragen sucht, benötigt eine große Menge an Daten. Dies sind zum einen Forschungsdaten etwa über den Erreger und zum andern medizinische Daten wie die individuellen Krankheitsverläufe und deren Behandlungen. Die Erfassung der Daten ist eine komplizierte und verantwortungsvolle Aufgabe, die sich nur im Zusammenspiel von Forschung und Versorgung bewältigen lässt und welche verlässliche Infrastrukturen braucht. Gemeinsam mit den Bundesministerien für Gesundheit und Wirtschaft haben wir die Innovationsinitiative „Daten für Gesundheit“ ins Leben gerufen, um diese Herausforderung anzugehen.

Auch in der Förderung setzt das Bundesministerium für Bildung und Forschung auf dem Weg zu einer personalisierten Patientenversorgung auf eine verstärkte Vernetzung der Forschung und auf eine vermehrte Digitalisierung in der Medizin. Diese Ansätze unterstützen wir unter anderem im „Deutschen Netzwerk für Bioinformatik-Infrastruktur de.NBI“ und der „Medizininformatik-Initiative“.

In der aktuellen Ausgabe von gesundhyte.de veranschaulichen vielfältige Projekte das breite Anwendungspotential aktueller digitaler Methoden in den Lebenswissenschaften und der Medizin. Ich wünsche Ihnen allen eine anregende Lektüre.

Anja Karliczek
Bundesministerin für Bildung und Forschung

Foto: Laurence Chaperon / BMBF

grußwort

Liebe Leserinnen und Leser,

mit dem Medizin- und Chemie-Nobelpreis wurde dieses Jahr die Hepatitis-Forschung als auch die Entdeckung der Genschere CRISPR/Cas9 ausgezeichnet. CRISPR/Cas9, zunächst eine Entdeckung der Grundlagenforschung, wird in der Zukunft als Methode des Genome Editing, immer weitere Möglichkeiten bieten, Krankheitsbilder durch das Editieren von Genen zu behandeln. Wissenschaftliche Forschung verändert kontinuierlich unsere Diagnostik- und Therapiemöglichkeiten. Ebenso wie sich die Medizin verändert, hat sich die systembiologie.de zu gesundhyte.de entwickelt, um den systembiologischen Charakter noch stärker auf klinische Anwendung, neue Präventionsmaßnahmen und Therapien für Patientinnen und Patienten zu richten.

Die aktuelle Pandemielage hat die Themen Infektionsschutz, Prävention und das Nachverfolgen von Infektionsketten, aber auch Fragen zur ausreichenden Personalausstattung an Kliniken und Pflegeheimen stärker in den Vordergrund gerückt. Wir an der Charité verstehen die aktuelle Covid-19 Pandemie nicht nur als eine große Herausforderung, sondern auch als eine Chance: Für den schnelleren Informationsaustausch und den Austausch von Forschungsdaten gilt es zum Beispiel, die Universitätskliniken besser miteinander zu verknüpfen, Patienten und Unternehmen digital zu ertüchtigen und aufbauend auf den Konsortien der Medizininformatik-Initiative bereits initiierte und laufende Prozesse zu beschleunigen.

Neben dem Schwerpunkt auf Infektionen gehen wir in diesem Jahrzehnt weitere wichtige Fragestellungen an: Insbesondere die Digitalisierung der Medizin, was sowohl das Einführen der elektronischen Patientenakte (ePA) und forschungskompatiblen Patientenakte als auch die personalisierte Patientenversorgung einschließt. Präzisionsmedizin wird nicht nur im Bereich der Onkologie, sondern auch für andere Krankheitsbilder wichtiger und wird sich zur Medizin der Zukunft entwickeln. Besonders in diesem Bereich treibt der Einsatz von Methoden der künstlichen Intelligenz (KI) neue Technologien und Therapien voran. Dabei spielen vor allem computergestützte Entscheidungshilfen - inklusive Methoden des maschinellen Lernens - nicht nur in der Forschung, sondern auch im Krankenhausalltag zunehmend eine Rolle, sei es zur Unterstützung bei Diagnosen, zur Wahl der aussichtsreichsten Therapie oder auch zur Optimierung im Pflegebereich. Es gibt viele Anwendungsmöglichkeiten, wir müssen sie nur zu nutzen wissen. Auch in der COVID-19 Krise birgt die KI enorme Möglichkeiten, die wir ergreifen sollten, ohne die damit verbundenen Risiken außer Acht zu lassen; auf all diese Themen werden Sie in dieser Spezial-Ausgabe „Künstliche Intelligenz“ von gesundhyte.de stoßen.

Ich wünsche Ihnen eine spannende Lektüre!

Prof. Dr. Heyo K. Kroemer
Vorstandvorsitzender der Charité – Universitätsmedizin Berlin

Foto: Wiebke Peitz / Charité

vorwort

Liebe Leserinnen und Leser,

wussten Sie, dass „Zeit“ das meistverwandte Substantiv der deutschen Sprache ist? In unserer schnelllebigen Zeit wird es dabei wahrscheinlich am häufigsten mit der Negativaussage „ich habe keine Zeit“ verbunden. Aber auch für die Andeutung eines Veränderungsprozesses wird es für die Beschreibung des Zustands „im Wandel der Zeit“ verwandt. So befindet sich auch das Feld der Systembiologie in einem ständigen Wandel, der sich in den bislang zwölf Ausgaben von systembiologie.de seit dem Ersterscheinen vor zehn Jahren widerspiegelt.

Nach einer etwas längeren Zeit des Innehaltens haben wir uns entschlossen einen Neuanfang zu wagen. Nachdem die biomedizinische Forschung schon seit langer Zeit ohne digitale Unterstützung nicht mehr denkbar ist, hält die Digitalisierung nun endlich auch mit Nachdruck Einzug in den Gesundheitssektor. Viele Hoffnungen konzentrieren sich dabei auch auf den Bereich der Künstlichen Intelligenz (KI), die zu großen, teils disruptiven Veränderungen im Gesundheitsbereich führen sollen. Die Versprechen sind groß, die Erwartungen gleichermaßen. Einige der Versprechungen werden sich in kurzfristigen Zeitfenstern umsetzen lassen, andere wiederum werden auf sich warten lassen. Spannend und allgegenwärtig ist das Thema allemal und wird uns noch für eine längere Zeit begleiten.

Wir haben uns entschlossen, den Titel des Magazins nach zehn Jahren zu verändern. Aus systembiologie.de wird gesundhyte.de. Diese Ihnen vorliegende Ausgabe können Sie von beiden Seiten betrachten, die Rückseite im alten Kleid, die Vorderseite im neuen. Die von uns gewählten Themen, die gewohnt von einer großen Zahl engagierter Autorinnen und Autoren mit Leben gefüllt wurden, bewegen sich ganz am Puls der Zeit. Natürlich kommt in dieser bewegten Zeit kein Magazin im Gesundheitsbereich ohne Betrachtung der COVID-19 Pandemie aus, so dass auch in der heutigen Ausgabe Themen rund um die Pandemie aus Sicht der KI beleuchtet werden. Denn auch in der laufenden Pandemie zeigt sich deutlich, dass die Digitalisierung erheblich dazu beiträgt, der Pandemie die Stirn zu bieten.

Ein Neuanfang ist immer mit einem Wandel verbunden. So hat sich das Redaktionsteam neu zusammengefunden und wir freuen uns über die zahlreichen neuen Mitglieder, die diesem Magazin neue, wichtige Impulse verliehen haben. Neuer zusätzlicher Herausgeber ist nun das Berliner Institut für Gesundheitsforschung, das sich selbst in einem spannenden Prozess des Wandels befindet. Manche Dinge wiederum haben sich nicht geändert. So bleiben wir weiterhin einer stets unterhaltsamen, aber immer auch fachlich orientierten Berichterstattung verbunden.

Sie haben sicher mitgezählt, das Wort „Zeit“ kommt in meinem Vorwort gleich dreizehnmal vor. Ich wünsche Ihnen auch in diesen schwierigen Zeiten hinreichend Muße und Zeit für die Lektüre dieses spannenden Magazins.

Es grüßt Sie herzlichst Ihr

Roland Eils
Chefredakteur gesundhyte.de

Foto: David Ausserhofer / BIH

inhalt

grußwort 3
Anja Karliczek, Bundesministerin für Bildung und Forschung

grußwort 4
Prof. Dr. Heyo K. Kroemer, Vorstandvorsitzender der Charité – Universitätsmedizin Berlin

vorwort 5
Prof. Dr. Roland Eils, Chefredakteur, Gründungsdirektor BIH-Zentrum für digitale Gesundheit an der Charité

spezial: KI – künstliche intelligenz

KI und ethik in der medizinforschung 8
KI-basierte Forschung in der Medizin und ihre ethischen Herausforderungen
von Olga Levina

IEAI: vorreiter bei der gestaltung ethischer KI 11
Globale, egalitäre und interdisziplinäre ethische Richtlinien für die Entwicklung und den Einsatz von Künstlicher Intelligenz
von Anastasia Aritzi und Caitlin Corrigan im Namen des IEAI

potenziale und herausforderungen von KI in der medizin – 15
Heilungschancen erhöhen, Krankheiten vorbeugen, Selbstbestimmung ermöglichen
von Klemens Budde und Karsten Hiltawsky für die Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme

„herr doktor, verstehen sie mich?“ 19
Wie lernende Systeme helfen medizinische Fachsprache zu verstehen und welche Rolle klinische Leitlinien dabei spielen
von Florian Borchert, Christina Lohr, Luise Modersohn, Udo Hahn, Thomas Langer, Gregor Wenzel, Markus Follmann und Matthieu-P. Schapranow

präzise diagnostik mit künstlicher intelligenz 23
Firmenporträt: Aignostics GmbH
von Stefanie Seltmann

alte datens(ch)ätze zu neuem leben erwecken 26
Künstliche Intelligenz ermöglicht, neue Informationen aus vorhandenen Daten zu gewinnen
von Jakob Simeth, Marian Schön, Michael Huttner, Paul Heinrich, Michael Altenbuchinger und Rainer Spang

big data und smartphones auf der intensivstation 30
Wie Smartphone und Künstliche Intelligenz dabei helfen, beatmete Patienten zu behandeln
von Gernot Marx, Sebastian Fritsch, Johannes Bickenbach, Julian Kunze, Oliver Maaßen, Saskia Deffge, Silke Haferkamp und Andreas Schuppert

#nCoVStats 34
Wie Data Science hilft die Coronavirus-Pandemie zu verstehen
von Matthieu-P. Schapranow

die chronic kidney disease nephrologist’s app 38
Personalisierte Systemmedizin für Patienten mit chronischem Nierenleiden
von Ulla T. Schultheiß, Johannes Raffler, Sahar Ghasemi, Robin Kosch, Michael Altenbuchinger und Helena U. Zacharias

alternatives spleißen aus sicht der systemmedizin 43
Welchen Einfluss hat alternatives Spleißen auf die Entstehung von Krankheiten?
von Tim Kacprowski, Nina Kerstin Wenke, Sabine Ameling, Kristin Wenzel, Olga v. Kalinina und Markus List im Namen des gesamten Sys_CARE Konsortiums

bioinformatik und systemkardiologie 48
Institutsporträt: Klaus Tschira Institute for Integrative Computational Cardiology
von Tobias Jakobi und Christoph Dieterich

das virtuelle gehirn 52
Interview: Petra Ritter
von Stefanie Seltmann

modellaustausch für die regulatorische genomik (MERGE) 56
Förderung der gemeinsamen Nutzung und Wiederverwendung von Vorhersagemodellen in der Genomik durch Software-Standardisierung
von Julien Gagneur, Oliver Stegle, Michael J. Ziller

„bench to bedside“: systempharmakologie in der praxis 60
Firmenporträt: esqLABS GmbH
von Stephan Schaller

der roboter im OP 64
Die Roboter-assistierte Chirurgie (RAC): vom Gimmick zum Standard
von Klaus-Peter Jünemann

meldungen aus dem BMBF 68

bioinformatik: junge frauen brauchen vorbilder 72
Interview: Janine Felden
von Melanie Bergs und Gesa Terstiege

ein motor für die analyse von COVID-19-forschungsdaten 75
Das Deutsche Netzwerk für Bioinformatik-Infrastruktur überzeugt durch Auswerteprogramme und eine Cloud-basierte Rechnerstruktur
von Alfred Pühler, Irena Maus, Vera Ortseifen und Andreas Tauch

de.NBI cloud als akademische lösung für lebenswissenschaftler 80
Cloud Computing bietet flexible skalierbare Rechen- und Speicherressourcen für Big Data Anwendungen
von Christian Lawerenz und Alexander Sczyrba

datenintegrationszentrum – drehscheibe für daten in der medizinischen forschung und versorgung 84
Datenintegration und ihre Voraussetzungen
von Björn Schreiweis, Danny Ammon, Martin Sedlmayr, Fady Albashiti und Thomas Wendt

einzug der massenspektrometrie in die systemmedizin 88
Vier Forschungskerne auf der Suche nach neuen Biomarkern
von Jeroen Krijgsveld, Ursula Klingmüller, Carsten Müller-Tidow, Bernhard Küster, Daniel Teupser, Ulrich Keilholz, Frederick Klauschen, Markus Ralser, Matthias Selbach, Philipp Wild und Stefan Tenzer

modellierung von infektionserkrankungen der atemwege 92
Systemmedizinische Modellierung von Therapieoptionen bei ambulant erworbener Pneumonie und COVID-19
von Peter Ahnert und Markus Scholz

news 98

events 100

impressum 101

wir über uns 102

kontakt 103

KI und ethik in der medizinforschung

KI-basierte Forschung in der Medizin und ihre ethischen Herausforderungen

von Olga Levina

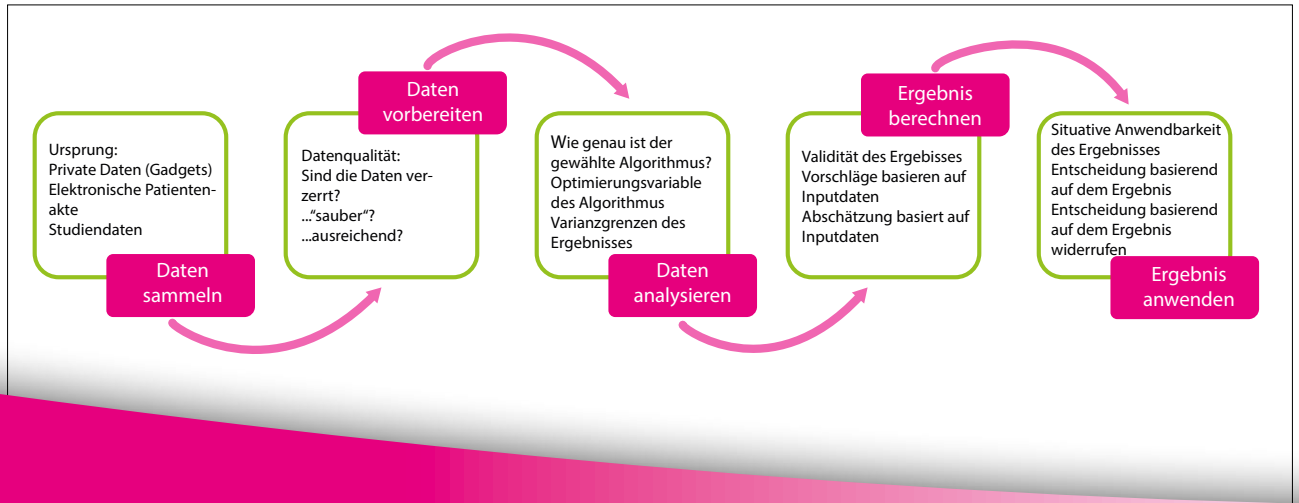
Das Thema Ethik im Kontext von Technologien der Künstlichen Intelligenz (KI) wird aktuell in der Wissenschaft und Öffentlichkeit aber auch in der Politik stark diskutiert. Die diskutierten Technologien sind meist der so genannten „schwachen" KI zuzuordnen. Das sind informationstechnische Systeme, die sich auf das Lösen eines bestimmten Problems mithilfe der Ansätze des Maschinellen Lernens (ML) fokussieren. Entsprechend trainiert und auf einen bestimmten Problembereich bezogen, können durch diese Algorithmen Muster in den Daten bestimmt und kontextbezogen in die Abläufe integriert werden. Als Technologien, die aus einer großen Menge an Daten Muster erkennen können, eignen sich die ML-basierten Technologien sehr gut dazu, medizinische Daten, die häufig komplex sind, zu strukturieren und zu analysieren. Bei ihrem Einsatz im realen Forschungsumfeld müssen jedoch einige Aspekte, die ihre Aussage unmittelbar beeinflussen können, beachtet werden.

Machine-Learning-Systeme in der medizinischen Forschung

Der Einsatz von ML-basierten Systemen in der medizinischen Forschung bedeutet z. B., dass Patient:innen effizienter bestimmten Studien zugeordnet oder Wirkstoffe effizienter identifiziert und kombiniert werden können. Die Auswertung der komplexen Kombination von klinischen Daten, genetischen Informationen sowie Informationen aus den medizinischen Studien kann durch ein ML-basiertes System effizienter durchgeführt werden, um so Aussagen z. B. über die Wirksamkeit von Medikamenten oder Therapien treffen zu können. Dabei gilt es stets den Schutz der verwendeten persönlichen und der sensiblen Daten zu gewährleisten.

Aktuelle Entwicklungen im Bereich der digitalen und mobilen Anwendungen wie z. B. Smart Watches und persönliche Fitnessanwendungen eröffnen für die Forschenden eine Möglichkeit auf eine große Menge von personenbezogenen, verhaltensbasierten Daten, die bisher hauptsächlich aus Selbstauskünften von Patient:innen oder Beobachtungsstudien bezogen werden konnten, zuzugreifen. Für die medizinische Forschung ist dies von Bedeutung, da die Datenmenge für die Ergebnisermittlung von ML-basierten Systemen eine wichtige

Abbildung 1: Gesellschaftliche Fragen im Datenverarbeitungsprozess (nach Levina, 2020)



Dr. Olga Levina, wissenschaftliche Mitarbeiterin am FZI Forschungszentrum Informatik in Berlin (Foto: Susanne Krautwald).

Rolle spielt. Je mehr Daten beim Training des Systems eingesetzt werden, desto genauer ist das ermittelte Ergebnis. Woher die Daten stammen, wie und wann sie gesammelt wurden und im welchem Maße sie sich auf die Forschungsfrage beziehen sind nur einige der Fragen, die hier relevant sind. So diktieren z. B. bei der Erfassung und Auswertung von Daten aus digitalen Anwendungen aktuell die herstellenden IT-Konzerne und nicht die Forschenden sowohl das Format, die Inhalte als auch den Zugang zu diesen Daten.

Ethische Herausforderungen für den Einsatz von ML-basierten Systemen in der Medizinforschung

Neben dem Zugang zu Daten stammen viele der ethischen Fragestellungen, die sich durch die Anwendung von ML-basierten Systemen in der medizinischen Forschung ergeben, aus dem Kontext der Anwendung von datenbasierten Systemen, Automatisierung und Vertrauen in Technologien. So werden diese Fragestellungen auch im Forschungskontext bereits bei der Entwicklung der Systeme relevant. Die Abbildung 1 zeigt einen vereinfachten Datenverarbeitungsprozess wie er bei der Gestaltung eines ML-basierten Systems durchlaufen wird.

Angefangen von der Sammlung der Daten über die Ermittlung und das Trainieren des Algorithmus, die tatsächliche Anwendung des Ergebnisses im Forschungsprozess sowie darüber hinaus, u. a. im Rahmen der Verantwortlichkeitsfragestellungen, müssen gesellschaftlich relevante Aspekte von den Systementwickler:innen und Anwender:innen betrachtet werden, um ein gesellschaftlich verträgliches System zu gestalten.

In der Phase der Datenanalyse der Abbildung 1 kommt die Gestaltung des Algorithmus zur Klassifikation der möglichen Ergebnisvarianten hinzu. Je nachdem wie z. B. die Fehlertoleranz oder die jeweiligen Grenzen der zu ermittelten Klassen definiert wurden, stellt sich die Frage nach der Verlässlichkeit der Ergebnisse. So hat ein Assistenzsystem in einer Studie den teilnehmenden Ärzten fälschlicherweise eine Empfehlung ausgestellt, Patient:innen nach Hause zu schicken, die noch nicht genesen waren (Caruana *et al.,* 2015). In einer anderen Studie, hat eine Applikation eine klinische Behandlung für Patient:innen angeregt, obwohl die erfassten Symptome keine solche erfordert hätten (Caruana *et al.,* 2015). Diese Fälle geben Hinweise darauf, dass die Rechenvorschrift in ihrer Klassifikation mathematisch richtig, jedoch nicht im Sinne der behandelten Ärzt:innen definiert wurde.

In der Phase der Ergebnisberechnung gehört u. a. die Frage nach der Transparenz über die im Datenverarbeitungsprozess bisher involvierten Entscheidungen. Um einen durch das ML-basierte System errechneten Vorschlag für eine Wirkstoffkombination einbeziehen zu können, müssen die Forschenden erkennen können, auf welcher Grundlage und unter Berücksichtigung welcher Faktoren dieser bestimmt wurde. Ohne diese Informationen können die Auswirkungen und die Anwendbarkeit des errechneten Ergebnisses nicht valide berücksichtigt werden. Hierdurch wird also die Validierung der Forschungsergebnisse gefährdet und damit ihre Integration in das medizinische Wissen.

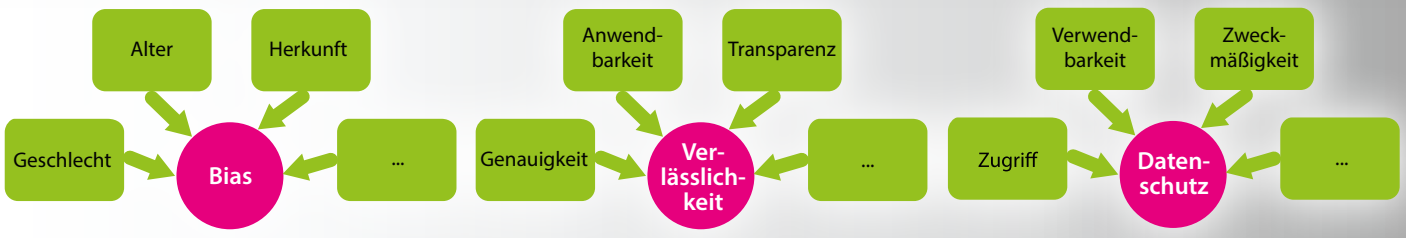


Abbildung 2: Einige ausgewählte ethische Herausforderungen in der Medizinforschung (nach Levina, 2020)

Daten bilden also die Grundlage für die durch ML-basierten Systeme berechneten Ergebnisse der Forschung und somit für deren künftigen Einsatz in diagnostischen Systemen, Versorgungsprozessen und Medikamenten. Die Qualität und Zusammensetzung der Daten kann demnach weitreichende Folgen haben. So wurde in zahlreichen Studien nachgewiesen, dass die Erkenntnisse aus der medizinischen Forschung häufig auf Daten basieren, die für die breite Bevölkerung wenig repräsentativ sind. So wurden z.B. Diagnosen und Medikamente für männliche Patienten entwickelt, die für weibliche Patienten unwirksam oder sogar schädlich sein können (siehe z. B. Popejoy and Fullerton, 2016). Auch ist die Sammlung und Auswertung von Daten aus den Bereichen oder Regionen, die noch keine ausreichende Digitalisierung ihrer Abläufe vorzeigen können, besonders schwierig und führt entsprechend dazu, dass diese Bereiche bei der Entwicklung der ML-basierten Systeme nicht berücksichtigt werden. So wird in den subsaharischen Ländern Afrikas im Durchschnitt bei jüngeren Frauen als in westlichen Ländern Brustkrebs diagnostiziert. Systeme, die darauf trainiert wurden, die Krankheit in einem weiter fortgeschrittenen Stadium und bei älteren Frauen zu identifizieren, können in diesem Szenario also nicht wirksam eingesetzt werden (Black and Richmond, 2019). Diese Verzerrung *(bias)* der Daten, d. h. die Über- oder Unter-Repräsentierung von bestimmten Gruppen im Datensatz, macht also die Fortschritte in der medizinischen Forschung nur für bestimmte Bevölkerungsgruppen anwendbar. Die Daten werden also aus dem Kontext heraus als Allgemeingut der Forschungsgemeinschaft betrachtet und manifestieren die Sicht auf die Gesellschaft und auf die Patient:innen in Softwarecode und damit in Prozesse, die die aktuelle und zukünftige Gesundheit zahlreicher Menschen unmittelbar betreffen.

Zusammenfassung

KI-Technologien können in vielen Bereichen der Medizin eingesetzt werden. Gleichzeitig schaffen sie eine Reihe ethischer Herausforderungen, die erkannt und abgemildert werden müssen, da diese Technologien die Präferenzen, die Sicherheit, die Gesundheit und die Privatsphäre der Patient:innen in hohem Maße bedrohen kann. Die derzeitige Politik und die ethischen Richtlinien hinken jedoch den Fortschritten hinterher, die die KI-Technologien im Bereich der Gesundheitsversorgung und -forschung gemacht haben. Obwohl es einige Bemühungen gibt, sich an diesen ethischen Gesprächen zu beteiligen, ist die medizinische Gemeinschaft nach wie vor noch ungenügend über die ethische Komplexität informiert, die die aufkeimenden KI-Technologien mit sich bringen können.

Referenzen:

Black, E., and Richmond, R. (2019). Improving early detection of breast cancer in sub-Saharan Africa: why mammography may not be the way forward. *Global Health* **15**, 3. https://doi.org/10.1186/s12992-018-0446-6

Caruana, R., *et al.* (2015). Intelligible models for healthcare, in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp1721-30.

Levina, O. (2020). A Research Commentary- Integrating Ethical Issues into the Data Process, in *WI2020 Community Tracks*. GITO Verlag, pp. 321–330. doi: 10.30844/wi_2020_z3-paper3.

Popejoy, AB., and Fullerton, SM. (2016). Genomics is failing on diversity. Nature 538:161.

Quartz (10 July 2017). If you're not a white male, artificial intelligence's use in healthcare could be dangerous.

The Society for Women's Health Research, The US FDA Office of Women's Health (2011) Dialogues on diversifying clinical trials.

Kontakt:

Dr. Olga Levina
wissenschaftliche Mitarbeiterin
FZI Forschungszentrum Informatik Berlin
levina@fzi.de

https://url.fzi.de/levina

IEAI: vorreiter bei der gestaltung ethischer KI

Globale, egalitäre und interdisziplinäre ethische Richtlinien für die Entwicklung und den Einsatz von Künstlicher Intelligenz

von Anastasia Aritzi und Caitlin Corrigan im Namen des IEAI

Unser Arbeits- und Lebensalltag und unsere Gesellschaft werden durch Künstliche Intelligenz (kurz KI) zunehmend verändert. „KI" dient dabei als Sammelbegriff für „Technologien und ihre Anwendungen, die durch digitale Methoden auf der Grundlage potenziell sehr großer und heterogener Datensätze in einem komplexen und die menschliche Intelligenz gleichsam nachahmenden maschinellen Verarbeitungsprozess ein Ergebnis ermitteln, das ggf. automatisiert zur Anwendung gebracht wird."[1].

Milliarden Menschen weltweit nutzen jeden Tag KI – und das meist, ohne sich dessen überhaupt bewusst zu sein. Zu bekannten KI-Anwendungen zählen autonom fahrende Autos, digitale Assistenten, Chatbots, Gesichtserkennungstechnologien und personalisierte Empfehlungsdienste. KI wird unsere Gesellschaft unaufhaltsam verändern. Das wirft dringende Fragen auf: Wie sieht eine solche Veränderung aus? Welche Folgen wird sie haben?

Das Institute for Ethics in Artificial Intelligence (Institut für Ethik in der Künstlichen Intelligenz / IEAI) untersucht Problemfelder im Zusammenhang mit Einsatz und Einfluss von KI und übernimmt so eine Führungsrolle in diesem Transformationsprozess. Das 2019 gegründete Institut unter der Leitung von Professor Christoph Lütge, Inhaber des Peter Löscher-Stiftungslehrstuhls für Wirtschaftsethik an der Technischen Universität München (TUM), ist Teil des Munich Center for Technology in Society (MCTS) der TUM.

[1] Deutsche Datenethikkommission – Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung (9. Oktober 2018) https://www.bmi.bund.de/SharedDocs/downloads/EN/themen/it-digital-policy/recommendations-data-ethics-commission.pdf?__blob=publicationFile&v=3

Das IEAI kurz gefasst

Stellen Sie sich vor, die KI-gestützte Bewerbermanagement-Software eines Unternehmens würde Sie aufgrund Ihres Geschlechts, Ihrer sexuellen Orientierung, Ethnie, Religionszugehörigkeit, Ihres Familienstands, Gesundheitszustands oder Alters aussortieren. Oder denken Sie an eine Situation, in der ein Algorithmus darüber entscheidet, wessen Sicherheit wichtiger ist: die der Insassen eines autonom fahrenden Autos oder die anderer Verkehrsteilnehmer, wie zum Beispiel eines Radfahrers.

In beiden Fällen muss die sinnvolle und sichere Entwicklung der Technologien durch eine angemessene Erwägung ethischer Fragen gestützt werden. KI und Algorithmen bieten die Möglich-

„AI cannot fly without ethics."
Prof. Christoph Lütge, Direktor des IEAI

Foto: Sonja Herpich

keit, Diskriminierungen entweder abzubauen oder aber zu verschärfen, und haben direkten Einfluss auf die Sicherheit und Unversehrtheit von Menschen. Themen wie Haftung, Transparenz und Datenschutzverstöße werden immer dann aktuell, wenn es um Systeme geht, die Entscheidungen auf der Grundlage großer – und teils personalisierter – Datenmengen treffen.

Der inter-, multi- und transdisziplinäre Forschungsansatz des IEAI zielt darauf ab, diese Herausforderungen zu erkennen, umsetzbare Ergebnisse hervorzubringen und konkrete Rahmenbedingungen, Methoden und Algorithmen zu erarbeiten, die KI-Entwicklern und Anwendern als Best Practices dienen können. Pro Forschungsprojekt arbeiten zwei Professoren aus unterschiedlichen Fakultäten, unterstützt durch ihre Forschungsteams, gemeinsam an der Entwicklung ethischer Leitlinien für die KI.

Die aktuellen Forschungsschwerpunkte des IEAI lassen sich in sieben Cluster unterteilen (siehe Tabelle 1). Die Forschungsprojekte widmen sich Fragestellungen wie zum Beispiel: Wie lässt sich KI am besten am Arbeitsplatz einsetzen? Können autonome Fahrzeuge ethisch handeln? Welche Folgen hat KI-gestütztes Marketing auf Entscheidungen, Autonomie und Verhalten von Konsumenten? Welche ethischen und leistungsbezogenen Probleme und Risiken ergeben sich aus dem Einsatz von KI im Gesundheitswesen?

Als Plattform für interdisziplinäre Zusammenarbeit richtet das IEAI Veranstaltungen aus und übernimmt die Federführung von verschiedenen Netzwerken, darunter das Global AI Ethics Consortium und Responsible AI Network-Africa, die dem Austausch zwischen unterschiedlichen Interessensgruppen und Fachbereichen dienen. Durch den Zusammenschluss mit Forschern und Entwicklern aus aller Welt ist das IEAI in der Lage, reale Probleme zu lösen und zum globalen Diskurs um Ethik und KI beizutragen.

KI, Ethik und die Corona-Krise

Von Frühwarnsystemen über die Erkennung von Ausbreitungstrends bis hin zur Kontaktverfolgung und Bestimmung wirksamer Behandlungsansätze: KI übernimmt in der Corona-Krise sehr viele Aufgaben. Der Einsatz einiger dieser Tools, beispielsweise von Apps zur Kontaktverfolgung, zur Quarantäne-Überwachung oder von Gesichtserkennungs-Software, hat jedoch Debatten rund um Datenschutz, Sicherheit, Menschen- und Freiheitsrechte befeuert und ernstzunehmende Fragen über die Auswirkungen dieser Technologien über die Krise hinaus

Tabelle 1: Die sieben Forschungscluster in der aktuellen IEAI Forschung

1. **KI UND DIE ZUKUNFT DER ARBEIT**
   A Human Preference-aware Optimization System Project
2. **KI, MOBILITÄT UND SICHERHEIT**
   ANDRE – AutoNomous DRiving Ethics Project
3. **KI, WAHLMÖGLICHKEIT UND AUTONOMIE**
   Consumer Perception and Acceptance of AI-enabled Recommender System Project
4. **KI IN MEDIZIN UND GESUNDHEITSWESEN**
   METHAD – Toward a MEdical ETHical ADvisor System for Ethical Decisions Project
5. **KI UND ONLINE-VERHALTEN**
   Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability, and Mitigation Project
   Online-Offline Spillovers-Potential Real-world Implications of Online Manipulation Project
   Personalized AI-based Interventions Against Online Norm Violations: Behavioural Effects and Ethical Implications Project
6. **KI, GOVERNANCE UND REGULIERUNG**
   TrustMLRegulation - Managing Trust and Distrust in Machine Learning with Meaningful Regulation Project
7. **KI UND NACHHALTIGKEIT**
   Artificial Intelligence for Earth Observation (AI4EO) Project

Für weitere Informationen über die Projekte besuchen Sie https://ieai.mcts.tum.de/research/
Letzte Aktualisierung: Sommer 2020

aufgeworfen. Natürlich gibt es immer Kosten-Nutzen-Abwägungen. Diese müssen jedoch stets transparent sein und klar benannt und erläutert werden.

KI ist ein mächtiges Instrument, dass Maßnahmen zur Pandemieeindämmung auf die oben beschriebene Weise unterstützen kann. Doch birgt sie bei nachlässigem oder unverantwortlichem Einsatz auch die Gefahr, dass Ungleichbehandlung und Voreingenommenheit zementiert und verschlimmert werden – weil beispielsweise falsche, unzureichende oder unausgewogene Daten zu diskriminierenden Ergebnissen führen. Die Corona-Krise hat deutlich gemacht, dass es einer koordinierten, zweckgebundenen Infrastruktur und eines ebensolchen Ökosystems für Daten und Technologie bedarf, um derart dynamischen Gefahren für Gesellschaft und Umwelt zu begegnen und dass wir zum einen uns dringend den ethischen Fragen stellen müssen, die der Einsatz von KI in der Medizin aufwirft, und zum anderen anwendbare ethische Leitlinien entwickeln müssen.

Abbildung 1: Das IEAI in Zahlen. Für weitere Informationen über Mitarbeiter des IEAI, besuchen Sie: https://ieai.mcts.tum.de/about-ieai/people
Letzte Aktualisierung: Oktober 2020 (Quelle: nach Aritzi und Corrigan, 2020).

Initiativen des IEAI angesichts der Corona-Krise

Research Brief über Ethische Dimension des Einsatzes von KI zur Eindämmung der Corona-Krise

„Klare und genau umrissene ethische Leitlinien können eine wichtige Rolle im Kampf gegen medizinische Krisen und Pandemien spielen. Um KI-Systeme erfolgreich einzusetzen, müssen die Anwender ihnen allerdings vertrauen können."[2]

Messen wir der Volksgesundheit und dem Allgemeinwohl einen höheren Wert bei, als der Privatsphäre und Autonomie des Individuums? Wie lassen sich KI-gestützte Entscheidungen zur Verteilung begrenzter medizinischer Ressourcen mit Grundsätzen der Gerechtigkeit und Schadensvermeidung vereinbaren? Diesen und weiteren Fragen gilt es im Zusammenhang mit der immer rasanteren Entwicklung von KI-gestützten Lösungen für die Bewältigung von Gesundheitskrisen nachzugehen. Das IEAI hat im Frühjahr 2020 einen Research Brief veröffentlicht, der einige aktuelle und potenzielle Einsatzbereiche von KI-gestützten Tools zur Pandemiebekämpfung beschreibt und die ethischen Dimensionen dieser Methoden erläutert. Der vollständige (englischsprachige) Research Brief ist auf der Website des IEAI unter dem Menüpunkt „Publications and Reports" abrufbar.

[2] IEAI. Ethical Implications of the Use of AI to Manage the COVID-19 Outbreak (April 2020). https://ieai.mcts.tum.de/wp-content/uploads/2020/04/April-2020-IEAI-Research-Brief_Covid-19-FINAL.pdf

Gründung des Global AI Ethics Consortium

„Die Zeit für eine Analyse des KI-Einsatzes ist jetzt – wer ist von welchen Auswirkungen wie betroffen, und mit welchen weiterreichenden gesellschaftlichen und wirtschaftlichen Folgen ist zu rechnen?"[3]

Die unterschiedlichen Einsatzgebiete von KI bei der Pandemiebekämpfung und die damit verbundenen ethischen Fragen erfordern einen multidisziplinären Ansatz unter Beteiligung verschiedenster Stakeholder. Richtlinien zur KI-Governance müssen auf internationaler Ebene gemeinsam erarbeitet werden. Mit der Gründung des Global AI Ethics Consortium (GAIEC) im April 2020 hat das IEAI einen wichtigen Schritt in diese Richtung getan. Das Konsortium bietet namhaften Mitgliedern universitärer und privater Forschungseinrichtungen aus aller Welt die Möglichkeit in eigenen Forschungsprojekten zusammenzuarbeiten und ihr Know-how auch anderen Stakeholdern zur Verfügung zu stellen.

[3] Global AI Ethics Consortium. Das globale AI-Ethik-Konsortium für Ethik und den Einsatz von Daten und künstlicher Intelligenz im Kampf gegen COVID-19 und andere Pandemien – Erklärung zur Zielsetzung (2020). https://ieai.mcts.tum.de/wp-content/uploads/2020/05/COVID-19-Statement-of-Purpose-FINAL-May-2020.pdf

Das Institute for Ethics in Artificial Intelligence ist eine Einrichtung der Technischen Universität München. (Foto: Andreas Heddergott/TUM).

Forschungsaufruf für multidisziplinäre Projekte zum Thema Ethik beim Einsatz von KI zur Pandemiebekämpfung und in Gesundheitskrisen

„Die wissenschaftliche Forschung spielt eine wichtige Rolle bei der Entwicklung von Tools und Rahmenrichtlinien, mit deren Hilfe Regierungen, Unternehmen, NGOs und andere Akteure zu vernünftigen, ethisch vertretbaren Lösungen für den KI-Einsatz bei der Pandemievorbeugung und -bekämpfung gelangen können." Das IEAI hat daher einen Forschungsaufruf für multidisziplinäre Projekte zur Untersuchung und Entwicklung forschungsbasierter, konkreter und praktischer Lösungen zu ethischen Problemen im Zusammenhang mit dem Einsatz von KI bei der Pandemiebekämpfung und in Gesundheitskrisen veröffentlicht. Die ausgewählten Projektvorschläge wurden im August 2020 bekanntgegeben und die Forschungsprojekte kurz darauf gestartet.

Was kommt

„TRAIF: Förderung von nachhaltigen, umfassenden und ganzheitlichen Leitlinien für einen weltweit vorteilhaften KI-Einsatz"

Das Responsible AI Forum (#TRAIF2021) verfolgt zwei Ziele. Erstens will das Forum Vertretern aus Wirtschaft und Industrie, Zivilgesellschaft, Politik und Wissenschaft eine Plattform bieten, auf der sie die wichtigsten und dringendsten Fragen im Zusammenhang mit dem verantwortungsvollen Einsatz von KI diskutieren und anhand von Erfahrungsaustausch, neusten Forschungsergebnissen und praktischen Anwendungen erörtern können. Zweitens soll der Austausch zwischen Wissenschaft und Praxis durch produktive Debatten und Vorführungen gefördert werden. TRAIF wird vom IEAI ausgerichtet und findet Anfang Juni 2021 in München statt. Mehr Informationen und Anmeldungen unter: responsibleaiforum.com

IEAI

Kontakt:

Prof. Dr. Christoph Lütge
Direktor des IEAI
INSTITUTE FOR ETHICS IN ARTIFICIAL INTELLIGENCE
München
ieai@mcts.tum.de

https://ieai.mcts.tum.de

potenziale und herausforderungen von KI in der medizin

Heilungschancen erhöhen, Krankheiten vorbeugen, Selbstbestimmung ermöglichen

von Klemens Budde und Karsten Hiltawsky
für die Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme

Egal ob bei der Prävention, frühzeitigen Diagnose oder der patientengerechten Therapie – Künstliche Intelligenz (KI) und Lernende Systeme können in der Medizin zu besseren Behandlungsergebnissen führen und somit unsere Gesundheitsfürsorge verbessern. Gleichzeitig stellt der Einsatz von KI-Assistenzsystemen im Gesundheitswesen auch hohe Anforderungen an die IT-Sicherheit und setzt das Vertrauen von medizinischem und pflegerischem Personal sowie von Patientinnen und Patienten voraus. Der Beitrag thematisiert die Potenziale, Herausforderungen und Voraussetzungen für einen sicheren, ethisch wünschenswerten und wirtschaftlich sinnvollen Einsatz von KI in Gesundheit und Pflege.

Die Corona-Pandemie zeigt aktuell nicht nur die Verwundbarkeit unserer medizinischen Versorgung auf, die Krise bietet auch eine Chance, die Resilienz unseres Gesundheitssystems durch technologische Innovationen zu erhöhen. So ermöglichen Methoden der Künstlichen Intelligenz und des Maschinellen Lernens gerade in der Epidemiologie, also der Untersuchung von Zusammenhängen und Verteilungen von Krankheiten und Risikofaktoren in der Bevölkerung, große Datenmengen effizient auswerten und neue Erkenntnisse daraus generieren zu können. Darüber hinaus kann KI in der Medizin bei der Prävention, der frühzeitigen Diagnose oder der passgenauen Therapiewahl einen wesentlichen Beitrag dazu leisten, dass Menschen schon in naher Zukunft medizinisch besser und individueller versorgt werden. Die Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme hat dazu Chancen und Herausforderungen Lernender Systeme im Gesundheitswesen in einem **Bericht** zusammengefasst, auf dem dieser Beitrag basiert. Die vom Bundesministerium für Bildung und Forschung (BMBF) initiierte **Plattform Lernende Systeme** ist in themenspezifische Arbeitsgruppen gegliedert und bündelt die vorhandene Expertise, um KI im Sinne der Gesellschaft zu gestalten.

Lernende Systeme
DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

Potenziale von KI in der Medizin

Ein Blick ins Krankenhaus der Zukunft zeigt die vielfältigen Einsatzmöglichkeiten von Lernenden Systemen: Dort stützen sich Ärztinnen und Ärzte beim Auswerten bildgebender Verfahren flächendeckend auf KI-Systeme und erhalten auf diese Weise präzisere Diagnosen. Neue Präventionsansätze können so entwickelt, seltene Erbkrankheiten erforscht und bislang unbekannte medizinische Zusammenhänge entdeckt werden. Aus vernetzten Daten können Lernende Systeme Vorschläge zu geeigneten Präventionsansätzen oder Therapiemaßnahmen ableiten – und auf diese Weise Medizinerinnen und Mediziner im Entscheidungsprozess unterstützen. Bilder aus Radiographien, nuklearmedizinischen Verfahren, Magnetresonanztomographien oder Ultraschall von Organsystemen können mit Hilfe von KI noch präziser, schneller und zuverlässiger analysiert werden. Da sich aus KI-Algorithmen nicht ungefiltert Handlungsentscheidungen ableiten lassen, kann und soll medizinisches und pflegerisches Personal durch diese technologischen Innovationen nicht ersetzt werden. Stattdessen stellen KI-

Assistenzsysteme in der Medizin eine komplementäre, wichtige Informationsquelle dar, etwa für die Therapiewahl. Zudem bleiben Ärztinnen und Ärzte für die Kommunikation mit den Betroffenen unverzichtbar, da sie nur gemeinsam mit der Patientin oder dem Patienten die bestmögliche Entscheidung treffen können.

Aber auch in der Pflege birgt das Zusammenspiel von Mensch und Maschine großes Potenzial: Eine KI-gestützte Spracherfassung könnte Pflegekräfte bei Routineaufgaben, etwa der Dokumentation, entlasten. Die dadurch gewonnene Zeit könnte für die menschliche Zuwendung genutzt werden. Assistenzroboter und KI-basierte Technologien (z. B. Exoskelette) könnten künftig außerdem ermöglichen, dass Menschen bis ins hohe Alter selbstbestimmt leben. Nicht nur Erkrankte oder Pflegebedürftige könnten durch bessere Therapiemöglichkeiten und höheren Heilungschancen von medizinischen KI-Anwendungen profitieren: Lernende Systeme weisen gerade in der Prävention und Früherkennung von Krankheiten großes Potenzial auf. So können die eigenen Gesundheitsdaten mit Hilfe von Smartphone-Apps oder Wearables erfasst und ausgewertet werden. Damit kann der Alltag gesünder gestaltet und Krankheitssymptome früher identifiziert werden.

Gesundheitsdaten austauschen und schützen

Damit Lernende Systeme in der Medizin zum Wohle von Patientinnen und Patienten eingesetzt werden können, müssen bestimmte Rahmenbedingungen geschaffen werden. Dazu gehört auch der Aufbau einer Gesundheitsdatenbasis: KI-Systeme sind nur so gut, wie die zur Verfügung stehenden Daten. Für Lernprozesse von KI-Systemen im Gesundheitswesen ist es daher wichtig, dass ausreichend diversifizierte, nutzbare Daten vorhanden sind. Mit jeder Datenerweiterung werden KI-Systeme in der Regel zuverlässiger und die Ergebnisse dadurch auch stabiler – das gilt auch für die Anwendung in der Medizin. Für eine geeignete Auswertung sollten zudem Daten aus verschiedenen Quellen vernetzt sein. Eine repräsentative, strukturierte und kontrollierte Gesundheitsdatenbasis gilt daher als Voraussetzung dafür, dass KI-basierte Lösungen im Gesundheitswesen entwickelt, angewendet und akzeptiert werden können. Personenbezogene Gesundheitsdaten sollten in der Regelversorgung für Methoden des Maschinellen Lernens kontinuierlich nutzbar gemacht werden. Beispielsweise können die Daten für die Unterstützung der ambulanten oder stationären Pflege eingesetzt werden. Damit im Zusammenhang stehen Fragen zur Nutzung, Speicherung und zur IT-Sicherheit von KI-Systemen und die damit verbundene Sorge, zum „gläsernen Patient“ zu werden. Damit genügend Menschen der Verwendung ihrer Daten in einer vernetzten Datenbasis zustimmen und KI in der Medizin gesellschaftlich akzeptiert wird, muss daher der selbstbestimmte Umgang der Betroffenen mit ihren Gesundheitsdaten garantiert werden.

Kompetenzaufbau in der medizinischen Ausbildung und Pflege

Berufliche Rollenprofile und der Berufsalltag von medizinischem und pflegerischem Personal werden sich durch die neuen technologischen Möglichkeiten verändern. Um KI-basierte Anwendungen in der Diagnostik oder Therapiewahl verantwortungsvoll nutzen zu können, sind Grundlagenkenntnisse im Bereich des Maschinellen Lernens und der Informationstechnik erforderlich. Auch die Arzt-Patienten-Beziehung wird

Publikation „Lernende Systeme im Gesundheitswesen“
Der Beitrag basiert auf dem Bericht der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme. Der Expertenbericht veranschaulicht das Potential der Technologien anhand von Forschungsbeispielen und einem Anwendungsszenario zum Thema Lungenkrebs.

Plattform Lernende Systeme (Hrsg.) Lernende Systeme im Gesundheitswesen – Bericht der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege, München 2019. Online verfügbar unter: www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG6_Bericht_23062019.pdf

Potenziale von KI in der Medizin nutzen (Foto: iStock © metamorworks)

sich verändern, wenn Algorithmen Entscheidungsprozesse unterstützen und Betroffene aktiver an der Datenerfassung und -auswertung teilnehmen. Um das Potential von KI im Gesundheitswesen auszuschöpfen, ist neben diesem Kompetenzaufbau in der medizinischen Ausbildung und Pflege auch eine geeignete Grundlagenforschung erforderlich. Ebenso können klinische Leuchtturmprojekte dazu dienen, Wissen aus der Forschung exemplarisch in die Praxis zu implementieren und auf Umsetzung und Nutzen zu prüfen. Um KI-Innovationen voranzubringen, müssten Unternehmen aus den Bereichen Medizintechnik, Pharmazie, Hard- und Software zudem in Gesundheits-KI investieren.

Innovationen fördern und zu Betroffenen bringen

Die Einführung von KI im Gesundheitswesen ist ein weitreichender Innovationsprozess. Sie bedingt, dass alle relevanten Akteure, wie etwa Patientenverbände, Arbeitnehmervertretungen und Ethik-Gremien, diesen Prozess aufmerksam begleiten. Zudem müssen regulatorische Fragen beantwortet werden. So muss etwa sichergestellt sein, dass medizinisches und pflegerisches Personal KI-Innovationen in einem klaren haftungsrechtlichen Rahmen verwenden kann. Bei der Zulassung entstehen spezifische Herausforderungen, da KI-Assistenzsysteme ständig weiter durch neue Daten dazulernen. Die bisherigen Regelungen für klinische Studien und die Zulassung von Medizinprodukten tragen dieser Veränderung bislang noch nicht Rechnung. Daher müssen neue Regeln für lernende Medizingeräte erarbeitet werden, die Unternehmen Rechtssicherheit geben.

Im Mittelpunkt jeglicher technologischer Errungenschaften muss aber der Nutzen für Betroffene stehen. Die größte Herausforderung besteht dabei darin, die Ergebnisse transparent darzulegen. Denn auch wenn das System nachweislich gute Ergebnisse liefert, muss gezeigt werden können, warum ein bestimmtes Ergebnis erzielt wird (Kausalität). Die Nachvollziehbarkeit wird zudem ein wichtiges Kriterium für Weiterverbreitung und Nutzung darstellen, denn Ärztinnen und Ärzte sowie Patientinnen und Patienten möchten wissen, aufgrund welcher Befunde ein Ergebnis erzielt wurde, gerade auch im Hinblick auf möglicherweise fehlerhafte Primärdaten. Um KI-Innovationen im Gesundheitswesen voranzubringen und um sie zum Wohle aller Patienten einzusetzen, muss die Forschung im Rahmen der Hochschulmedizin und in anderen Forschungsinstitutionen gestärkt werden. Zusätzlich werden Unternehmen benötigt, die die Technologie zur Marktreife bringen, sodass Krankenhäuser, Arztpraxen und Pflegeeinrichtungen sichere und nützliche KI-Innovationen einsetzen können. Voraussetzung für die Investition in die Entwicklung von KI-Medizinprodukten ist dabei die Aufnahme in Leistungskataloge der Krankenkassen. Die Erstattung setzt einen klaren Rechtsrahmen voraus, der die Nutzung und Entwicklung von KI-Anwendungen im Gesundheitswesen ermöglicht und reguliert. Bedingung für die Erstattung sind neue regulatorische und technische Rahmenbedingungen, mit denen die Ergebnisse KI-basierter Innovationen transparent, messbar und vergleichbar werden. Dazu gehört auch die Regulierung und Förderung klinischer Studien

(Forschungs-, Zulassungs- und Marktbeobachtungsstudien), die die Sicherheit, Wirksamkeit und Nützlichkeit von lernenden Medizinprodukten valide nachweisen.

Ausblick: Diskurs über ethische Fragen notwendig

Darüber hinaus sind die kritische Auseinandersetzung mit ethischen Fragen und die gesellschaftliche Akzeptanz notwendige Bedingungen für den Einsatz von KI: Nur, wenn ein effektiver Datenschutz sowie die Einhaltung gesellschaftlicher Normen und Werte garantiert sind, können Menschen KI vertrauen. Mit dem Einsatz von KI sind zahlreiche ethische Fragen verbunden – gerade im Gesundheitswesen, wo die körperliche und seelische Unversehrtheit des Menschen im Mittelpunkt steht und es um sehr sensible Daten geht. Daher ist es nötig, dass die technologischen Innovationen von einer breiten gesellschaftlichen Debatte begleitet werden: Patientenverbände und Ethik-Gremien sollten frühzeitig in die Diskussion über ethische Standards und in politische Entscheidungsgremien einbezogen werden. Dilemma-Situationen wird es weiterhin geben und auf manche Fragen gibt es derzeit noch keine eindeutigen Antworten, da noch nicht alle Auswirkungen dieser technologischen Neuerungen absehbar sind. Bislang fehlen wichtige rechtliche Rahmenbedingungen für die Verwendung von Gesundheitsdaten für KI, sie setzen eine sorgfältige Interessenabwägung zwischen medizinischen Vorteilen und dem Schutz der Persönlichkeitsrechte voraus. Der Einzug von KI im Gesundheitswesen kann zudem zu neuen Interessenkonflikten verschiedener Akteure führen (z. B. Wissenschaft, Wirtschaft, Start-Ups, Pharmaindustrie, Medizintechnik, Wohlfahrtsverbände, Krankenkassen). Unklar ist bislang auch, wie sich die Entscheidungsautonomie von Ärztinnen und Ärzten, Pflegerinnen und Pflegern sowie Patientinnen und Patienten durch den Einsatz von KI wandelt. Das Urteilen und Entscheiden nach ethischen Kriterien ist integraler Bestandteil des medizinischen und pflegerischen Alltags. Lernende Systeme können diese Aufgabe nicht übernehmen, weshalb KI auch in Zukunft lediglich als Unterstützung für den Menschen eingesetzt werden kann. Dieser Tatsache gilt es durch entsprechende politische Regulierung Rechnung zu tragen.

Kontakt:

Prof. Dr. med. Klemens Budde ist Co-Leiter der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme und Leitender Oberarzt an der Medizinischen Klinik mit Schwerpunkt Nephrologie und Internistische Intensivmedizin an der Charité –Universitätsmedizin Berlin.
klemens.budde@charite.de

https://nephrologie-intensivmedizin.charite.de/metas/person/person/address_detail/budde-1/

Dr.-Ing. Dr. med. Karsten Hiltawsky ist Co-Leiter der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme und seit 2012 bei Drägerwerk AG & Co. KGaA, momentan leitet er dort die Abteilung Technology und Intellectual Property.

karsten.hiltawsky@draeger.com

Über die Plattform Lernende Systeme

Die Plattform Lernende Systeme wurde 2017 vom Bundesministerium für Bildung und Forschung (BMBF) initiiert, vereint Expertise aus Wissenschaft, Wirtschaft, Politik und Gesellschaft, und unterstützt den weiteren Weg Deutschlands zu einem international führenden Technologieanbieter für Lernende Systeme. Die rund 200 Mitglieder der Plattform sind in Arbeitsgruppen und einem Lenkungskreis organisiert. Sie zeigen den persönlichen, gesellschaftlichen und wirtschaftlichen Nutzen von Lernenden Systemen auf und benennen Herausforderungen und Gestaltungsoptionen.

„herr doktor, verstehen sie mich?“

Wie lernende Systeme helfen medizinische Fachsprache zu verstehen und welche Rolle klinische Leitlinien dabei spielen.

von Florian Borchert, Christina Lohr, Luise Modersohn, Udo Hahn, Thomas Langer, Gregor Wenzel, Markus Follmann und Matthieu-P. Schapranow

Relevantes medizinisches Wissen für die Behandlung Krebskranker wird heute in Form sogenannter klinischer Leitlinien bereitgestellt. Deren Inhalte werden durch qualifizierte Onkologie-Experten der Krebsgesellschaft nach Indikation sortiert und oftmals manuell recherchiert. Sobald neue Erkenntnisse verfügbar sind, werden die bestehenden Leitlinien ergänzt und neue Versionen publiziert.

Liebe Leserinnen und Leser, wir möchten Sie auf eine Reise in eine nicht allzu ferne Zukunft mitnehmen. Digitale Lösungen unterstützen das Gesundheitssystem an vielen Stellen. Auch die Behandlung Krebskranker profitiert davon, z. B. durch den Einsatz lernender Systeme bei der klinischen Dokumentation, bei der Entscheidungsunterstützung oder bei der Prognose klinischer Endpunkte. Abbildung 1 zeigt den Status Quo bei der Erstellung klinischer Leitlinien. Abbildung 2 zeigt wie künftig auch onkologische Leitlinien durch lernende Systeme auf Basis von Künstlicher Intelligenz (KI) noch schneller und evidenzbasiert in die Praxis gelangen.

Der Austausch medizinischer Daten zwischen Kliniken erfolgt über die digitalen Infrastrukturen, die durch die Medizininformatik-Initiative geschaffen wurden. Kliniken sind mehr als nur reine Datenspeicher: sie sind auch Quelle neuer medizinischer Erkenntnisse. Sie bieten evidenzbasiert Einblick in den Erfolg von Therapien anhand konkreter Fälle und sind daher wichtige Partner bei der Bewertung klinischer Leitlinien. Mit Hilfe von KI-Algorithmen können die Behandlungsverläufe hunderter Krebspatienten automatisiert analysiert werden. Experten der onkologischen Fachgesellschaften bewerten regelmäßig die Ergebnisse der KI-Algorithmen und prüfen, inwiefern Leitlinien ergänzt werden sollten. So können neue Versionen klinischer Leitlinien häufiger publiziert werden. Maschinenlesbare Versionen der Leitlinien gewinnen auch immer mehr an Bedeutung, z. B. um die Wissensvermittlung und Qualitätssicherung durch klinische Informationssysteme zu unterstützen. An dieser Stelle ist die automatische Verarbeitung menschlicher Sprache – auch als „Natural Language Processing“ (NLP) bezeichnet – eine wichtige technische Komponente.

Abbildung 1: Heutiger linearer Prozess zur Erstellung klinischer Leitlinien (vereinfachte Darstellung)

Quelle: Matthieu-P. Schapranow

Abbildung 2: Lernendes KI-Leitlininien-System. Iterativer Prozess zur Erstellung von Leitlinien mit Hilfe eines lernenden KI-basierten Leitlinien-Systems (Quelle: M. Schapranow, www.flaticon.com/packs/medical-icons und www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Ministerium/Flyer_Poster_Etc/BMG-Infografik_Gesundheitssystem_2019.pdf).

Maschinelles Lernen für die Verarbeitung biomedizinischer Texte

Die Verarbeitung menschlicher Sprache durch Computer ist seit jeher ein zentrales Ziel der KI. Menschliche Sprache ist allerdings auf vielen Ebenen für die maschinelle Verarbeitung komplex, z. B. kann sie vieldeutig, kontextabhängig und oft nur durch implizite Informationen interpretierbar sein. In speziellen Domänen wie der Medizin ergeben sich zusätzliche Herausforderungen, wie das ausgeprägte Fachvokabular, die häufige Verwendung von Abkürzungen und ein deutlich von der Alltagssprache abweichender, stark verdichteter Sprachstil.

Durch maschinelle Lernverfahren können auch beim Forschungsgebiet NLP beachtliche Fortschritte erreicht werden. Bei vielen Aufgaben, zum Beispiel der Extraktion von medizinisch relevanten Informationen aus Texten, basieren die aktuell erfolgreichsten Verfahren auf überwachten Lernverfahren *(supervised learning)*. Bei diesen Verfahren werden die verfügbaren Sammlungen realer Texte (sog. Textkorpora) durch inhaltliche Metadaten mit Bezug zur zu lernenden Aufgabe durch menschliche Annotatoren ergänzt, um den Algorithmen geeignete Beispiele für das Lernen zur Verfügung zu stellen. Die Verfügbarkeit großer, möglichst reich annotierter Textkorpora ist im Zusammenspiel mit Lernverfahren, die auf sog. tiefen neuronalen Netzen *(deep learning)* beruhen, wesentlicher Treiber der aktuellen NLP-Forschung (Wu *et al.*, 2020). Bisher ist die NLP-Forschung dominiert von englischsprachigen Texten, weil englischsprachige Textkorpora zahlreicher verfügbar sind als für alle anderen Sprachen. Dieser Fokus auf englischsprachige Dokumente führt jedoch zu Problemen. So sind z. B. sprachanalytische Lernmodelle, die auf englischen Texten trainiert wurden, in aller Regel nicht auf Texte in anderen Sprachen anwendbar.

Für die automatische Analyse klinischer Leitlinien ist der Mangel an einzelsprachspezifischen NLP-Modellen ein veritables Hindernis, da Leitlinien in der jeweiligen Landessprache veröffentlicht werden. In Deutschland ist das Problem besonders ausgeprägt, da nahezu keine anderen medizinischen Texte in deutscher Sprache für NLP-Lernmodelle zur Verfügung stehen: Fachtexte werden meist in englischsprachigen Fachzeitschriften publiziert oder unterliegen Copyright-Beschränkungen. Klinische Texte, wie Arztbriefe oder Pathologiebefunde, sind nur unter sehr hohen Datenschutzauflagen zugänglich und stehen Forschern nicht in dem erforderlichen Maße de-identifiziert und öffentlich zugänglich zur Verfügung.

GGPONC: Ein deutschsprachiges Textkorpus basierend auf klinischen Leitlinien

Die deutschsprachige NLP-Forschung kann nun dank Partner der Konsortien „HiGHmed“ und „SMITH“ der BMBF-Medizininformatik-Initiative richtig Fahrt aufnehmen. Das Hasso-Plattner-Institut für Digital Engineering der Universität Potsdam, das JULIE Lab der Friedrich-Schiller-Universität Jena und das Leitlinienprogramm Onkologie haben mit GGPONC (Borchert *et*

al., 2020) erstmals einen deutschsprachigen medizinischen Textkorpus basierend auf klinischen Leitlinien erstellt und unter der folgenden URL Forschern in kürzester Zeit zugänglich gemacht:

www.leitlinienprogramm-onkologie.de/projekte/ggponc

Klinische Leitlinien der deutschen Krebsgesellschaft stellen für die NLP-Forschung eine wichtige Datenquelle dar. Sie eignen sich perfekt als Grundlage zur Erstellung eines deutschsprachigen medizinischen Textkorpus, das ganz ohne Personenbezug und damit ohne Datenschutzbeschränkungen auskommt.

Die Leitlinienempfehlungen und -hintergrundtexte, die die Basis der neuen Leitlinien-App des Leitlinienprogramms Onkologie bilden, werden dabei direkt von der deutschen Krebsgesellschaft bereitgestellt. In der ersten Version enthält das Korpus über 1,3 Mio. Token aus 25 Leitlinien und bewegt sich damit in der Größenordnung der bisher umfangreichsten deutschsprachigen klinischen Textkorpora, die im Gegensatz jedoch nicht öffentlich zugänglich sind.

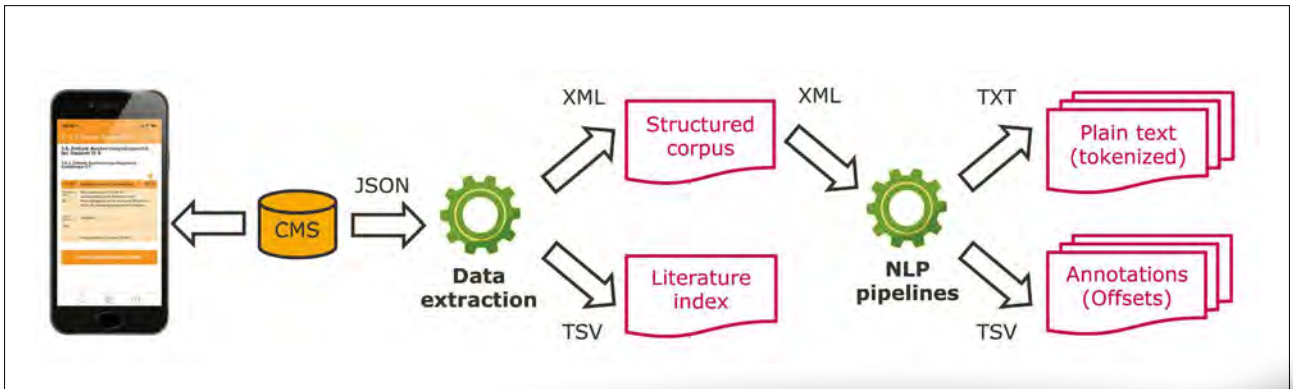
Abbildung 3 zeigt, wie nach der Erstellung der Textressource durch NLP-Verfahren medizinisch relevante Informationstypen automatisch aus den Leitlinientexten identifiziert werden. Grundlage hierfür sind u. a. die deutschsprachige UMLS-Terminologie sowie TNM-Ausdrücke. Ein Teil dieser automatisch bestimmten Informationstypen wurde durch medizinische Experten validiert und steht nun als „Goldstandard" für weitere Arbeiten zur Verfügung. Auch durch die Abbildung der gefundenen Informationstypen auf weitere Ontologien, wie SNOMED CT, wird eine Anreicherung der Texte mit semantischen Informationen möglich, um sie besser maschinenlesbar zu machen.

Wie können lernende Systeme Kliniker unterstützen?

Für die Zukunft stellt sich die Frage, in welcher Form Kliniker und medizinische Experten in die Entwicklung von NLP-Systemen frühzeitig eingebunden werden können. Die Abhängigkeit überwachter Lernverfahren von manuell annotierten Daten stellt die Skalierbarkeit dieser Ansätze in Frage. Auch erscheint es ineffizient, das Fachwissen von Experten nur in Form von zeitaufwändigen Annotationen in solche Systeme einfließen zu lassen.

Ein Lösungsansatz kann in der Adaption von Modellen aus anderen Domänen an den Kontext der Medizin *(transfer learning)* liegen. So ist es denkbar, hochwertige Repräsentationen von Wörtern im Kontext des deep learning *(word embeddings)* auf öffentlichen Daten wie GGPONC zu lernen. Diese würden es dann in einem klinischen Kontext erlauben, Systeme mit weniger Trainingsdaten aus der Zieldomäne zu entwickeln.

Abbildung 3: Pipeline zur automatischen Erstellung des annotierten GGPONC-Korpus aus den Daten des Content-Management-Systems des Leitlinienprogramms Onkologie



Quelle: Florian Borchert

Außerdem bieten sich verschiedene Modelle der Einbindung von Menschen in den Lernprozess *(human-in-the-loop)* an. So kann ein NLP-System zur Informationsextraktion aus Leitlinientexten direkt in das Redaktionssystem bei der Leitlinienerstellung eingebunden werden, um z. B. eine automatische Verschlagwortung zu unterstützen. Der menschliche Redakteur kann mit seinem Feedback unmittelbar die Modellvorhersagen korrigieren und so eine kontinuierliche Verbesserung des Modells ermöglichen. Vielversprechend sind ebenfalls neuartige Verfahren, die es ermöglichen, NLP-Systeme mit weniger starker menschlicher Überwachung *(weak supervision)* zu trainieren. Statt kleinteiliger Annotationen können Domänenexperten hier ihr Wissen in Form von Heuristiken, Metadaten oder unter Zugriff auf existierende kuratierte Wissensdatenbanken kodieren und so Annotationen automatisch für eine große Menge an Texten erzeugen (Ratner, Hancock & Ré 2019). Wie die Schnittstellen für eine derartige Interaktion mit einem lernenden System aussehen können, ist Gegenstand aktueller Forschung.

Steckbrief Forschungsprojekt:

Das German Guideline Program in Oncology NLP Corpus (GGPONC) ist eine 2019 gestartete Kooperation der Medizininformatikinitiative des BMBF, vertreten durch das Digital Health Center des Hasso-Plattner-Instituts und das JULIE Lab der Friedrich-Schiller-Universität Jena, gemeinsam mit der deutschen Krebsgesellschaft.

Referenzen:

Seufferlein, T., Kopp, I., Post, S., Jonat, W., Kreienberg, R., Nothacker, M., Marks, A., Nettekoven, R., Langer, T., Follmann, M., Bamberg, M. (2019). Onkologische Leitlinien – Herausforderungen und zukünftige Entwicklungen. Forum. 34(3) : 277–283.

Borchert, F., Lohr, C., Modersohn, L., Langer, T., Follmann, M., Sachs, J. P., Hahn, U., & Schapranow, M.-P. (2020). GGPONC: A Corpus of German Medical Text with Rich Metadata Based on Clinical Practice Guidelines. ArXiv:2007.06400 [Cs]. http://arxiv.org/abs/2007.06400

Cohen, K. B., and Demner-Fushman, D. (2014). Biomedical Natural Language Processing (Vol. 11). John Benjamins Publishing Company.

Ratner, A.J., Hancock, B., and Ré, C. (2019). The Role of Massively Multi-Task and Weak Supervision in Software 2.0. CIDR.

Wu, S. T. *et al.* (2020). Deep Learning in Clinical Natural Language Processing: A Methodical Review in Journal of the American Medical Informatics Association, 27, 457–470.

Kontakt:

Florian Borchert, M.Sc.
HPI Digital Health Center,
Hasso-Plattner-Institut, Universität Potsdam
florian.borchert@hpi.de

https://hpi.de/digital-health-center/members/working-group-in-memory-computing-for-digital-health/florian-borchert.html

Dr.-Ing. Matthieu-P. Schapranow
Leiter der Arbeitsgruppe
„In-Memory Computing for Digital Health“
Scientific Manager Digital Health Innovations
HPI Digital Health Center,
Hasso-Plattner-Institut, Universität Potsdam
schapranow@hpi.de

https://hpi.de/digital-health-center/members/working-group-in-memory-computing-for-digital-health/dr-ing-matthieu-p-schapranow.html

Fotos: privat / Schapranow und privat / Borchert

präzise diagnostik mit künstlicher intelligenz

Firmenporträt der Aignostics GmbH

von Stefanie Seltmann

Ob es sich bei einem verdächtigen Knoten um Krebs handelt oder nicht, wird in der Regel im pathologischen Labor unter dem Mikroskop geklärt. Um bei steigenden Erkrankungszahlen und immer feineren molekularen Details den Pathologen sowohl die Arbeit zu erleichtern als auch maßgeschneiderte Behandlungen zu ermöglichen, haben Wissenschaftler*innen vom Pathologischen Institut der Charité – Universitätsmedizin Berlin gemeinsam mit Kollegen von der TU Berlin ein digitales Bildanalyse-System entwickelt, das mit künstlicher Intelligenz mikroskopische Aufnahmen beurteilen kann. Der Digital Health Accelerator (DHA) des Berlin Institute of Health (BIH) unterstützte die Wissenschaftler*innen dabei, das langjährige Forschungsprojekt zum Translationserfolg zu führen. Im April 2020 konnten die Wissenschafter*innen nun die Aignostics GmbH gründen.

AIGNOSTICS

Professor Frederick Klauschen hat Physik und Medizin studiert, „eine gute Kombination, um die Digitalisierung in der Medizin voranzutreiben“, erzählt der stellvertretende Direktor des Instituts für Pathologie der Charité am Campus Charité Mitte. „Bei immer mehr Proben von immer mehr Patientinnen und Patienten können digitale Assistenzsysteme dabei helfen, Fehler zu vermeiden. Und der Mensch ist auch schlechter im Schätzen, wenn es etwa darum geht, zu beurteilen, wie viel Prozent eines Gewebes Tumor ist oder auf welchem Anteil der Tumorzellen sich ein bestimmter, therapeutisch relevanter Rezeptor befindet. Da kann uns der „digitale Kollege“ helfen, weil er sowohl schneller als auch präziser beim Zählen ist.“

Um den „digitalen Kollegen“ in der pathologischen Diagnostik auszubilden, arbeitet das Team um Klauschen mit vielen menschlichen Kolleg*innen von verschiedenen Universitätskliniken zusammen. Diese zeichnen auf Tausenden von digitalen mikroskopischen Aufnahmen von Gewebeschnitten die pathologischen Veränderungen ein. Mit diesen Befunden „füttern“ die Pathologen die Software, so dass diese „lernen“ kann, wie sich z. B. Tumorgewebe optisch von gesundem Gewebe unterscheidet. „Diese so genannten Annotationen haben wir für verschiedene Krankheiten generiert“, erklärt Klauschen. Bisher haben die Wissenschaftler*innen die Software so trainiert, dass sie zuverlässig Lungen-, Brust und Darmkrebs erkennen kann, sowie Immunzellen im Tumorgewebe und auch verschiedene Tumormarker. Ebenso können Infektionen, degenerative, Bindegewebs- oder Autoimmunerkrankungen analysiert werden, „alles was im Gewebe zu sichtbaren Veränderungen führt.“

Erklärbare Künstliche Intelligenz

„Der von uns entwickelte Ansatz zeigt uns darüber hinaus, wie die KI zu ihrer Entscheidung gekommen ist“, erklärt Professor Klaus-Robert Müller, Mitgründer von Aignostics, der an der TU Berlin den Lehrstuhl für Maschinelles Lernen innehat. „Die Software erzeugt sogenannte „Heatmaps“, welche präzise zeigen, welche Zellen oder Bildbereiche für die Klassifikation z. B. von Krebs vs. Nicht-Krebs für den Algorithmus entscheidend waren.“ Anhand dieser „Heatmaps“ können die Pathologen anschließend beurteilen, ob die Analyse der KI nachvollziehbar und korrekt ist. Aber auch in der Forschung ergeben sich durch die Tech-

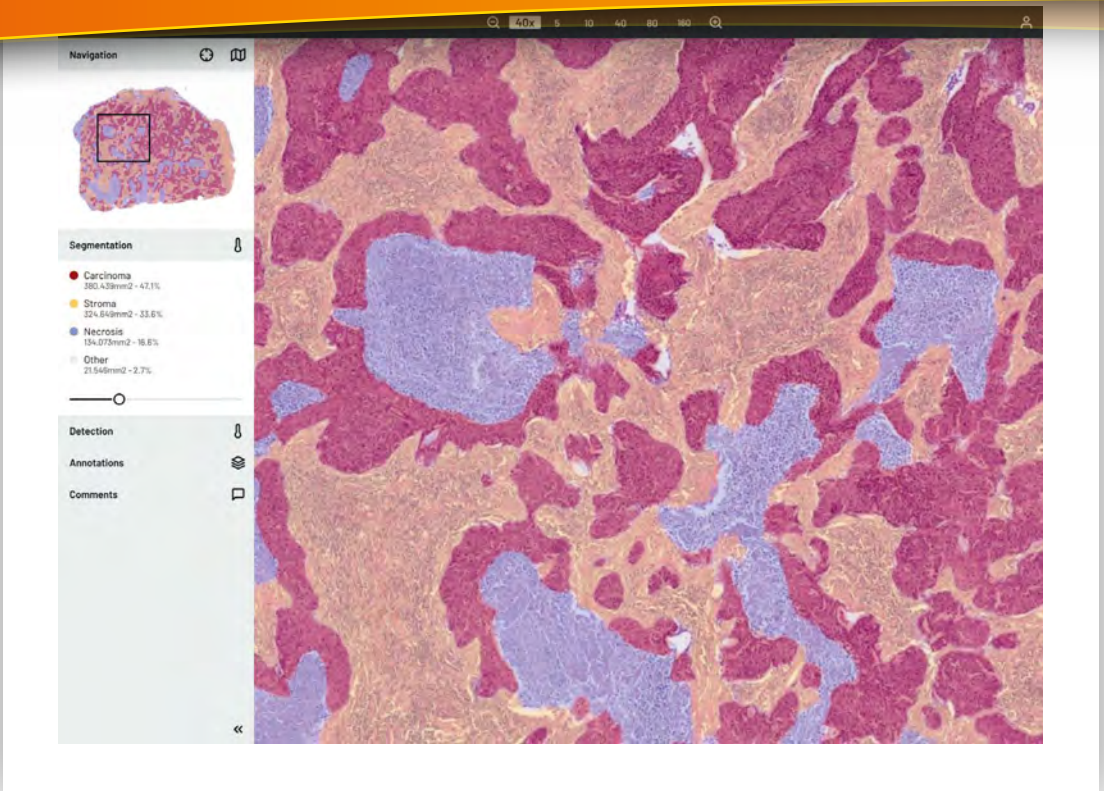


Abbildung 1: Aignostics Portal (Quelle: Aignostics GmbH).

nologie ganz neue Möglichkeiten. Trainiert man zum Beispiel die KI mit positiven und negativen Verläufen einer bestimmten Therapie, können die resultierenden „Heatmaps" den Pathologen ermöglichen, neue Merkmale („Biomarker") zu entdecken, die den Therapieerfolg vorhersagen könnten. „Diesen Ansatz nennen wir „Explainable AI" oder erklärbare künstliche Intelligenz", so Müller.

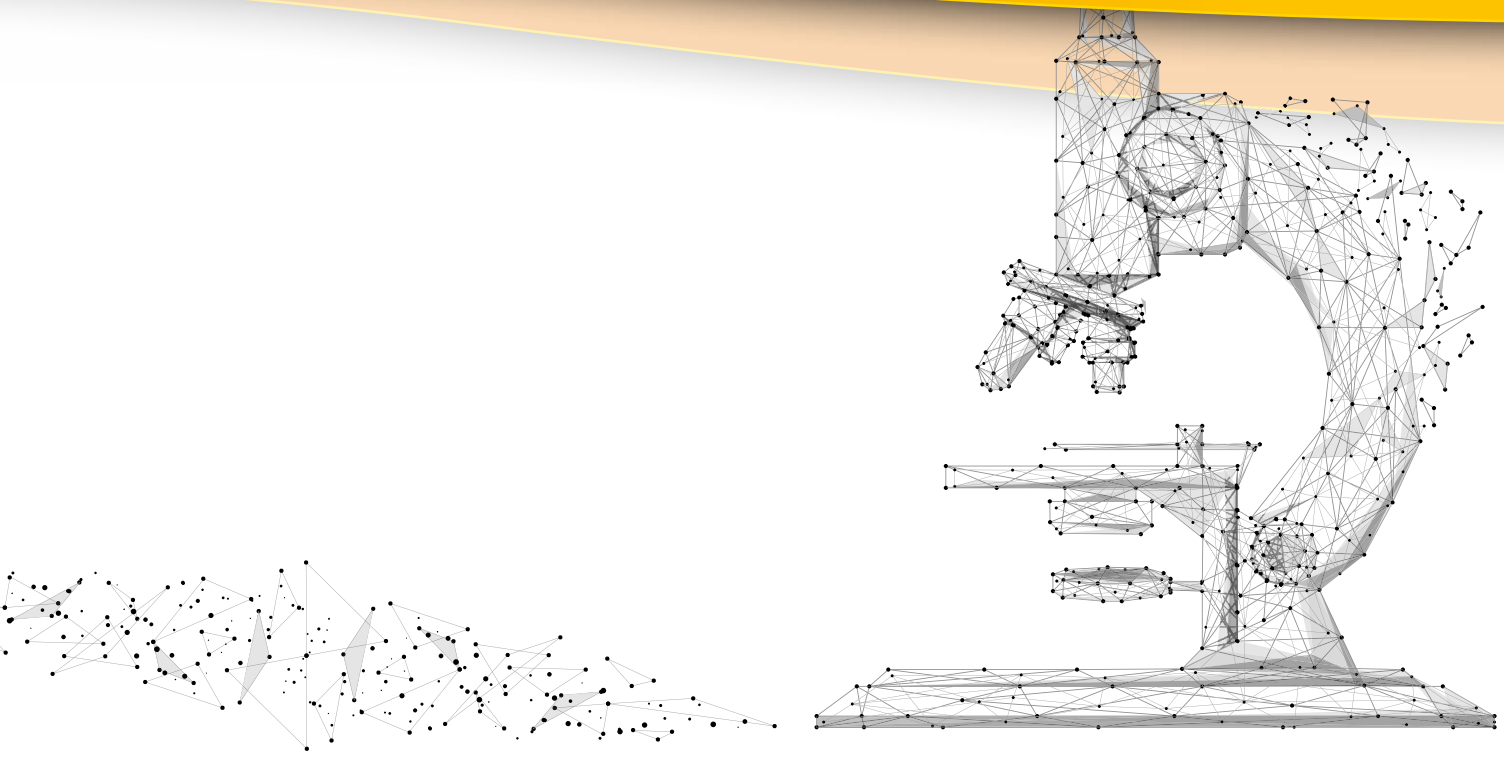
In der Routinediagnostik soll das KI-Verfahren nicht nur Zeit sparen und dabei helfen, Fehler zu vermeiden, sondern auch den Weg für die personalisierte Medizin ebnen. Immer mehr Therapieentscheidungen benötigen den genauen Nachweis und die Quantifizierung von bestimmten Eigenschaften der Gewebeproben. Am Institut für Pathologie an der Charité wird die entwickelte Software bereits in der Diagnostik eingesetzt. Der Einsatz in anderen Instituten und Praxen soll folgen, die Zertifizierung der Software für die breite Anwendung ist in Arbeit.

Auch für Forschung und Medikamentenentwicklung geeignet

„Unsere Software ist auch für die aufwändigen Zulassungsverfahren für neue Medikamente geeignet", erklärt Frederick Klauschen. „Wenn etwa ein Pharmaunternehmen eine neue Substanz entwickelt, müssen umfangreiche Studien, teilweise in mehreren Ländern, sowohl beim Tier als auch beim Menschen zeigen, dass das Medikament gut wirkt und ein akzeptables Nebenwirkungsprofil hat. Dabei sind pathologische Untersuchungen eine wichtige Komponente. Und gerade hier kann unsere Software natürlich helfen." Erste Kooperationen mit verschiedenen Herstellern bei der Entwicklung von Wirkstoffen sind bereits angelaufen. Frederick Klauschen ergänzt: „Unser Ziel ist es auch, so genannte „Companion Diagnostics" zu entwickeln". Darunter versteht man die Kombination eines zielgerichteten Medikaments mit dem Nachweis des passenden Zielmoleküls, etwa auf einem Tumor: Der Arzt darf das Medikament nur dann verschreiben, wenn das passende Zielmolekül auch vorhanden ist. „Wenn dieser Nachweis mithilfe unserer Software schneller und präziser gelingen könnte, wäre das natürlich eine tolle Sache, sowohl für den Patienten als auch für die Hersteller."

Digital Health Accelerator des BIH unterstützt Entwicklung digitaler Produkte

Von der Idee bis zur Ausgründung ist es jedoch ein langer Weg. Deshalb bietet das BIH genau an dieser Schnittstelle zwischen Forschung und Anwendung Hilfe mit seinem Digital Health Accelerator an. „Das BIH unterstützt Innovator*innen aus der Charité und dem Max-Delbrück-Centrum dabei, aus ihren Konzepten digitale Produkte zu entwickeln und diese in die medizinische Anwendung zu überführen. Dies kann neben Lizensierung oder Industriekooperation auch durch Ausgründung erfolgen und so Arbeitsplätze schaffen", erklärt Tim Huse, der bei BIH Innovations den Digital Health Accelerator leitet. „Bei diesem Projekt ging es insbesondere darum, das Team zu komplettieren, eine Unternehmensstrategie sowie das Produkt zu entwickeln und die Ausgründung vorzubereiten." Die Unterstüt-



Grafik: shutterstock / anttoniart

zung durch das Digital Health Accelerator-Programm und des Teams von BIH Innovations reicht von finanzieller Förderung über Coaching und Mentoring durch Experten, Netzwerkzugang zu Talenten, Entwicklungspartnern, Industrie und Investoren bis zur Bereitstellung eines Co-Workingspace für Innovation und Translation und einer engen Begleitung von Ausgründungen.

Forschung und Routinediagnostik

Das Team von Aignostics besteht momentan aus 15 Mitarbeiter*innen sowie einem Netzwerk aus über 20 Pathologen und Forschungspartnern, welche die Entwicklung unterstützen und die Software laufend testen. „Wir rechnen aber damit, unser Team im Laufe des Jahres noch einmal signifikant zu vergrößern“, erklärt der Geschäftsführer Viktor Matyas. „Damit sind wir in der Lage, die bestehenden Anwendungen auszubauen sowie neue zu entwickeln”, ergänzt der technische Leiter Dr. Maximilian Alber. Dabei ist die Nähe zum klinischen Routinebetrieb Frederick Klauschen besonders wichtig: „Wir richten die Entwicklung unserer Lösung ganz klar an den Anforderungen der Routinediagnostik und klinischen Medizin aus. Da ich unsere Software täglich nutze, können wir Verbesserungspotentiale sofort erkennen. Wir hoffen, so den Patient*innen am besten zu helfen.“

Firmenprofil:

Aignostics ist eine Ausgründung der Charité – Universitätsmedizin Berlin und entwickelt KI-basierte Lösungen für die Pathologie. Bereits 2011 legte ein Ärzteteam rund um Prof. Dr. Frederick Klauschen gemeinsam mit Forschern der Fraunhofer Gesellschaft sowie der TU Berlin, unter der Leitung von Prof. Dr. Klaus-Robert Müller, mit einem ersten Patent den Grundstein für die Pathologie-Software. Anfang 2018 wurde das Team schließlich in das Digital Health Accelerator (DHA) Programm des Berlin Institute of Health aufgenommen, Anfang 2020 erfolgte die Ausgründung. Ein besonderer Fokus liegt auf der Lösung des oftmals kritisierten „Black-Box“-Problems von KI in der Pathologie, für das Aignostics eine proprietäre „Explainable AI“-Plattform entwickelt hat.

Kontakt:

Prof. Dr. Frederick Klauschen
Lead Medical Adviser | Aignostics GmbH, Berlin
info@aignostics.com

www.aignostics.com

Foto: privat / Frederick Klauschen

alte datens(ch)ätze zu neuem leben erwecken

Künstliche Intelligenz ermöglicht, neue Informationen aus vorhandenen Daten zu gewinnen

von Jakob Simeth, Marian Schön, Michael Huttner, Paul Heinrich, Michael Altenbuchinger und Rainer Spang

Gewebe bestehen aus vielen verschiedenen Zellen mit den unterschiedlichsten Funktionen. Moderne Einzelzell-RNA Expressionsdaten haben uns erstmals gezeigt, was die einzelnen Zellen eines Gewebes tun und wie sie sich in einem Tumor verändern. Diese Daten sind jedoch neu und klinische Verläufe fehlen noch – denn diese brauchen Zeit. Gleichzeitig gibt es aber schon umfangreiche Sammlungen von Genexpressionsprofilen der Gewebe (Bulk) mit klinischen Verläufen, die jedoch nicht zellulär aufgelöst sind. In unserem Projekt TissueResolver entwickeln wir Tools, die von Einzelzell-Daten lernen, wie man aus Bulkprofilen die Zellzusammensetzung des Gewebes und zelltypspezifische Expression rekonstruiert.

Die Sequenzierung und Quantifizierung von RNA auf Einzelzellebene (single cell RNA sequencing, scRNA-seq) entwickelt sich rasend schnell: Als vor ein paar Jahren die Einzelzellsequenzierung aufkam, konnten nur wenige hundert Zellen aus einem Gewebe isoliert und anschließend sequenziert werden. Inzwischen erlauben neue Microdroplet Technologien die gleichzeitige Sequenzierung von Millionen von Zellen in einem Durchlauf. Die Technologie kann damit Auskunft über die Zusammensetzung und Funktion eines Gewebes geben. So wie bei der Durchflusszytometrie ist es möglich, Zellen bestimmten Typen und Subtypen zuzuordnen und deren Anzahl im Gewebe zu zählen. Das verrät bereits viel über den Zustand des Gewebes; man kann so beispielsweise herausfinden, ob die Anzahl infiltrierender T-Zellen in einem Tumor die Sterblichkeit beeinflusst, oder ob sich deren Anzahl nach Immuntherapie verändert hat. Darüber hinaus ermöglicht scRNA-seq von Geweben außerdem die Genexpression von nur einem Teil aller Zellen getrennt zu analysieren. Dies ist durchaus relevant: Sieht man in einem Bulkprofil eine Hochregulation von Genen des Zellzyklus spricht dies für proliferierende Zellen. Es macht jedoch einen großen Unterschied, ob die Tumorzellen wachsen oder ob unser Immunsystem gerade mehr T-Zellen produziert, die den Tumor bekämpfen.

Solche Fragen werden moderne scRNA-seq-Verfahren beantworten – es gibt da nur einen Haken: Diese Technologie ist noch sehr neu und sehr teuer. Bislang sind nur wenig Daten verfügbar. Andererseits gibt es abertausende Datensätze von Bulkgewebe, Sequenzdaten von vielen verschiedenen Geweben und Tumoren. Diese Daten wurden über die letzten beiden Jahrzehnte frei und öffentlich zugänglich gemacht. Bis scRNA-seq-Daten in ähnlichem Umfang vorliegen, wird es noch etwa ein Jahrzehnt dauern.

Der Nutzen von Bulk-Genexpression ist jedoch beschränkt. Aus der Überlagerung der Expression von Millionen von Zellen sind zelltypspezifische Signale nicht einfach herauszufiltern. So ähnlich wie es für uns Menschen schwierig ist, einzelne Stimmen aus dem Stimmengewirr einer großen Menschenmenge herauszufiltern. Es braucht also spezialisierte Software-Werkzeuge, um zelltypspezifische Signale, etwa von spezifischen Immunzellen, vom Expressionsgewirr des Gewebes verlässlich zu trennen. In unserem Projekt wollen wir diese Lücke schließen und solche Werkzeuge für die Wissenschaftsgemeinschaft frei zur Verfügung stellen.

Einzelzelldaten sind der Schlüssel, um Gewebedaten zu verstehen

Um existierende Gewebedaten *in silico* zellulär aufzulösen, benutzen wir die wenigen Einzelzelldatensätze, die es schon gibt.

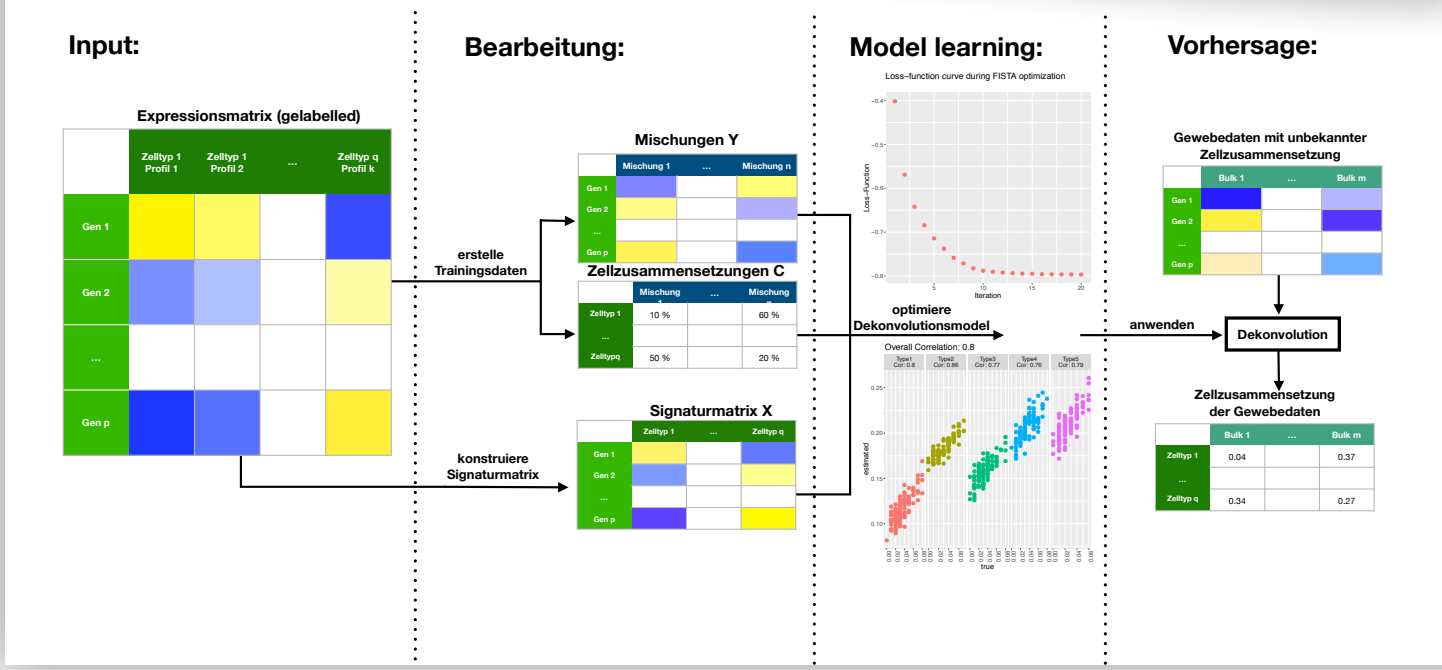


Abbildung 1: Schematische Darstellung des Dekonvolutionsprozesses. Aus Einzelzelldaten bekannten Zelltyps werden einerseits Mischungen bekannter Zusammensetzung und gleichzeitig eine Signaturmatrix erzeugt. Die Gewichte der einzelnen Gene werden daraufhin durch maschinelles Lernen so bestimmt, dass die aus der Simulation bekannte Zusammensetzung der künstlichen Bulks am besten rekonstruiert werden kann. Schließlich können mit dem so bestimmten Modell auch unbekannte Gewebe analysiert werden (Quelle: Marian Schön / Schön *et al., 2020*).

Die Grundidee dabei ist einfach: Wir kombinieren die Genexpression von Zellen, deren Typ wir kennen, so dass sie einem simulierten Gewebebulk von bekannter Zusammensetzung so ähnlich wie möglich sind. Das wird für viele Zellen und Gewebe wiederholt, bis wir wissen, wie Einzelzelldaten kombiniert werden müssen, um einem Gewebe „ähnlich“ zu sein. Mit diesem Wissen können wir dann das Problem umkehren und uns direkt die Gewebeexpression von unbekannter Zusammensetzung ansehen und darin die Anteile bestimmter Zelltypen schätzen. Uns interessiert zunächst: Welche Zellen kommen in einem Gewebe vor und in welcher Menge? Die Beantwortung dieser Frage wird Digital Tissue Deconvolution (DTD) genannt und dafür ist es wichtig, die charakteristischen Merkmale in der Genexpression eines Zelltyps einzufangen. Diese „Signatur“ eines Zelltyps bestimmen wir mit Hilfe von Machine Learning Methoden. Einen ersten Algorithmus dazu haben wir publiziert (Görtler *et al.*, 2020) und zugehörige Software frei zugänglich gemacht (Schön *et al.*, 2020).

Gewebe-Dekonvolution und die Rolle von maschinellem Lernen

Die Genauigkeit von DTD hängt vom Design der Signaturmatrix ab. Ganz so wie unser Gehirn sich auf bestimmte, ihm bekannte Stimmen mit Hilfe deren Charakteristika konzentrieren kann und andere Stimmen dabei ausblendet, müssen auch in diesem Fall jene Gene ausgewählt werden, die besonders charakteristisch für einen Zelltyp sind und andere ausgeblendet werden, die allgemeinere Aufgaben in Zellen kodieren. Von diesen ist keine Information über den Zelltyp zu erwarten. Beispielsweise müssen die genetischen Programme zur Zellteilung von allen Zellen abrufbar sein und eine hohe Expression dieser Gene gibt deshalb keinen Aufschluss darüber, um welchen Zelltyp es sich handelt. Bei der Auswahl solcher Signaturgene muss man also etwas Vorsicht walten lassen: Einerseits dürfen nicht zu wenige Gene ausgewählt werden, denn die Dekonvolution ist dann anfällig für Messungenauigkeiten und technische Artefakte. Andererseits werden die Ergebnisse auch ungenau, wenn zu viele Gene mit einbezogen werden, denn dann sind die schwachen Signale von kleinen Zellpopulationen nicht sichtbar. Die richtige Balance kann durch maschinelles Lernen gefunden werden: Wir weisen den Genen das Gewicht zu, das wir aus den Daten selbst gelernt haben, d. h. wir optimieren die Gewichte so, dass Trainings- und Validierungsdaten bekannter Zusammensetzung am besten in ihre Zellinhalte aufgespalten werden können. Dadurch werden automatisch nur jene Gene ausgewählt, die besonders viel Auskunft über den Zelltyp geben. Das führt zu einer guten Dekonvolutionsleistung und verbessert die Detektion seltener Zellen deutlich (vgl. Abbildung 1, Goertler *et al.*, 2020).

Virtuelle Gewebemodelle

Wir gehen im Design der Signaturmatrix noch einen Schritt weiter. Wir haben Daten vieler Einzelzellen und wissen *a priori* nicht, was wir in einem Gewebe finden werden. Wir entwickeln deshalb Algorithmen, die aus geeigneten Einzelzelldaten diejenigen aussuchen, die, wenn man sie aufaddiert, der zu analysierenden Gewebeexpression möglichst ähnlich sind. Damit nutzen wir das volle Spektrum existierender Zellen im Menschen und müssen diese nicht einmal einem bekannten Zelltyp zuordnen. Das Resultat bezeichnen wir als „virtuelles Gewebemodell“: Eine begrenzte Anzahl von repräsentativen Zellen, denen ein Gewichtsfaktor zugeordnet ist, der in etwa der Häufigkeit ähnlicher Zellen im analysierten Gewebe entspricht. Dieses Ensemble aus Zellen ist also nah an dem, was man sonst mit neuesten Hochdurchsatz scRNA-seq Technologien erhielte.

So ist es möglich, zwei Gewebe über ihre Modelle zu vergleichen, indem man zur Visualisierung der Einzelzellen ein Verfahren zur Dimensionsreduktion nutzt und die identifizierten Zellen beider Gewebemodelle markiert (vgl. Abbildung 2). Wenn man dann jeder Zelle noch einen Typ zuweist – entweder aus dem Originaldatensatz oder durch geeignetes Clustering – werden Unterschiede deutlich, die darüber hinausgehen, was durch die zelluläre Zusammensetzung allein erklärt werden könnte. So kann man zum Beispiel untersuchen, wie sich die B-Zellen im Tumorgewebe von denen im peripheren Blutkreislauf unterscheiden, auch wenn keine Einzelzellmessungen aus den Geweben zur Verfügung stehen.

Zelltypspezifische Expression aus virtuellen Geweben

Seit Jahrzehnten vergleichen wir die Expression von Genen in Geweben. Dabei wissen wir nie, ob Unterschiede auf Genregulationen beruhen oder nur eine unterschiedliche zelluläre Zusammensetzung der Gewebe widerspiegeln. Selbst wenn sie auf Genregulationen beruhen, wissen wir nicht, in welchen Zellen des Gewebes diese Regulation stattfand. Proliferationssignale können aus Tumorzellen stammen oder auf eine Immunantwort im Tumor hinweisen. Wir möchten diesen Unterschied im Nachhinein sichtbar machen. Dazu müssen wir die Expressionsmuster einzelner Zellen im Gewebe entwirren und wieder benutzen wir Einzelzellendaten und Gewebemodelle, um zu lernen wie das geht. Mit diesem Trick lassen sich im Nachhinein völlig neue Fragen beantworten: Wie unterscheidet sich die Genexpression von T-Zellen aus Tumorpatienten mit guter Prognose von jenen mit schlechter? Wie stark exprimieren B-Zellen bestimmte Gene, wenn sie den Tumor infiltrieren? Erzeugen die Tumor- oder die Immunzellen, die ihn bekämpfen, die Proliferationssignale?

Gewebenormalisierung ermöglicht besseren Blick auf den Tumor

Häufig stehen natürlich nicht Einzelzelldaten für dieselben Zellen zur Verfügung, wie sie sich im Gewebe befinden. Wenn zum Beispiel in der Einzelzellreferenz nur Immunzellen verfügbar sind und das zu analysierende Gewebe aus Melanomen stammt, kann auch nur der Immunzellanteil des Gewebes detailliert erklärt werden. Das mag zunächst wie ein Nachteil erscheinen, aber tatsächlich kann man den Tumor so trotzdem besser verstehen. Indem wir die durch die vorhandenen Zellen erklärte Expression von der Gesamtexpression entfernen, entsteht ein klareres Bild. So können wir zum Beispiel viel der Variation in den Daten erklären, die einfach durch die zufällige Wahl des Ortes, wo im Tumor Gewebe entnommen wurde, entsteht. Durch dieses Entfernen von einfach erklärbaren Expressionsunterschieden schaffen wir gleiche Bedingungen und können die zurückbleibenden Daten besser miteinander vergleichen.

Datens(ch)ätze zu neuem Leben erwecken

Sequenzierte Gewebe sind öffentlich und in großem Umfang verfügbar und je mehr Einzelzellmessungen zugänglich sind, umso mehr können wir – mit unseren neuen Werkzeugen – auch aus diesen alten Daten lernen. Das ist besonders wichtig, weil viele dieser Daten bereits vor Jahren erhoben wurden und dadurch bereits klinische Follow-up Information verfügbar ist, d. h. die Daten können mit einer Prognose verknüpft werden. Wir gehen davon aus, dass es noch ein Jahrzehnt dauern wird, bis ähnliche Aussagen mit Einzelzellmessungen vorliegen. Auch für die klinische Praxis und personalisierte Medizin werden Gewebemessungen auf absehbare Zeit nicht ersetzt werden können. Einzelzelldiagnostik ist schlicht zu aufwendig für die klinische Routine.

Umso wichtiger ist es, den vorhandenen Datenschatz zu heben und Werkzeuge zu entwickeln, die dabei hilfreich sind. Der Entwicklung solcher Software widmen wir uns in unserem Projekt und veröffentlichen sie frei und für jeden verfügbar. Einen Anfang hat dabei unser Tool zur Gewebedekonvolution gemacht (DTD), das als Paket für die Programmiersprache „R“ zur Verfügung steht (Schön *et al.*, https://github.com/spang-lab/DTD). Parallel können auch fertige Modelle unkompliziert mit einem Webtool verwendet werden (https://dtd.spang-lab.de). Diese Werkzeuge erweitern wir im Laufe der Zeit immer wieder, so dass auch unsere neuen Algorithmen und Entwicklungen bald Anwendung finden und die wertvollen existierenden Datensätze weiter und tiefer analysiert werden können. Wir hoffen, dass

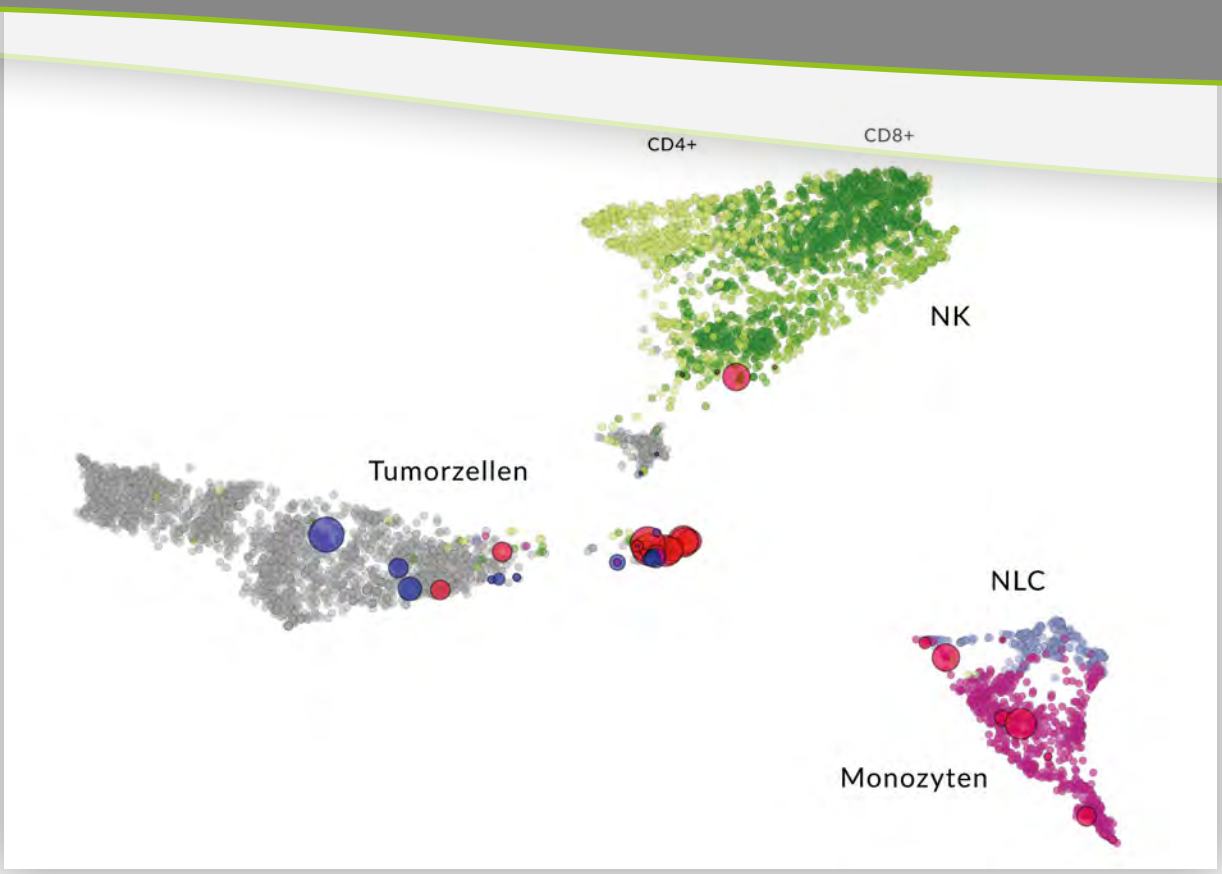


Abbildung 2: Darstellung von Einzelzellprofilen und Bulkgewebe von Patienten mit chronischer lymphatischer Leukämie. Kleine Punkte entsprechen einzelnen Zellen und ihre Farben visualisieren den Zelltyp (grüner Cluster oben: Immunzellen, violetter Cluster unten rechts: Monozyten, graue Cluster links und in der Mitte: Tumorzellen). Mit diesen Einzelzellprofilen wurden zwei Bulkgewebe analysiert und repräsentative Zellen gefunden, die den Bulks am besten ähneln. Repräsentative Zellen der ersten Gewebeprobe sind in blau dargestellt, die der anderen Probe in rot. Die Größe der roten und blauen Kreise entspricht der Häufigkeit ähnlicher Zellen). Die beiden Gewebeproben unterscheiden sich deutlich: Während der Tumor des ersten (blauen) Patienten vor allem im großen Cluster links vertreten ist, ähnelt der Tumor des zweiten (roten) Patienten mehr den (ebenfalls malignen) Zellen in der Mitte. Auch unterschiedliche Häufigkeiten und Expressionsmuster bei den Immunzellen und Monozyten werden deutlich (Quelle: Jakob Simeth).

hiervon eine breite Gemeinschaft aus Wissenschaftlern – in Tumorbiologie, Immunologie und vielen mehr – profitiert und daraus neue Erkenntnisse gewinnen kann.

Steckbrief Forschungsprojekt:

„TissueResolver“ ist ein vom Bundesministerium für Bildung und Forschung gefördertes Projekt (FKZ 031L0173), das 2019 am Institut für Statistische Bioinformatik an der Universität Regensburg startete. Ziel ist die Algorithmenentwicklung und Bereitstellung von interaktiven Werkzeugen, um Bulkgewebedaten besser untersuchen zu können.

Referenzen:

Schön, M., Simeth, J., Heinrich, P., Görtler, F., Solbrig S., Wettig T., Oefner, P.J., Altenbuchinger, M., und Spang, R. (2020). DTD: An R Package for Digital Tissue Deconvolution. Journal of Computational Biology, 386-389.

Görtler, F., Schön, M., Simeth, J., Solbrig S., Wettig T., Oefner, P.J., Spang, R. und Altenbuchinger, M. (2020). Loss-Function Learning for Digital Tissue Deconvolution. Journal of Computational Biology, 342-355.

Kontakt:

Prof. Dr. Rainer Spang
Statistische Bioinformatik
Institut für Funktionelle Genomik
Universität Regensburg
rainer.spang@ur.de
www.spang-lab.de

Jakob Simeth
Statistische Bioinformatik
Institut für Funktionelle Genomik
Universität Regensburg
jakob.simeth@ur.de

Fotos: privat / Rainer Spang und privat / Jakob Simeth

big data und smartphones auf der intensivstation

Wie Smartphone und Künstliche Intelligenz dabei helfen, beatmete Patienten zu behandeln

von Gernot Marx, Sebastian Fritsch, Johannes Bickenbach, Julian Kunze, Oliver Maaßen, Saskia Deffge, Silke Haferkamp und Andreas Schuppert

Als im Januar im chinesischen Wuhan die ersten Fälle einer neuen Lungenkrankheit auftauchten, ahnte wohl noch kaum jemand, was bald auf die Welt zukommen würde. Früh zeichnete sich ab, dass Patienten, die besonders schwer erkranken oder sogar versterben, mehrheitlich an einem akuten Lungenversagen (engl. *acute respiratory distress syndrome*; ARDS) erkranken. Dabei ist das Krankheitsbild für Intensivmediziner keineswegs neu. Auch abseits von COVID-19 tötet es etwa 40 % aller Erkrankten. Doch obwohl es eigentlich klare Kriterien für die Diagnosestellung gibt, wird es noch immer zu selten diagnostiziert. Dieser Herausforderung stellt sich der klinische Anwendungsfall ASIC im Rahmen des SMITH-Projektes.

ARDS und Beatmung

Um ein ARDS zu verstehen, muss man sich die Lunge wie einen Schwamm vorstellen, in dem luftgefüllte Räume sich mit feinen Gewebestrukturen abwechseln. Kommt es zu einer Schädigung der Lunge, z. B. durch eine Lungenentzündung, strömt Flüssigkeit in das Gewebe ein: Der Schwamm saugt sich voll, es entsteht ein sog. „Lungenödem". In der Folge, füllen sich luftgefüllte Räume mit Flüssigkeit, die Gewebestrukturen schwellen an und der Gasaustausch verschlechtert sich. Die Lunge wird „nass und schwer" und so die Beweglichkeit des Gewebes deutlich herabgesetzt. Häufig kann dann die Lunge den Sauerstoffbedarf des Körpers nicht mehr decken. Gleichzeitig nimmt die muskuläre Atemarbeit des Patienten soweit zu, dass vor allem alte oder geschwächte Patienten nicht mehr in der Lage sind, diese aufzubringen.

Beides macht letztlich häufig eine maschinelle Beatmung erforderlich, bei der sauerstoffangereichterte Luft über einen Beatmungsschlauch (Tubus) mit Überdruck in die Lunge geleitet wird. Nach dem Ende der „Ausatmung" bleibt ein Überdruck auf dem Lungengewebe – man spricht von *positive endexpiratory pressure* (PEEP) – um so Flüssigkeit aus dem Lungengewebe „herauszudrücken" und wieder mehr Gewebe zu belüften. Mit gleichem Ziel werden beatmete Patienten mit fortgeschrittenem ARDS in Bauchlage gebracht. So gelangen Lungenareale, die sich in Rückenlage durch das Gewicht der nassen Lunge verschlossen haben, nach oben und öffnen sich erneut. Häufig kann dadurch eine Verbesserung des Gasaustausches erreicht werden.

Die maschinelle Beatmung ist häufig lebensrettend. Dennoch hat sie auch Schattenseiten: Durch hohe Beatmungsdrücke und zu groß eingestellte Volumina der jeweiligen Atemzüge kann die Lunge geschädigt werden. Das Gewebe der unelastischen Lunge wird durch die einwirkenden Kräfte überdehnt und verletzt. Dies wiederum setzt Entzündungsprozesse in Gang, welche die Lunge noch weiter schädigen. Es entsteht ein Teufelskreis aus Beatmung, weiterer pulmonaler Verschlechterung und erneuter Intensivierung der Beatmung. Kernziel der Therapie ist also, die

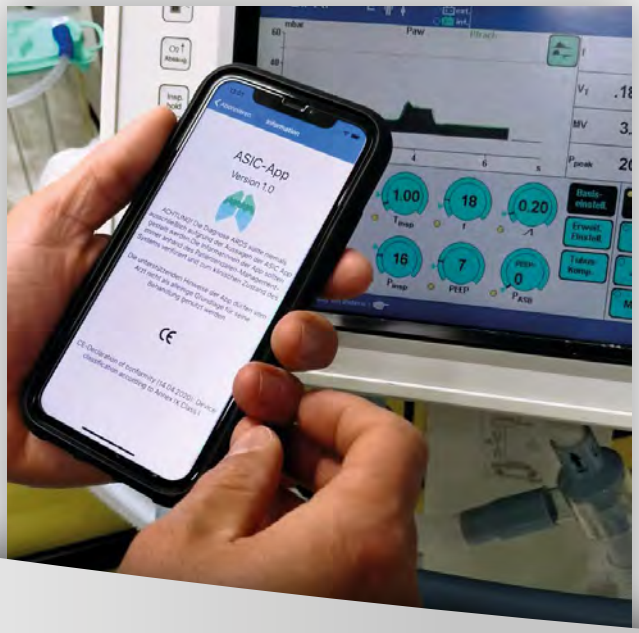


Abbildung 1: Die ASIC App unterstützt die behandelnden Intensivmediziner bei Diagnose und Therapie des ARDS (Foto: Uniklinik RWTH Aachen).

Abbildung 2: Ein erheblicher Teil der ARDS-Erkrankungen wird nicht oder erst zu spät erkannt (Foto: Uniklinik RWTH Aachen).

Beatmung lungenprotektiv, also mit kleinen Atemzugvolumina und niedrigen Spitzendrücken, zu gestalten. Leider sterben trotz Fortschritten in der Therapie noch immer 40 % aller ARDS-Patienten an ihrer Erkrankung (Bellani *et al.,* 2016).

Die „Berlin-Definition“

Das Krankheitsbild ARDS wurde im Jahre 1967 erstmals beschrieben und seither mehrfach unterschiedlich definiert. Die aktuelle Fassung aus dem Jahr 2012 wurde nach dem Tagungsort der Expertengruppe benannt und ist daher als Berlin-Definition bekannt (The ARDS Definition Task Force, 2012). Wesentlicher Teil der Definition ist der sog. Horovitz-Quotient. Er drückt die Fähigkeit der Lunge aus, Sauerstoff ins Blut zu bringen und wird aus dem Sauerstoffpartialdruck im arteriellen Blut und dem Sauerstoffanteil in der Einatemluft berechnet. Er bestimmt auch den Schweregrad des ARDS. Weitere Kriterien für eine ARDS-Diagnose umfassen den zeitlichen Verlauf, vorhandene Veränderungen im Röntgenbild und den Ausschluss anderer Zustände, wie einer Überwässerung oder eines Herzversagens.

Leider wird eine Vielzahl von ARDS-Fällen nicht oder zu spät erkannt. In einer großen Beobachtungsstudie mit über 29.000 Patienten, die auf 439 Intensivstation in 50 Ländern über 4 Wochen beobachtet wurden, wurden nur knapp die Hälfte aller milden ARDS-Fälle erkannt (Bellani *et al.,* 2016). Es ist leider davon auszugehen, dass eine nicht diagnostizierte Erkrankung auch nicht korrekt behandelt wird.

ASIC
Algorithmic Surveillance

SMITH
Smart Medical Information Technology for Healthcare

Die ASIC-App

Hier setzt die ASIC-App an. Sie wird von den Intensivmedizinern an den Universitätskliniken des SMITH-Konsortiums auf dienstlichen Smartphones genutzt. Die App kommuniziert mit dem Patientendaten-Management-System (PDMS) und überprüft regelmäßig, ob der Horovitz-Quotient den Grenzwert für ein ARDS unterschreitet. Ist dies der Fall, wird der Arzt mittels Push-Nachricht über das potenzielle Vorliegen eines ARDS bei einem seiner Patienten informiert und aufgefordert, die übrigen Diagnose-Kriterien zu bewerten. Sind die Kriterien erfüllt, zeigt die App die wichtigsten Empfehlungen zur ARDS-Therapie aus der entsprechenden Leitlinie an. Der Arzt kann somit die Diagnose deutlich früher stellen, die bisherige Therapie überprüfen und ggf. evidenzbasiert anpassen. Die ASIC-App hat vor ihrem Einsatz am Patientenbett das CE-Konformitätsverfahren durchlaufen und wurde als Medizinprodukt der Klasse I eingestuft.

COVID-19

Dass ARDS ein großes Problem bei der Therapie von COVID-19 ist, wurde bereits früh deutlich. Auch nach etwa 20 Mio. Erkrankten scheint offenkundig zu sein, dass ein ARDS das Behandlungsergebnis massiv negativ beeinflusst. Wie viele SARS-CoV-2-Infizierte tatsächlich ein ARDS entwickeln, ist aufgrund der dynamischen Lage und stark schwankender Zahlen bisher nicht klar. Gleichzeitig postulieren erste Autoren, dass sich das ARDS bei COVID-19 von einem „konventionellen“ ARDS unterscheidet (Marini and Gattinoni, 2020). Einige Patienten, so die Autoren im renommierten JAMA, zeigen zunächst zwar eine ausgeprägte Gasaustauschstörung, aber nur ein minimal ausgeprägtes Lungenödem. Das Lungengewebe bliebe damit eher elastisch. Da die etablierten Therapieempfehlungen von einer „nassen“ Lunge mit deutlichem Lungenödem ausgehen, könnten diese Empfehlungen bei einem COVID-ARDS nicht nur ohne Vorteil, sondern u. U. sogar schädlich sein. Um diese Theorie aber anhand klinischer Verläufe zu untersuchen, wären die Daten einer großen Anzahl von Patienten nötig. Intensivpatienten eignen sich hierzu sehr gut, da pro Tag und Patient ca. 1200

Abbildung 3: Die Autoren (Foto: Uniklinik RWTH Aachen)

Datenpunkte gespeichert werden. Leider sind die Systeme, die auf den Intensivstationen deutscher Universitätskliniken verwendet werden, zumeist nicht miteinander kompatibel. Ebenso dürfte die Zahl der intensivpflichtigen COVID-19-Patienten an einer deutschen Universitätsklinik für o. g. statistische Auswertungen eher zu niedrig sein. Die Erstellung eines großen Datenpools anonymisierter Behandlungsdaten von Intensivpatienten, der sich auch für den Einsatz neuer Methoden der Datenwissenschaften eignet, stellt somit eine große Herausforderung dar. Dies ist das zweite wichtige Feld, auf dem ASIC Innovation herbeiführen und so Patientenversorgung verbessern will.

Datenwissenschaften und Datenintegrationszentren

Künstliche Intelligenz, Big Data-Analyse, Machine Learning und Data Science werden große Bedeutung für die Zukunft der Medizin zugesprochen, da sie es ermöglichen, auch bei sehr komplexen, nicht vollständig verstandenen Zusammenhängen noch gute Prognosen aus Daten zu extrahieren. So ist es z. B. schon heute möglich, anhand von kleinen, für Menschen unmerklichen Veränderungen der Vitalparameter das Auftreten eines septischen Schocks mehrere Stunden im Voraus vorherzusagen (Ghalati *et al.,* 2019). Hinzu kommen systemmedizinische Modelle, sogenannte Virtuelle Patienten, die auf Basis medizinisch etablierter Prozesse die Entwicklung von komplexen Erkrankungen beschreiben. Für die Entwicklung solcher Modelle sind an die spezielle Situation adaptierte Algorithmen in Kombination mit großen Datenbanken erforderlich. Die aktuell am weitesten verbreitete Datenbank für Intensivpatienten, die MIMIC-III-Datenbank, enthält aber ausschließlich US-amerikanische Patienten. Ihre Daten sind nur begrenzt auf deutsche Patienten übertragbar. Ziel des Use Cases ASIC ist daher, die Daten der Intensivpatienten aller beteiligten Universitätskliniken zusammenzuführen und für die Forschung nutzbar zu machen. Hierzu sind eine Vielzahl von Herausforderungen zu lösen, die im Rahmen von ASIC angegangen werden. So wurden an jedem beteiligten Standort sog. Datenintegrationszentren (DIZ) geschaffen. Ein solches DIZ extrahiert die Daten aus den Primärsystemen der Krankenversorgung, depersonalisiert sie und überführt sie in ein interoperables Format, um die Daten zusammenführen zu können. Durch die Analyse dieser Daten können dann Unterschiede zwischen bestimmten Patientengruppen, wie etwa mit COVID-19-ARDS und konventionellem ARDS, erkannt werden. Die erfolgreichsten Therapiestrategien können so identifiziert und künftige Therapien darauf angepasst werden. Mithilfe dieser im Entstehen begriffenen Datenbank soll in ASIC ein Frühwarnsystem für das ARDS, analog zum bestehenden System für den septischen Schock, sowie ein Virtuelles Patientenmodell für ARDS entwickelt werden. Letzteres kann die Wirkung bestimmter therapeutischer Maßnahmen unter bestimmten Rahmenbedingungen vorhersagen, ohne dies mithilfe aufwendiger klinischer Studien und Versuchen an Patienten überprüfen zu müssen.

Datenschutz

Um diese ambitionierten Projekte umsetzen zu können, wurde ein besonderer Schwerpunkt auf ein datenschutzkonformes Vorgehen gelegt. So mussten für die Nutzung medizinischer Da-

ten neben übergreifend gültigen Vorgaben, wie z. B. der Datenschutzgrundverordnung (DSGVO), auch die unterschiedliche Gesetzgebung aus 6 Bundesländern berücksichtigt werden. Hierzu wurden mehrere juristische Gutachten eingeholt, ein umfassendes Datenschutzkonzept durch externe Experten erstellt sowie ein Dialog mit den Datenschutzbeauftragten aller beteiligten Universitätskliniken geführt. Auch die Ethikkommissionen der teilnehmenden Häuser wurden konsultiert und gaben ihre Zustimmung.

Ausblick

Der Use Case ASIC stellt ein Anwendungsbeispiel im Rahmen des SMITH-Medizininformatik-Konsortiums dar. Er macht beispielhaft deutlich, welche Rolle Datenintegrationszentren der Medizininformatik für die Forschung und Verbesserung der Versorgung spielen (s. auch Artikel **Datenintegrationszentrum - Drehscheibe für Daten in der medizinischen Forschung und Versorgung**, Seite 84). Die Fokussierung auf das Krankheitsbild ARDS soll primär die Funktionalität der eingesetzten Verfahren demonstrieren. Sollten sich diese im Rahmen des Projektes bewähren, können sie in Zukunft mit überschaubarem Aufwand auf weitere Krankheitsbilder, wie z. B. die Sepsis oder das akute Nierenversagen ausgeweitet werden. Bis dahin wird die App wohl noch oft den Arzt auf der Intensivstation vor einem möglichen ARDS warnen.

Das SMITH-Konsortium in Kurzform:

Der klinische Anwendungsfall ASIC des SMITH-Konsortiums wird im Rahmen der Medizininformatik-Initiative (MI-I) vom Bundesministerium für Bildung und Forschung (BMBF) gefördert. Im SMITH-Konsortium haben sich 10 deutsche Universitätskliniken mit weiteren Partnern aus Forschung und Industrie zusammengeschlossen, um den Austausch und die Nutzung von gesundheitsrelevanten Daten aus Forschung und Patientenversorgung standortübergreifend in einem interoperablen Format zu realisieren (Winter *et al.,* 2018). An den Universitätskliniken werden dafür DIZ aufgebaut und miteinander vernetzt. Gleichzeitig werden in 2 klinischen und einem methodischen Anwendungsfall innovative IT-Lösungen entwickelt, mit denen die Möglichkeiten moderner digitaler Dienstleistungen und die Funktionsfähigkeit und der Nutzen der neuen Infrastruktur demonstriert werden sollen.

Konsortialpartner von SMITH
(www.smith.care/konsortium)
Universität Leipzig
Universitätsklinikum Leipzig AöR
Friedrich-Schiller-Universität Jena
Universitätsklinikum Jena
Uniklinik RWTH Aachen
RWTH Aachen
Fraunhofer-Institut für Software- und Systemtechnik ISST Dortmund
Bayer AG Leverkusen
März Internetwork Services AG Essen
Averbis GmbH Freiburg
ID GmbH & Co. KGaA Berlin
Forschungszentrum Jülich
Universitätsklinikum Halle (Saale)
Universitätsklinikum Bonn
Universitätsklinikum Hamburg-Eppendorf
Universitätsmedizin Essen
Universitätsmedizin Rostock
Universitätsklinikum Düsseldorf
Verband des Universitätsklinikums der Ruhr-Universität Bochum

Referenzen:

Bellani, G., Laffey, J.G., Pham, T., Fan, E., Brochard, L., Esteban, A., Gattinoni, L., van Haren, F., Larsson, A., McAuley, D.F., *et al.* (2016). Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. Journal of the American Medical Association (JAMA) 315, 788-800.

Ghalati, P.F., Samal, S.S., Bhat, J.S., Deisz, R., Marx, G., and Schuppert, A. (2019). Critical Transitions in Intensive Care Units: A Sepsis Case Study. Scientific reports 9, 12888.

Marini, J.J., and Gattinoni, L. (2020). Management of COVID-19 Respiratory Distress. JAMA 323, 2329-2330.

The ARDS Definition Task Force (2012). Acute Respiratory Distress Syndrome: The Berlin Definition. JAMA 307, 2526-2533.

Winter, A., Staubert, S., Ammon, D., Aiche, S., Beyan, O., Bischoff, V., Daumke, P., Decker, S., Funkat, G., Gewehr, J.E., *et al.* (2018). Smart Medical Information Technology for Healthcare (SMITH). Methods of information in medicine 57, e92-e105.

Kontakt:

Univ.-Prof. Dr. med Gernot Marx, FRCA
Direktor der Klinik für Operative Intensivmedizin und Intermediate Care
Uniklinik RWTH Aachen
gmarx@ukaachen.de

www.smith.care

#nCoVStats

Wie Data Science hilft die Coronavirus-Pandemie zu verstehen

von Matthieu-P. Schapranow

Die durch das Coronavirus Ende 2019 ausgelöste Pandemie bestimmt auch nach mehr als einem halben Jahr das Leben von uns allen. Im Kampf gegen das Virus sind öffentlich zugängliche Daten für Experten und Laien wichtig, um sich unabhängig informieren zu können. Hierbei unterstützt eine Gruppe von Forschern am Hasso-Plattner-Institut für Digital Engineering (HPI) der Universität Potsdam. Sie stellen Analyse-Werkzeuge zur grafischen Auswertung auf der Webseite https://we.analyzegenomes.com sowie täglich aktuelle Daten rund um die Coronavirus-Pandemie und dem Hashtag #nCoVStats Interessierten kostenlos zur Verfügung.

Virus eindämmen durch systematisches Prüfen von Verdachtsfällen

Das Leben in Zeiten grenzenloser Globalisierung bietet zahlreiche Vorteile: Wir können reisen, wohin wir wollen und selbst exotischste Güter sind über das Internet schnell bestellt und treffen binnen weniger Tage in den eigenen vier Wänden ein. Aber auch Viren können sich heute schneller denn je über den Globus verteilen. Dabei helfen ihnen die unzähligen Flugverbindungen, die selbst entlegenste Orte des Planeten in nur wenigen Stunden mit Metropolen verbinden.

Am HPI hat man bereits Erfahrungen bei der Erforschung von Epidemien. Beispielsweise arbeiteten HPI-Forscher bei der Eindämmung der Ebola-Epidemie 2014 in Westafrika gemeinsam mit internationalen Wissenschaftlern. Damals – wie auch heute – ist das Nachverfolgen von Kontakten (*Contact Tracing*) eine wichtige Maßnahme zur Eindämmung des Virus. Dabei werden Kontaktpersonen von Infizierten über den Zeitraum der Inkubation isoliert und regelmäßig auf krankheitsspezifische Symptome befragt. Nur durch konsequentes Identifizieren von Kontaktpersonen und deren Isolation kann die Gefahr der Ansteckung weiterer Personen reduziert werden.

Quelle: CDC, Macau Photo Agency, M. Schapranow

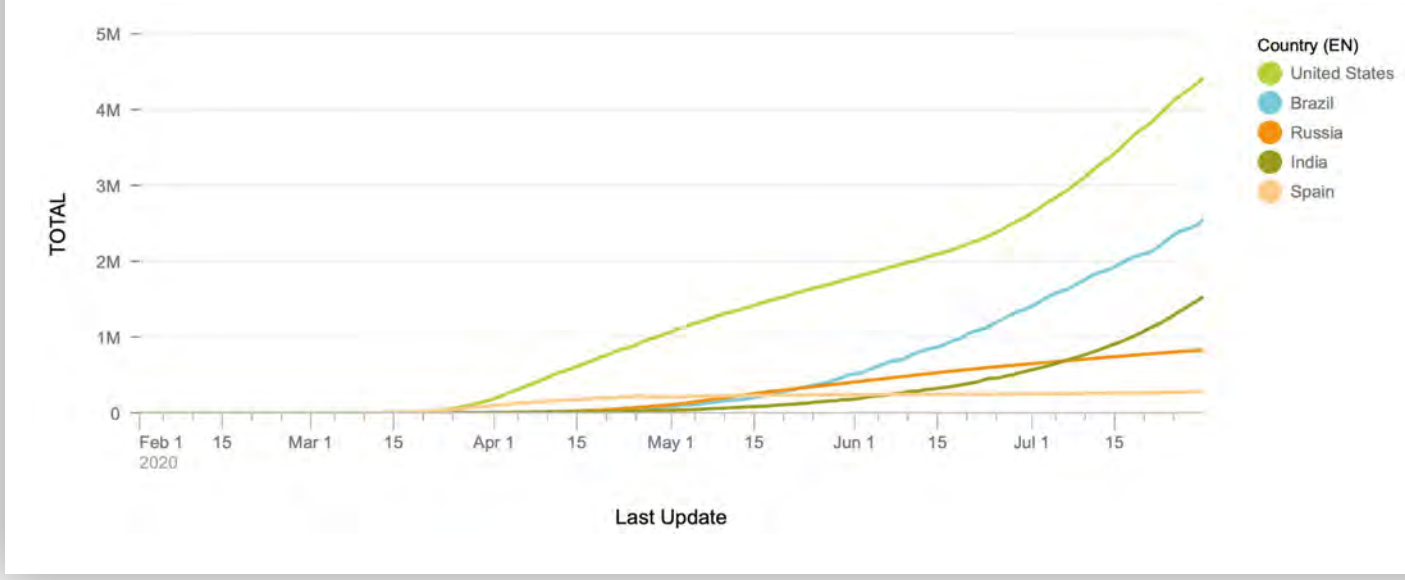


Abbildung 1: Entwicklung der Infektionszahlen der Top-5-Länder (Stand 23. Juli 2020, Quelle: https://we.analyzegenomes.com/apps/ncovstats/).

Auch während der aktuellen Coronavirus-Pandemie ist das Contact Tracing ein Schlüssel zum Erfolg. Gerade bei den ersten Fällen in Deutschland wurde sehr erfolgreich mit persönlichen Interviews versucht zu rekonstruieren, mit welchen Personen Infizierte in den vorangegangenen Tagen in Kontakt gewesen waren. Je konsequenter dies erfolgte, desto schneller wurde auch klar, dass dafür viele Ressourcen nötig waren. Schon bei der Ebola-Epidemie 2014 hatte sich gezeigt, dass qualifiziertes Personal für das Contact Tracing rasch knapp wird. Daher war schon damals am HPI gemeinsam mit einem internationalen Wissenschaftlerteam eine App für das Contact Tracing entwickelt und vor Ort in Nigeria erprobt worden. Nach einer kurzen Schulung konnten auch Nicht-Mediziner die App nutzen und das Contact Tracing unterstützen. Auch 2020 unterstützt das HPI bei der Entwicklung der sogenannten CovApp, die bei der Erfassung relevanter Symptome bei Verdachtsfällen hilft.

Auch diesmal zeigt der Einsatz digitaler Anwendungen, wie in Zeiten knapper Ressourcen, diese effektiver genutzt werden können und so medizinisches Fachpersonal entlastet wird, um Notfälle zu versorgen.

Wichtige Entscheidungsgrundlage: Stets aktuelle Daten

Neben den Daten aus dem Contact Tracing sind Behandlungsdaten aus Krankenhäusern eine wichtige Datenquelle. Da sie aus gesicherten Labortests stammen, stellen Krankenhausdaten qualitativ hochwertige Daten dar. Sie können beispielsweise präzise Auskunft über die Zahl der Neuerkrankten, Gesundeten oder Verstorbenen geben. Doch diese wichtigen Zahlen werden dezentral erhoben und liegen in verschiedenen IT-Systemen vor. Ein zentrales Register, in das die Daten automatisiert ohne Verzögerung erfasst werden, gibt es bislang nicht. Dabei stellen landesweit, aktuelle Daten die Grundlage für viele wichtige Entscheidungen dar. Beispielsweise nutzen Epidemiologen aktuelle Daten über Infizierte je Region, um Einschätzungen zur Ausbreitung und passende Handlungsempfehlungen zu geben. Abbildung 1 zeigt die Entwicklung der Infektionszahlen der Top-5-Länder weltweit. Man erkennt gut, dass die USA, Brasilien und Indien noch immer in einem exponentiellen Wachstum stecken, das heißt ein Infizierter steckt im Schnitt mehrere Personen an. Als Ergebnis steigt der Druck auf das Gesundheitssystem, z.B. bei der Behandlung schwerer Fälle. Im Gegensatz dazu verzeichnen Russland und Spanien nur noch einen linearen Anstieg, d. h. das dortige Infektionsgeschehen verläuft signifikant langsamer und belastet das nationale Gesundheitssystem weniger stark.

Am HPI haben die Forscher den Ernst der Lage frühzeitig erkannt. Schon im Januar 2020 wurde damit begonnen, verfügbare internationale Datenquellen mit Fallzahlen zu SARS-CoV-2 zu identifizieren. Da zu dieser Zeit das Zentrum der Epidemie noch in China lag, konzentrierten sie sich auf chinesische Internetquellen. Die weltweiten Fallzahlen wurden in eine Hauptspeicherdatenbank integriert. Diese am HPI erforschte Datenbanktechnologie ermöglicht es selbst riesige Datenmengen nach beliebigen Kriterien blitzschnell zu analysieren. In der Datenbank werden u. a. die aktuell berichteten Fallzahlen zu Erkrankten, Gesundeten und Verstorbenen je Land oder Region

zusammen mit Zeitstempeln gespeichert. Um die Daten nicht händisch erfassen zu müssen, kommen sogenannte Crawler zum Einsatz. Dabei handelt es sich um Computerprogramme, die in regelmäßigen Abständen die Datenquellen nach aktualisierten Fallzahlen absuchen und aktualisierte Meldungen automatisch in die Datenbank importieren. Auf diese Weise entstand eine longitudinale Datenbank zu den weltweiten Fallzahlen, die mittlerweile etwa 35.000 Einträge für knapp 600 Regionen und Länder weltweit umfasst.

Visualisierung hilft Zusammenhänge zu erkennen

Nach dem Zusammentragen aktueller Daten, ist deren Auswertung der nächste Schritt. So können nicht nur Aussagen zur aktuellen weltweiten Lage getätigt werden, sondern auch retrospektiv Daten analysiert werden, um z. B. Trends in einzelnen Ländern oder Regionen zu erkennen. Hierbei kommen Softwaresysteme zum Einsatz, die durch interaktive Visualisierungen bei der Exploration großer Datenmengen unterstützen. Abbildung 2 zeigt ein Beispiel, das die Fallzahlen vom 20. April und 20. Juli 2020 je Land anhand von Kreisdiagrammen vergleicht. Man erkennt rasch, wie stark die Fallzahlen vor allem in Nord- und Südamerika, aber auch in Teilen Europas und Russlands binnen kurzer Zeit zugenommen haben. Sie übersteigen die Fallzahlen im Ursprungsland China bei weitem, das in der Abbildung nur noch als verhältnismäßig winziger Fleck zu sehen ist.

Unterschiedliche Qualität der Daten

Gleichzeitig sieht man aber auch, dass z. B. der afrikanische Kontinent vermeintlich geringe Fallzahlen meldet. Doch ist das wirklich der Fall? Hier stößt man auf eine weitere Herausforderung: die Datenqualität. Zwar können wir auf gemeldete Daten aus fast allen Ländern zurückgreifen, jedoch haben wir keinen Einfluss auf die Qualität der dort erfassten Daten. Dabei geht es nicht nur um die Korrektheit der übermittelten Zahlen, sondern insbesondere um Definitionen und Annahmen je Land. Anhand welcher Kriterien wird beispielsweise entschieden, ob ein Verdachtsfall als Infizierter gemeldet wird oder nicht? Gerade zu Beginn des Jahres fehlten Kapazitäten für das flächendeckende Testen von Verdachtsfällen. Statt eines PCR-Tests auf Viren-RNA wurden auch andere Indikatoren, wie CT-Bilder der Lunge, zur Fallbestimmung herangezogen. Das unterschiedliche Vorgehen führt aber dazu, dass die gemeldeten Zahlen je Land mit unterschiedlichen Messfehlern behaftet sind.

In afrikanischen Ländern mit einem weniger gut aufgestellten Gesundheitssystem ist das systematische Testen von COVID-19 Verdachtsfällen extrem schwierig. Auch die Dokumentation von Verdachtsfällen und die Akquise der Daten aus regionalen medizinischen Zentren ist für die Regierung mit logistischen Hürden verbunden. Aufgrund der Erfahrungen aus früheren Epidemien ist daher davon auszugehen, dass die öffentlich gemeldeten Zahlen nur einen Bruchteil der Realität abbilden. Hinzukommt, dass glücklicherweise nur ein verhältnismäßig geringer Teil der Infizierten an schweren Symptomen erkrankt, die eine Hospitalisierung erforderlich machen.

Akkurate Prognosen durch Künstliche Intelligenz

Auch in Deutschland wissen wir, dass viele Infizierte mitunter nur leichte oder gar keine Symptome zeigen, also auch nicht durch einen Arztbesuch registriert werden. Um diesen Fehler in nationalen Zahlen zu berücksichtigen, wurden in den Regionen Deutschlands, die als Corona-Hotspots gelten, flächendeckend Einwohner befragt und getestet. Aus diesen regionalen Studien erhofft man sich eine präzisiere Prognose für die realen Fallzahlen in Deutschland und ein besseres Verständnis für die Übertragungswege des Virus zu erhalten. Auch am HPI nutzt man die aufgebaute Datenbank der erfassten COVID-19-Daten als Grundlage für Prognosen. So kommen Verfahren des maschinellen Lernens und der künstlichen Intelligenz zum Einsatz, um beispielsweise anhand der Entwicklungen in China die Fallzahlen für weitere Länder zu prognostizieren oder die Wirksamkeit von getroffenen Maßnahmen zu bewerten.

Einige Dinge konnten wir bereits jetzt aus der Pandemie lernen: Je zeitiger man landesweit auf aktuelle Daten zugreifen kann, desto schneller können angemessene Maßnahmen zur Eindämmung der Pandemie getroffen werden. Etablierte klinische Prozesse, systematisches Testen von Verdachtsfällen, ein zentrales Register zur Erfassung von Verdachtsfällen sowie geeignete IT-Werkzeuge zur interaktiven und flexiblen Auswertung der Daten sind nur einige Bausteine, die unser Gesundheitssystem für künftige Pandemien wappnet. Darum sollten wir diese jetzt systematisch aufbauen, und nicht die Erkenntnisse aus der aktuellen Coronavirus-Pandemie einfach *ad acta* legen.

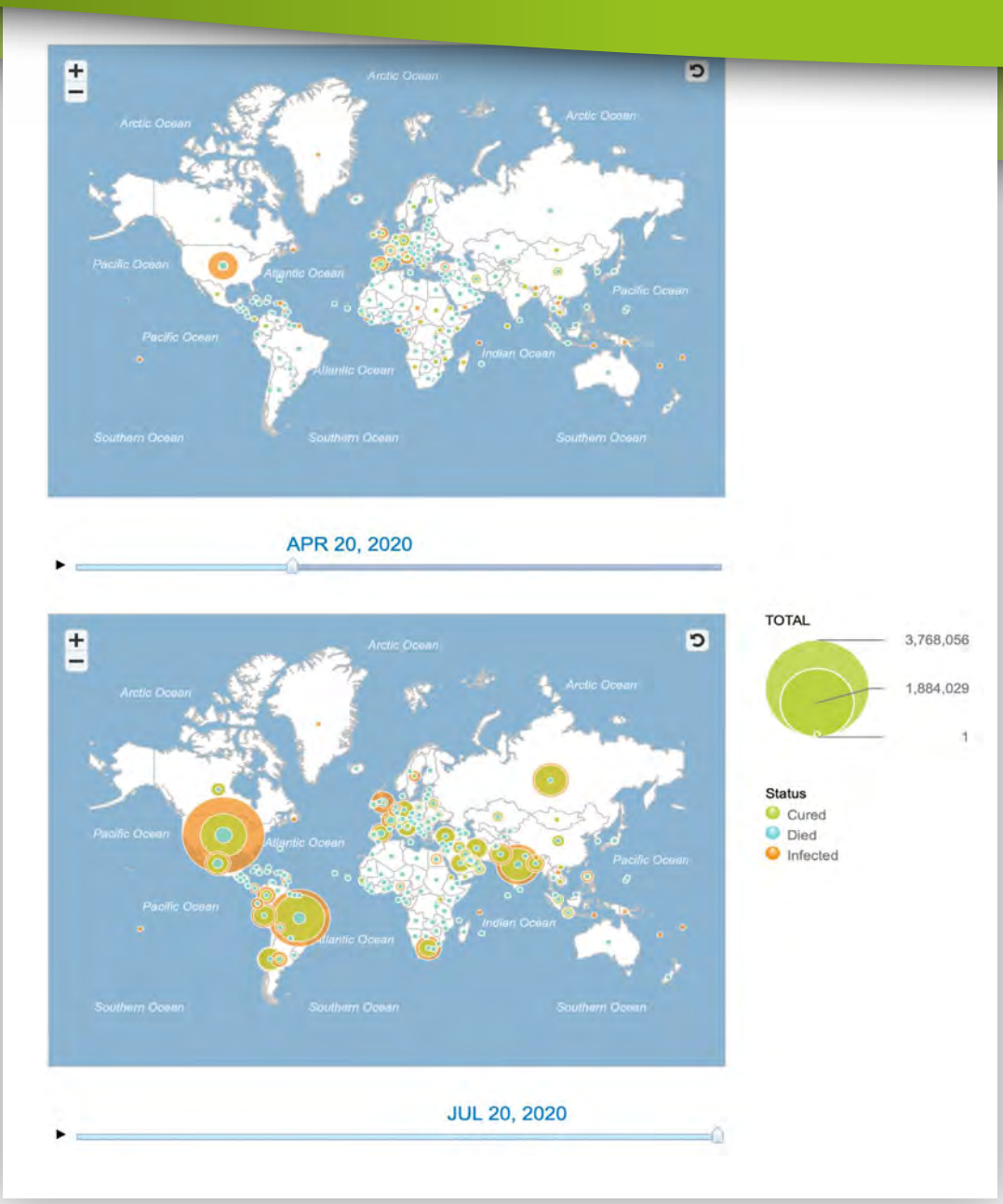


Abbildung 2: Visualisierung der weltweiten COVID-19-Fallzahlen auf einer Weltkarte, oben 20. April, unten 20. Juli 2020. Die Kreisfläche spiegelt die absoluten Zahlen wider (Quelle: https://we.analyzegenomes.com/apps/ncovstats/).

Steckbrief Forschungsprojekt:

Das Projekt **#nCoVStats** wurde Anfang 2020 am Digital Health Center des Hasso-Plattner-Instituts in Potsdam ins Leben gerufen, um stets aktuelle Zahlen über die Virusausbreitung zusammen mit grafischen Analysewerkzeugen der interessierten Öffentlichkeit, politischen Entscheidungsträgern, Pressevertretern, sowie Wissenschaftlern über das Internet kostenfrei zugänglich zu machen.

Kontakt:

Dr.-Ing. Matthieu-P. Schapranow
Leiter der Arbeitsgruppe
„In-Memory Computing for Digital Health“
Scientific Manager Digital Health Innovations
HPI Digital Health Center,
Hasso-Plattner-Institut, Universität Potsdam
schapranow@hpi.de

https://hpi.de/digital-health-center/members/working-group-in-memory-computing-for-digital-health/dr-ing-matthieu-p-schapranow.html

Foto: privat / Schapranow

die chronic kidney disease nephrologist's app

Personalisierte Systemmedizin für Patienten mit chronischem Nierenleiden

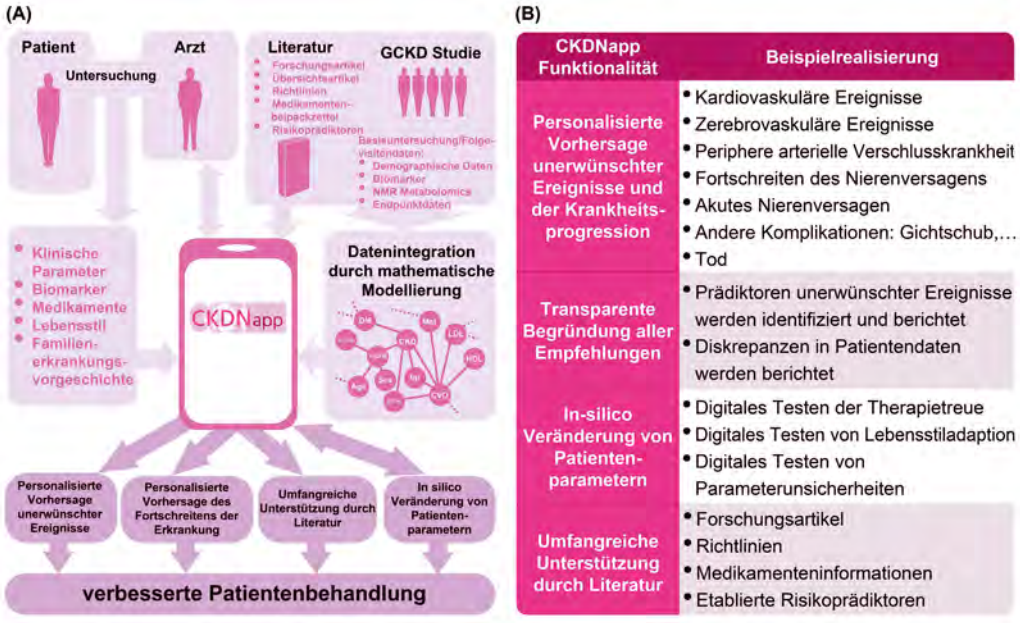
von Ulla T. Schultheiß, Johannes Raffler, Sahar Ghasemi, Robin Kosch, Michael Altenbuchinger und Helena U. Zacharias

Die Niere spielt eine Schlüsselrolle in der Regulation zahlreicher systemischer Prozesse im menschlichen Körper. Die chronische Niereninsuffizienz, eine der häufigsten Todesursachen weltweit, ist eine komplexe Erkrankung, die durch einen sehr individuellen Krankheitsverlauf und zahlreiche Begleiterkrankungen gekennzeichnet ist. Dies erschwert die Vorhersage von unerwünschten Ereignissen sowie des Krankheitsverlaufs der Betroffenen, die Behandlungsplanung und das Medikationsmanagement. Der e:Med Juniorverbund „CKDNapp" hat sich das Ziel gesetzt, eine App für Nephrologen zu entwickeln, um sie bei der personalisierten Behandlung von Patienten mit chronischer Niereninsuffizienz zu unterstützen.

Transparente Entscheidungsunterstützung für praktizierende Nephrologen

Der aktuelle Zustand sowie der zu erwartende Krankheitsverlauf eines Patienten mit chronischer Niereninsuffizienz (engl. Chronic Kidney Disease, Abk. CKD) hängen von zahlreichen demographischen, krankheitsgeschichtlichen, Lebensstil- und Medikationsparametern ab. Über den zukünftigen Patientenzustand können traditionelle, aber auch neuartige Biomarker Aufschluss geben. Um eine optimale Versorgung betroffener Patienten gewährleisten zu können, muss der behandelnde Arzt all diese unterschiedlichen und komplexen Daten auf Basis medizinischen Wissens gemeinsam bewerten und integrieren. Nur wenn dies gelingt, kann die Krankheitstherapie auf den jeweiligen Patienten zugeschnitten, das heißt personalisiert werden.

Abbildung 1: (A) CKDNapp – Schematischer Workflow der Entwicklung und Anwendung und (B) CKDNapp – detaillierte Funktionen



Quelle: Michael Altenbuchinger und Helena U. Zacharias

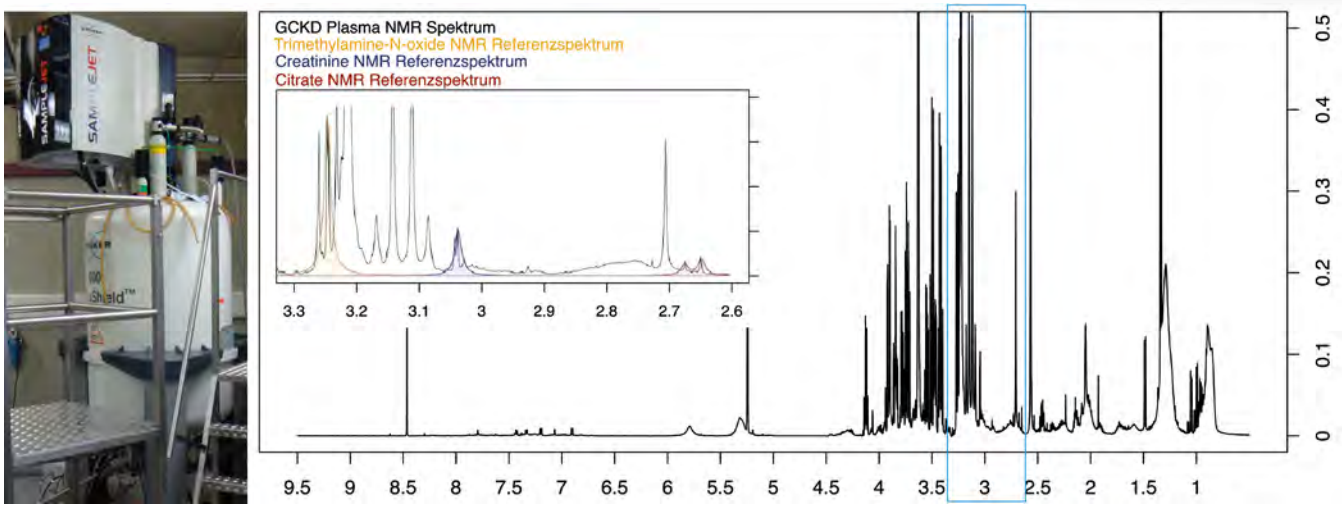


Abbildung 2: Das in diesem Projekt zum Einsatz kommende NMR Spektrometer (links) liefert hochaufgelöste NMR Spektren der zu untersuchenden GCKD Plasmaproben (rechts). Signale im Spektrum reflektieren bestimmte Metabolite, die aufgrund ihrer Signalposition auf der x-Achse identifiziert werden können. Die Signalhöhe auf der y-Achse reflektiert die Konzentration eines Metaboliten in der untersuchten Plasmaprobe. Die Metabolite, z. B. Kreatinin, können durch Vergleiche mit Referenzspektren identifiziert werden (s. Vergrößerung des hellblau markierten Spektrumausschnittes).
(Quelle: Helena U. Zacharias und Wolfram Gronwald).

Der e:Med Juniorverbund „CKDNapp", koordiniert von Helena Zacharias, wird die „Chronic Kidney Disease Nephrologist's App" (Abk. CKDNapp) für praktizierende Nephrologen entwickeln, um sie bei dem komplexen Prozess der Datenintegration und somit der personalisierten Behandlung von Patienten mit chronischer Niereninsuffizienz zu unterstützen. Bei der Untersuchung eines Patienten erhebt der behandelnde Arzt umfangreiche Patientendaten, die er in die CKDNapp einspeisen kann (Abbildung 1). Die App selbst wird auf zwei komplementären Säulen aufgebaut: (1) umfangreiche mathematische Diagnostik- und Prädiktionsmodelle, beispielsweise zur personalisierten Vorhersage von kardiovaskulären Ereignissen, einem terminalen Nierenversagen oder dem Tod des Patienten, und (2) eine umfassende Sammlung aus bereits in der Literatur etablierten Risikoprädiktoren. Im Besonderen wird die CKDNapp alle relevanten Prädiktoren dieser Patientenereignisse identifizieren und dem Nephrologen berichten, der somit stets die volle Kontrolle über alle Entscheidungsfindungsprozesse behält. Außerdem wird der behandelnde Arzt die Möglichkeit haben, innerhalb der App virtuell Patientenparameter zu adjustieren, Diagnosen zu verändern, oder virtuell die Krankheitsprogression zu sondieren. Solch virtuelle Veränderungen der Lebensweise eines Patienten, die laut der CKDNapp zu einer Verringerung des Patientenrisikos führen, könnten beispielsweise die Motivation des Betroffenen stärken, diese Veränderungen auch tatsächlich umzusetzen.

Aus kleinen Molekülen werden große Biomarker

Das Herzstück der CKDNapp bilden unterschiedliche mathematische Modelle zur personalisierten Diagnoseverfeinerung sowie zur Prädiktion von unerwünschten medizinischen Patientenereignissen und Krankheitsverläufen. Der e:Med Juniorverbund „CKDNapp" nutzt zur Erstellung dieser Modelle umfassende demographische, phänotypische, klinische Endpunkt- und Metabolomikdaten der *„German Chronic Kidney Disease"* (Abk. GCKD) Studie (https://www.gckd.de/), eine der weltweit größten CKD Beobachtungsstudien mit über 5.000 Patienten, die über 10 Jahre lang beobachtet werden. Vor allem die Metabolomikdaten, gewonnen aus Plasmaproben der GCKD Studie, haben es den Wissenschaftlern angetan.

Doch warum konzentriert man sich überhaupt auf den Metabolismus von Patienten mit chronischer Niereninsuffizienz? Metabolite, kleine organische Moleküle wie zum Beispiel Aminosäuren oder Zucker, sind Zwischen- und/oder Endprodukte des körpereigenen Stoffwechsels und in sämtlichen Körperflüssigkeiten und -geweben vorzufinden. Die Niere ist einer der Hauptregulatoren des Metabolismus, da sie Stoffwechselendprodukte aus dem Blut filtriert und über den Urin ausscheidet. Ist dieser wichtige Mechanismus aufgrund einer eingeschränkten Nierenfunktion gestört, gerät der Stoffwechsel in Unordnung. Letzteres spiegelt sich dann in einer veränderten Zusammensetzung von zahlreichen Metaboliten im Blut und/oder Urin wider. Metabolite können so als Biomarker für den derzeitigen physiologischen Patientenzustand, aber auch zur Vorhersage eines zukünftigen Patientenereignisses dienen.

Um ein aussagekräftiges Bild des menschlichen Metabolismus zu erhalten, müssen möglichst viele Metabolite gemessen wer-

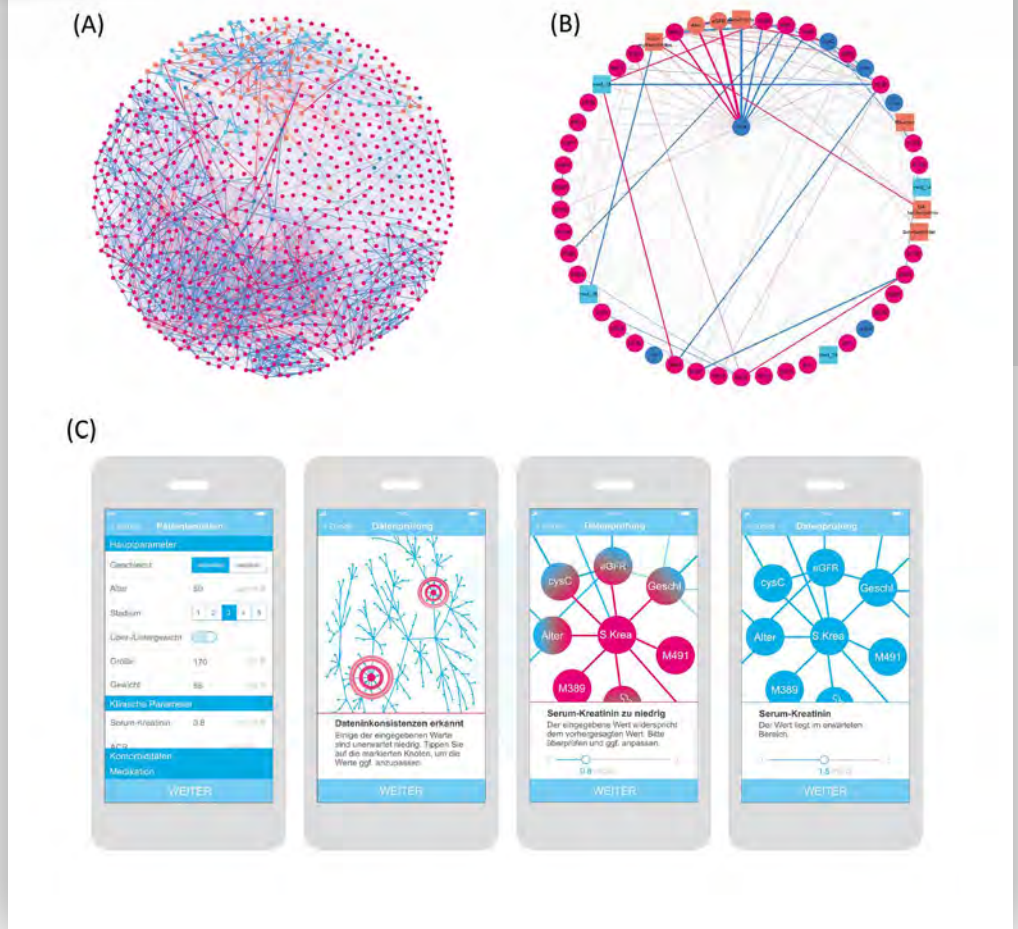


Abbildung 3: (A) Der Juniorverbund „CKDNapp“ berechnet mittels neu entwickelter Algorithmen ein mathematisches CKD-Modell aus den GCKD-Daten, das alle direkten Abhängigkeiten zwischen den einzelnen Patientenparametern aufdeckt. (B) Dieses Modell ermöglicht es z.B., alle Patientenparameter aufzuzeigen, die direkt mit dem Serum-Kreatinin-Wert eines Patienten zusammenhängen. (C) Die von dem Nephrologen in die CKDNapp hochgeladenen Patientendaten werden auf Dateninkonsistenzen überprüft. Hierbei vergleicht die App die eingelesenen Daten mit dem mathematischen Modell und weist den Benutzer darauf hin, dass der übergebene Serum-Kreatinin-Wert möglicherweise zu niedrig ist. Der Arzt kann nun „virtuell“ diesen Parameter verändern, um die Dateninkonsistenz aufzulösen (Quelle: (A) adaptiert von (Altenbuchinger *et al.*, 2019, CC BY 4.0) (B) Helena U. Zacharias (C) Johannes Raffler, Michael Altenbuchinger und Helena U. Zacharias).

den. Der Juniorverbund **„CKDNapp“** verfolgt deshalb einen hypothesenfreien Messansatz mit Hilfe dessen er alle detektierbaren Metabolite in einer Probe ohne vorherige Selektion bestimmt (engl. *„non-targeted metabolomics“*). Die in diesem Projekt zum Einsatz kommende Kernspinresonanzspektroskopie (engl. *„nuclear magnetic resonance spectroscopy“*, Abk. NMR Spektroskopie) ermöglicht solch einen *„non-targeted“* Metabolomik-Messansatz (Abbildung 2). Im Laufe des Projektes werden diese Metabolomikmessungen in allen Plasmaproben einer Folgevisite zwei Jahre nach der Basisuntersuchung wiederholt, um die CKDNapp-Modelle mit Zeitverlaufsdaten anzureichern. Kontinuierlich werden unterschiedlichste Daten über Patientenereignisse wie Herzinfarkt, Nierenversagen und Schlaganfall gesammelt und für die weitere Datenauswertung aufbereitet.

Diese umfangreichen Patienteninformationen der GCKD Studie mit aktuell über 15.100 unterschiedlichen Datenpunkten stellen einen wahren Schatz dar, der jetzt nur noch gehoben werden muss.

Computergestützte Datenintegration entwirrt hochkomplexe systembiologische Zusammenhänge im menschlichen Körper

Die Integration all dieser Patientenparameter wird mit Hilfe aktuellster Methoden des maschinellen Lernens durchgeführt. So erarbeitet der Juniorverbund „CKDNapp“ neue mathematische Modelle, die beispielsweise die präzise Vorhersage eines terminalen Nierenversagens bei chronischer Niereninsuffizienz ermöglichen (Zacharias *et al.*, 2019).

Wichtige Patienteninformationen wie medizinische Komplikationen, Begleiterkrankungen, demographische Faktoren, Medikamentengabe, Laborparameter und Metabolomikdaten sind sehr stark miteinander verwoben. Dies führt zu komplexen Abhängigkeiten zwischen den Variablen, die die Interpretation dieser mathematischen Modelle erschweren können. CKDNapp wird jedoch diese komplexen Assoziationsmuster ausnutzen und sie in einem konsistenten mathematischen Modell abbilden, um

Abbildung 4: Organisation des Juniorverbundes „CKDNapp“ (Fotos: Ulla T. Schultheiß, Helena U. Zacharias, Sahar Ghasemi, Michael Altenbuchinger, Robin Kosch, Johannes Raffler).

so ein detailliertes Bild der chronischen Nierenerkrankung zu erstellen (Abbildung 3A). Derartige, aus den GCKD-Daten mittels neu entwickelter Algorithmen „gelernte“ mathematische Modelle ermöglichen es den Wissenschaftlern herauszufinden, welche Patientenparameter direkt voneinander abhängen (Altenbuchinger *et al.*, 2019, 2020). Zum Beispiel verändert sich der Serum-Kreatinin-Wert mit dem Patientenalter und hängt zudem von unterschiedlichen Metaboliten, Medikamenten und Nierenerkrankungen ab (Abbildung 3B). Durch solch eine umfassende Modellierung komplexer Parameterzusammenhänge können neue Erkenntnisse über wichtige Pathomechanismen, die chronischer Niereninsuffizienz zugrunde liegen, gewonnen werden.

Insbesondere wird die CKDNapp mit Hilfe dieses Modells in der Lage sein, die Übereinstimmung zwischen neuen Patientendaten und Modell abzuschätzen und dadurch mögliche Abweichungen aufzudecken. Als Beispiel stelle man sich vor, dass der Serum-Kreatinin-Wert eines Patienten falsch gemessen wurde. Der Nephrologe lädt zunächst alle Patientendaten in die CKDNapp und lässt von der App eine automatische Datenüberprüfung durchführen (Abbildung 3C). CKDNapp erkennt eine Inkonsistenz zwischen den hochgeladenen Patientendaten und dem mathematischen CKD-Modell und weist den Nephrologen darauf hin, dass anscheinend der eingegebene Serum-Kreatinin-Wert zu niedrig ist. Der Arzt hat nun die Möglichkeit, diesen Wert „virtuell“ zu verändern, bis die Dateninkonsistenz aufgelöst ist. Somit weist die CKDNapp auf mögliche Dateninkonsistenzen hin, verringert potenzielle Vorhersagefehler und verbessert Therapieempfehlungen. Außerdem kann dem Patienten mit Hilfe dieses CKDNapp Features beispielsweise aufgezeigt werden, wie sich die Anpassung seines Lebensstils positiv auf seine Prognose auswirken würde. Da das basierend auf den Daten der GCKD Studie gelernte mathematische Modell alle Abhängigkeiten zwischen den einzelnen Patientenparametern gleichzeitig „gelernt“ hat, wird die CKDNapp in der Lage sein, derartige Dateninkonsistenzen aufzuzeigen, ohne von einer spezifischen Hypothese geleitet zu werden.

Die CKDNapp bringt metabolische Biomarkersignaturen in die klinische Praxis

Die CKDNapp wird als nutzerfreundliche Software mit intuitiven Benutzeroberflächen kostenlos zur Verfügung gestellt werden. Bereits während des Designprozesses wird der Juniorverbund „CKDNapp“ sowohl die Vorhersagekraft der neu entwickelten mathematischen Modelle, als auch die Anwendbarkeit der App selbst in einem kontrollierten klinischen Umfeld testen.

Aber können die metabolischen CKDNapp-Modelle, die basierend auf hochspezialisierten NMR-Spektroskopie-Messungen entwickelt wurden, überhaupt in der klinischen Praxis angewendet werden? Dieser Frage wird der „CKDNapp“ Juniorverbund vor der finalen Realisierung der App auf den Grund gehen. NMR-Spektroskopie wird im Augenblick kaum in der Routinelabordiagnostik eingesetzt, wohingegen die Massenspektrometrie hierfür durchaus verbreitet ist. Der Juniorverbund wird zunächst testen, ob die NMR-basierten Metabolit-CKDNapp-Modelle auf

Massenspektrometriedaten angewendet werden können, ohne erheblich an Vorhersagekraft zu verlieren. Er setzt dabei auf einen neu entwickelten Algorithmus des maschinellen Lernens namens „Zero-sum“ Regression (Altenbuchinger *et al.,* 2017a). Mit Hilfe dieses Algorithmus konnten bereits Proteomics-Biomarkersignaturen erfolgreich von einer analytischen Messplattform auf eine andere transferiert werden (Altenbuchinger *et al.,* 2017b). Im Rahmen des „CKDNapp“-Projektes wird ausführlich untersucht werden, ob ein derartiger Plattformtransfer auch für metabolische Biomarkersignaturen möglich ist. Letztlich sollen die neuen metabolischen CKDNapp-Biomarker auch mit standardisierten analytischen Methoden gemessen werden können, um eine breite und kostenoptimierte Anwendung der App in der klinischen Routinediagnostik zu ermöglichen. Auch dies wird der Juniorverbund „CKDNapp“ erarbeiten und bringt so die Systemmedizin in die klinische Anwendung.

Steckbrief Forschungsprojekt:

Der e:Med Juniorverbund „CKDNapp: Eine Toolbox zur Beobachtung und maßgeschneiderten Therapie von Patienten mit chronischem Nierenleiden – ein personalisierter, systemmedizinischer Ansatz“ wird vom Bundesministerium für Bildung und Forschung seit 2019 für fünf Jahre gefördert. Das interdisziplinäre Team umfasst die Arbeitsgruppen von vier Juniorgruppenleiterinnen und -leitern, die sich folgenden Teilprojekten widmen (Abbildung 4):

- **TP 1:** Konsolidierung von Datensätzen, Generierung von klinischen Input-Variablen, medizinische Interpretation und Validierung, Ulla T. Schultheiß (Universitätsklinikum Freiburg)
- **TP 2:** Berechnung der mathematischen Modelle für CKDNapp, Helena U. Zacharias (Universitätsmedizin Greifswald)
- **TP 3:** Algorithmische Grundlagen der CKDNapp Modelle, Michael Altenbuchinger (Universität Hohenheim)
- **TP 4:** Entwicklung der CKDNapp App und Webserver, Johannes Raffler (Helmholtz Zentrum München)

Weitere Informationen: https://www.ckdn.app

Referenzen:

Altenbuchinger, M., Rehberg, T., Zacharias, H.U., Stämmler, F., Dettmer, K., Weber, D., Hiergeist, A., Gessner, A., Holler, E., Oefner, P.J., *et al.* (2017a). Reference point insensitive molecular data analysis. Bioinformatics.

Altenbuchinger, M., Schwarzfischer, P., Rehberg, T., Reinders, J., Kohler, C.W., Gronwald, W., Richter, J., Szczepanowski, M., Masqué-Soler, N., Klapper, W., Oefner, P.J., Spang, R. (2017b). Molecular signatures that can be transferred across different omics platforms. Bioinformatics.

Altenbuchinger, M., Zacharias, H.U., Solbrig, S., Schäfer, A., Büyüközkan, M., Schultheiß, U.T., Kotsis, F., Köttgen, A., Spang, R., Oefner, P.J., *et al.* (2019). A multi-source data integration approach reveals novel associations between metabolites and renal outcomes in the German Chronic Kidney Disease study. Sci. Rep.

Altenbuchinger, M., Weihs, A., Quackenbush, J., Grabe, H.J., and Zacharias, H.U. (2020). Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools. Biochim. Biophys. Acta - Gene Regul. Mech.

Zacharias, H.U., Altenbuchinger, M., Schultheiss, U.T., Samol, C., Kotsis, F., Poguntke, I., Sekula, P., Krumsiek, J., Köttgen, A., Spang, R., *et al.* (2019). A Novel Metabolic Signature to Predict the Requirement of Dialysis or Renal Transplantation in Patients with Chronic Kidney Disease. J. Proteome Res.

Kontakt:

Dr. Helena U. Zacharias
Wissenschaftliche Koordinatorin des e:Med Juniorverbundes „CKDNapp“
Nachwuchsgruppenleiterin „Computational Biomarker Discovery“
Universitätsmedizin Greifswald
Klinik und Poliklinik für Psychiatrie und Psychotherapie
Greifswald
helena.zacharias@med.uni-greifswald.de

Fotos: privat / Zacharias

alternatives spleißen aus sicht der systemmedizin

Welchen Einfluss hat alternatives Spleißen auf die Entstehung von Krankheiten?

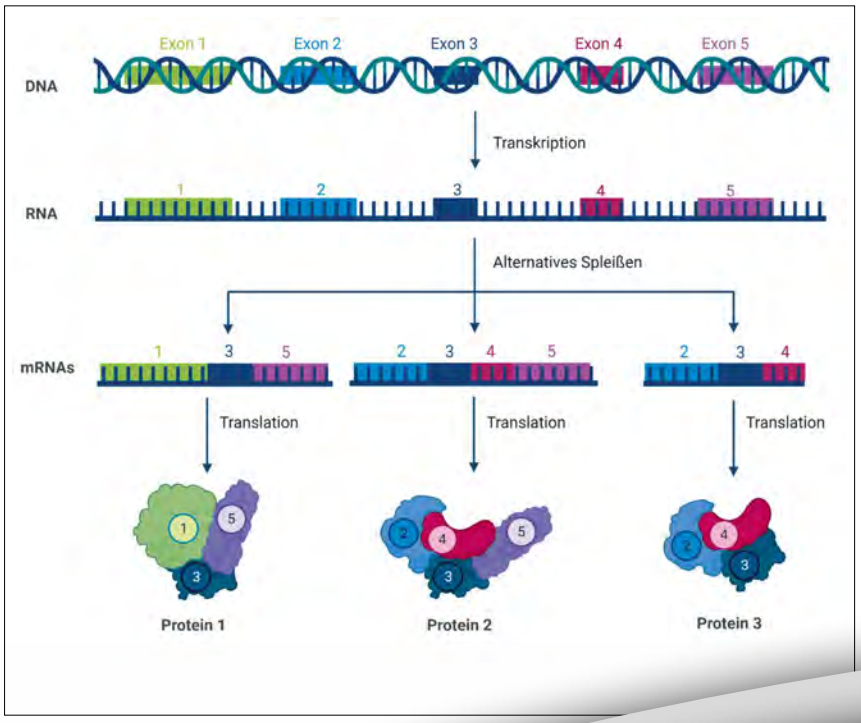
von Tim Kacprowski, Nina Kerstin Wenke, Sabine Ameling, Kristin Wenzel, Olga V. Kalinina und Markus List im Namen des gesamten Sys_CARE Konsortiums

Durch alternatives Spleißen (AS) entstehen aus dem selben Gen verschiedene Transkripte und somit verschiedene Proteine (Isoformen). Bisher wird in gängigen Studien, die auf Molekulardaten (sog. Omics) basieren, kaum zwischen verschiedenen Isoformen unterschieden. Dadurch ist die Bedeutung von AS für die Entstehung von Krankheiten noch unzureichend erforscht. Zum Verständnis der weitreichenden Auswirkungen von AS ist ein systemmedizinisches, integratives Vorgehen unerlässlich. **Sys_CARE** verfolgt das Ziel, die Bedeutung von AS für die Entstehung und den Verlauf von Krankheiten systemmedizinisch zu untersuchen und hierfür nötige bioinformatische Werkzeuge zu entwickeln.

Alternatives Spleißen bei Dilatativer Kardiomyopathie (DCM) und Hypertensiver Nephrosklerose (HN)

AS hat weitreichende Folgen, da verschiedene Proteinisoformen unterschiedliche Strukturen und Funktionen ausbilden können (Abbildung 1) und somit auch in Protein-Protein Interaktionen und komplexen Signalwegen unterschiedlich agieren. Einzelne Isoformen werden bisher selbst in molekularbasierten (Omics-) Datensätzen nur selten unterschieden. Welche Auswirkungen AS für die Entstehung und den Verlauf von Krankheiten hat, konnte daher noch nicht im Detail erforscht werden. *„Bei der Integration von Multi-Omics-Daten muss neben der Genexpression auch die Proteinstruktur und deren mögliche Auswirkungen auf die Funktion und Protein-Protein Interaktionen betrachtet werden, um die molekularen Mechanismen der Krankheitsentstehung zu verstehen“,*

Abbildung 1: Alternatives Spleißen



Aus einem Gen können durch alternatives Spleißen unterschiedliche Transkripte entstehen. Durch Verwendung unterschiedlicher Exons der RNA entstehen verschiedene mRNAs, die wiederum während der Translation zu strukturell und funktional verschiedenen Proteinen (Isoformen) führen. (Erstellt mit BioRender.com.)

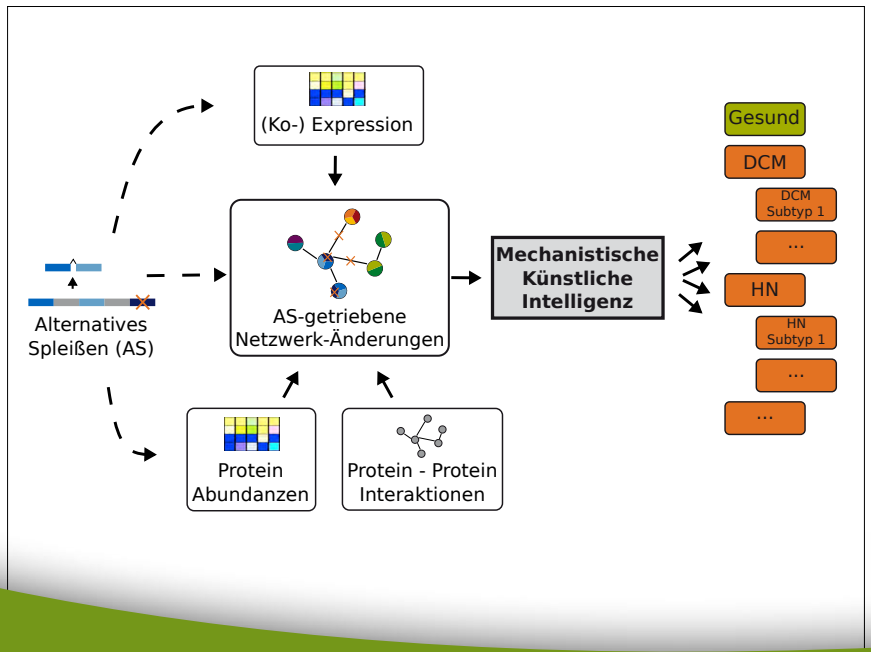
Quelle: Sys_CARE Konsortium

betont Professorin Olga V. Kalinina (Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) / Helmholtz-Zentrum für Infektionsforschung (HZI)).

Wir konzentrieren uns auf zwei Krankheitsbilder, bei denen AS mutmaßlich eine Rolle spielt. Die dilatative Kardiomyopathie (DCM) ist eine schwerwiegende Herzerkrankung, die durch das Vorhandensein von links- (LV) oder biventrikulärer Dilatation und systolischer Dysfunktion gekennzeichnet ist. DCM ist die häufigste Form der Kardiomyopathie und die häufigste Ursache für nicht-ischämische Herzinsuffizienz und Herztransplantation im Endstadium. DCM hat sowohl genetische als auch nicht-genetische Ursachen. Es ist bekannt, dass eine zunehmende Anzahl von Mutationen in bestimmten Genen wie z. B. Titin, Laminin oder Myosin, mit der Entwicklung von DCM in Zusammenhang stehen. *„Neben genetischen Ursachen spielen auch Entzündungsvorgänge bei der Entstehung der DCM eine Rolle. Die molekularen Mechanismen der Pathogenese dieser häufigen Myokarderkrankung sind nicht vollständig verstanden und müssen genauer untersucht werden“*, erklärt Professor Stephan B. Felix (Universitätsmedizin Greifswald). Bereits bekannt ist, dass AS für die Entwicklung von Erkrankungen des Herzens eine Rolle spielt (Beqqali 2018). Jedoch konnte bisher kein kausaler Zusammenhang zwischen AS und der Entwicklung einer DCM im Speziellen nachgewiesen werden. Auch im zweiten untersuchten Krankheitsbild, der hypertensiven Nephrosklerose (HN), sind die Krankheitsmechanismen nicht vollständig verstanden. Die Rolle von AS in HN wurde bisher nicht untersucht. Es wurden jedoch in nierenspezifischen Zellen, den Podozyten, AS für zelltypische Gene wie z. B. WT-1 (Wilm's Tumorprotein 1) beobachtet, die zu verschiedenen Isoformen führen (Barbaux *et al.*, 1997).

„Diese Beispiele zeigen die Notwendigkeit einer systematischen Analyse der Beteiligung von AS bei Herz- und Nierenerkrankungen“, bestätigt Professor Uwe Völker (Universitätsmedizin Greifswald). Derzeit basieren vorhandene Daten überwiegend auf Genexpressions-Array-Technologien, die die Identifizierung neuer Genprodukte nicht zulassen. Zudem wurden die meisten transkriptomweiten AS-Studien zur Untersuchung von Kardiomyopathien an Mausmodellen durchgeführt. Obwohl AS in Tiermodellen für Nierenerkrankungen nachgewiesen wurde, beschränkten sich transkriptomweite AS-Analysen auf Patienten mit genetisch bedingten Nierenerkrankungen und wurden bisher nicht auf die häufigeren, nicht-genetischen Nierenerkrankungen wie HN angewendet. Daher ist eine systemmedizinische Forschung unter Verwendung größerer Kohorten und hochsensibler Sequenzierungsmethoden dringend erforderlich.

Abbildung 2: Systemmedizinische Interpretation von Alternativem Spleißen



Quelle: Sys_CARE Konsortium

Sys_CARE vereint Daten aus verschiedenen Omics Ebenen, die komplementäre Beobachtungen von alternativem Spleißen darstellen. Expressionsdaten, Proteinabundanzen und Protein-Protein-Interaktionsdaten werden zu heterogenen Netzwerken integriert, um die Effekte von AS auf Netzwerkebene abzubilden. Diese dienen dann als Prädiktoren einer künstlichen Intelligenz (KI) die mechanistische Muster erkennt, d. h. die nicht nur die Klassifizierung vornimmt, sondern diese auch durch molekulare Mechanismen erklären kann, um die Krankheiten diagnostisch zu unterscheiden und neue Krankheitssubtypen zu identifizieren.

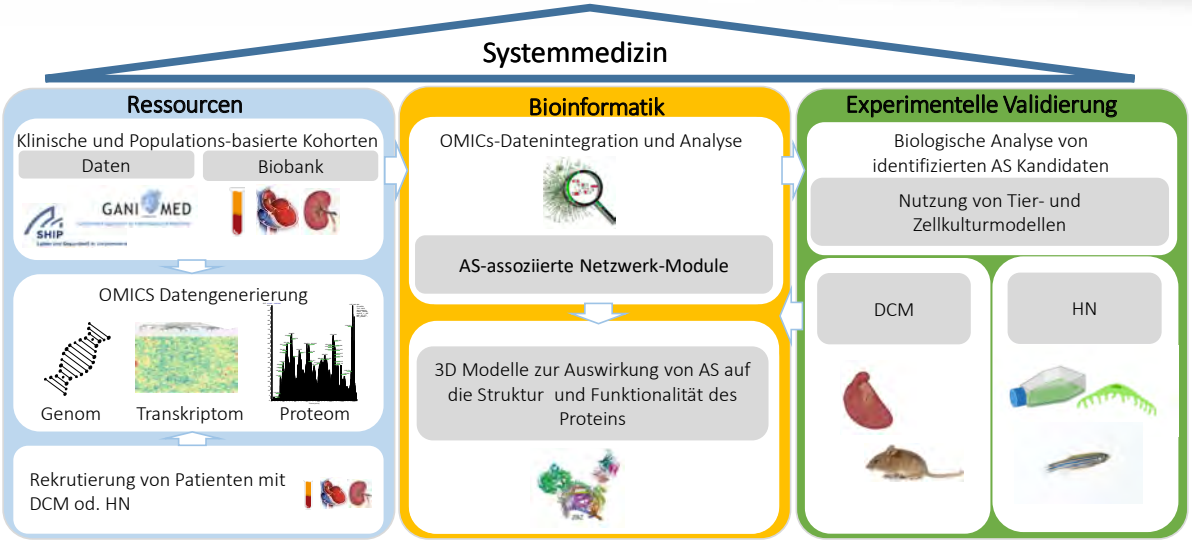


Abbildung 3: Systemmedizin für die Detektion und funktionelle Charakterisierung von alternativen Spleißvarianten. Im Sys_CARE Konsortium werden durch die enge interdisziplinäre Zusammenarbeit zwischen Klinikern, Naturwissenschaftlern und Bioinformatikern die Datenintegration, Analysen der AS-Varianten sowie Netzwerkanalysen und 3D-Modellierung möglich. Ziel ist es, mit humanen Omics-Daten ein besseres Verständnis von AS-Varianten und deren funktionelle Relevanz in Erkrankungen wie DCM und HN zu erhalten und die gewonnenen Erkenntnisse für die klinische Anwendung nutzbar zu machen. (Quelle: Sys_CARE Konsortium)

„Die dilatative Kardiomyopathie und die hypertensive Nephrosklerose sind zwei heterogene Erkrankungen, die derzeit phänotypisch, d. h. über ihre Symptome, definiert werden. Das Paradigma der Systemmedizin ist jedoch, dass komplexe Erkrankungen stattdessen mechanistisch, d. h. auch über die betroffenen molekularen Prozesse und Stoffwechselwege, charakterisiert werden sollten“ unterstreicht Professor Jan Baumbach (Technische Universität München (TUM)). Das mechanistische Verständnis der Krankheit und deren Entstehung ermöglicht zielgerichtete Therapien, die sich der Krankheitsursache annehmen, anstatt sich nur der Symptombekämpfung zu widmen. Deshalb nutzt Sys_CARE innovative Methoden zur Netzwerkanalyse, um nicht nur durch AS-betroffene Proteinisoformen, sondern auch deren Einfluss auf Protein-Protein Interaktionen und funktionelle Netzwerke mit heterogenen Multi-Omics-Daten zu identifizieren (Abbildung 2). Dies kann zur Identifizierung von Krankheitssubtypen führen, für die dann eine optimale Therapie im Sinne der personalisierten Medizin entwickelt werden kann.

Sys_CARE – innovative, interdisziplinäre und translationale Ansätze in der Systemmedizin

Die Basis der systemmedizinischen Forschung in Sys_CARE bilden klinische und populationsbasierte Kohorten. Mit Zugriff auf und Generierung von umfangreichen Multi-Omics Daten werden systemmedizinische Analysen ermöglicht. Im Mittelpunkt steht dabei der Einfluss von AS auf die Bildung von Proteinkomplexen sowie auf regulatorische Mechanismen. Dabei werden insbesondere Veränderungen in molekularen Interaktionsnetzwerken untersucht, die die Entstehung und den Verlauf der Krankheiten beeinflussen. Neu zu entwickelnde bioinformatische Methoden werden es dabei ermöglichen, Genexpressionsdaten in diese Netzwerke zu integrieren, um zu beleuchten, wie AS zelluläre Mechanismen und regulatorische Programme beeinflusst.

Parallel dazu werden AS-Ereignisse in Zeitverlaufsstudien mit Maus- und Zebrafischmodellen der DCM bzw. der HN *in vivo* und *in vitro* verfolgt. Dies ermöglicht ein Screening auf AS-Ereignisse während des Fortschreitens der Krankheit, was die Identifizierung von dynamischen AS-Mechanismen ermöglicht. Die Resultate wiederum dienen der iterativen Verbesserung der Netzwerkanalyse-Methoden (Abbildung 3).

Ein wichtiger Schritt ist auch die Prüfung von AS-Ereignissen durch den Nachweis von spezifischen Proteinisoformen oder Peptiden, die für die klinische Diagnostik geeignet sind. Ferner wird der Vergleich von Blut-basierten und Biopsie-basierten AS-Ereignissen aufzeigen, inwieweit im Blut gefundene krankheitsassoziierte AS-Ereignisse mechanistische Vorgänge im Herz- oder Nierengewebe widerspiegeln. Diese könnten Biomarker für die klinische Diagnostik aufzeigen, die auch ohne invasive Eingriffe zur Gewinnung von Gewebsbiopsien (Bioptaten) direkt im Blut messbar sind.

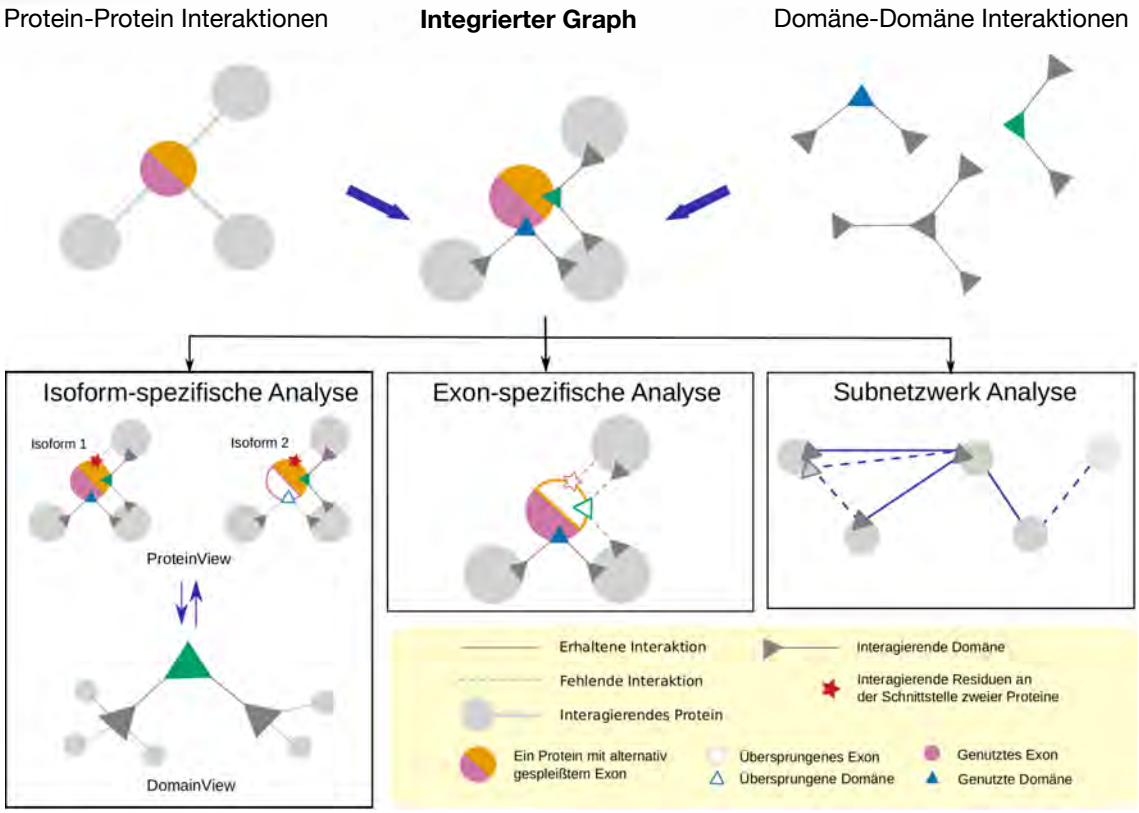


Abbildung 4: DIGGER Übersicht: Protein-Protein-Interaktionsdaten werden mit Domäne-Domäne-Interaktionsdaten angereichert, um strukturell annotierte Interaktionsdaten für jedes Gen zu erhalten. Die angereicherten Daten erlauben den Vergleich verschiedener Proteinisoformen (links), die Aufklärung funktioneller Effekte ausgespleißter Exons (Mitte) oder die Untersuchung von AS auf Netzwerkebene (rechts). (Quelle: https://www.exbio.wzw.tum.de/digger/)

Erste Ergebnisse von Sys_CARE

Um den Einfluss von AS auf Netzwerkebene untersuchen zu können, wurde im Rahmen von Sys_CARE die Webanwendung DIGGER (www.exbio.wzw.tum.de/digger , Louadi *et al.*, 2020) entwickelt. Bisherige Tools und Datenbanken beschränken sich meist auf Protein-Protein-Interaktionen und vernachlässigen dadurch den Einfluss von AS, was dazu führen kann, dass die für eine Interaktion benötigte Protein-Domäne nicht Teil einer Protein-Isoform wird. Um hier eine detaillierte Analyse zu erlauben, bildet DIGGER solche Interaktionen auf Ebene der einzelnen Protein-Domänen und der darunter liegenden Exon-Strukturen ab (Abb. 4). Zu diesem Zweck integriert DIGGER Informationen zu Protein-Protein-Interaktionen und Domäne-Domäne-Interaktionen in einem gemeinsamen Netzwerk. Darüber hinaus zeigt DIGGER interagierende Residuen, d. h. Aminosäuren, die sich an der Schnittstelle zweier interagierender Proteine finden. Diese wurden mittels bioinformatischer Methoden aus der experimentellen Röntgenkristallographie und Kernspinaufnahmen von interagierenden Proteinstrukturen ermittelt und liefern damit einen komplementären Beleg für eine mögliche Protein-Interaktion. DIGGER erlaubt es den Benutzern, nach einzelnen oder Gruppen von Exons oder Isoformen zu suchen, um ihre Interaktionen unter dem Einfluss von AS visuell zu erforschen.

Translationales Potenzial von Sys_CARE für die Präzisionsmedizin

Die in Sys_CARE entwickelten bioinformatischen Methoden werden Studien zu AS in Krankheiten entschieden voranbringen und das Feld der Systemmedizin über das laufende Projekt hinaus prägen. Wir erwarten, dass die durch Sys_CARE gewonnenen Einblicke in die Rolle von AS bei der DCM und der HN zur Entwicklung der Präzisionsmedizin beitragen. Das translationale Potenzial der hier geleisteten Grundlagenforschung zeigen aktuelle klinische Studien, die Antisense-Oligonukleotide (AONs) als therapeutisches Mittel zur effektiven Kontrolle des Spleißens einsetzen (Spitali and Aartsma-Rus 2012). Dies legt nahe, dass eine wirksame Behandlung von Erkrankungen wie der DCM und der HN schon bald nach einem mechanistischen Verständnis der Rolle von AS möglich sein wird.

Steckbrief Forschungsprojekt:

Sys_CARE wird im Rahmen von e:Med Systems Medicine durch das BMBF gefördert und vereint die Expertisen aus der medizinischen Grundlagenforschung (Universitätsmedizin Greifswald, Prof. Dr. Uwe Völker, Prof. Dr. Karlhans Endlich, Prof. Dr. Nicole Endlich), der funktionellen Proteinmodellierung (Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) / Helmholtz-Zentrum für Infektionsforschung, Prof. Dr. Olga V. Kalinina), der multi-Omics Netzwerkanalyse (Technische Universität München, Prof. Dr. Jan Baumbach, Dr. Markus List, Dr. Tim Kacprowski) und der klinischen Anwendung (Universitätsmedizin Greifswald, Prof. Dr. Stephan B. Felix, Prof. Dr. Sylvia Stracke). Das Gesamtziel von Sys_CARE ist es, die Bedeutung von AS für die Krankheitsentstehung und ihren Verlauf mit systemmedizinischen Ansätzen zu untersuchen und die hierfür nötigen bioinformatischen Methoden zu entwickeln.

Wichtige Teilprojekte sind:

TP 1: Klinische Omics Analyse (Prof. Dr. Uwe Völker)
TP 2: Modellierung von Proteinstruktur und -funktion (Prof. Dr. Olga V. Kalinina)
TP 3: Alternatives Spleißen zugehöriger Netzwerkmodule (Prof. Dr. Jan Baumbach)
TP 4: Experimentelle Studien zu krankheitsassoziierten alternativen Spleißereignissen (Prof. Dr. Stephan B. Felix)

Referenzen:

Barbaux, S., Niaudet, P., Gubler, M.C., Grünfeld, J.P., Jaubert, F., Kuttenn, F., Fékété C.N., *et al.* (1997). „Donor Splice-Site Mutations in WT1 Are Responsible for Frasier Syndrome." *Nature Genetics* 17 (4): 467–70.

Beqqali, A. (2018). „Alternative Splicing in Cardiomyopathy." *Biophysical Reviews* 10 (4): 1061–71.

Louadi, Z., Yuan, K., Gress, A., Tsoy, O., Kalinina, O.V., Baumbach, J., *et. al.* (2020). „DIGGER: exploring the functional role of alternative splicing in protein interactions." *Nucleic Acids Research.* 10.1093/nar/gkaa768.

Spitali, P., and Aartsma-Rus, A. (2012). „Splice Modulating Therapies for Human Disease." *Cell* 148 (6): 1085–88.

Kontakt:

Prof. Dr. Jan Baumbach
Lehrstuhl für Experimentelle Bioinformatik,
Technische Universität München (TUM), Freising
jan.baumbach@wzw.tum.de
www.exbio.de

Prof. Dr. Uwe Völker
Universitätsmedizin Greifswald
voelker@uni-greifswald.de
https://biologie.uni-greifswald.de/struktur/institute/interfakultaeres-institut-fuer-genetik-und-funktionelle-genomforschung/department-of-functional-genomics

Prof. Dr. Olga V. Kalinina
Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) / Helmholtz-Zentrum für Infektionsforschung (HZI),
Universität des Saarlandes, Saarbrücken
olga.kalinina@helmholtz-hips.de
https://www.helmholtz-hzi.de/en/research/research-topics/anti-infectives/drug-bioinformatics

Prof. Dr. Stephan B. Felix
Klinik und Poliklinik für Innere Medizin B,
Universitätsmedizin Greifswald
felix@uni-greifswald.de
http://www2.medizin.uni-greifswald.de/inn_b/

Prof. Dr. Karlhans Endlich
Institut für Anatomie und Zellbiologie,
Universitätsmedizin Greifswald
karlhans.endlich@uni-greifswald.de
http://www2.medizin.uni-greifswald.de/anatomie

bioinformatik und systemkardiologie

Ein Institutsporträt des Klaus Tschira Institute for Integrative Computational Cardiology

von Tobias Jakobi und Christoph Dieterich

Das Klaus Tschira Institut für Integrative Computergestützte Kardiologie ist in drei Themenfeldern aktiv. Erstens: RNA-Reifung und Verarbeitung. Insbesondere die Entwicklung und Physiologie des Herzens erfordern eine strenge Kontrolle der RNA Biologie. Unserem Labor ist es gelungen, zahlreiche Software-Lösungen zur Untersuchung der komplexen RNA-Welt zu veröffentlichen. Zweitens haben wir das Gebiet der Systemkardiologie für *in vitro* und *in vivo* Modelle der Herzinsuffizienz etabliert. Drittens wird durch das HiGHmed Konsortium im Rahmen der Medizininformatikinitiative eine Brücke in den Bereich der klinischen Datenwissenschaft eröffnet. An dieser Stelle sind insbesondere unsere Arbeiten mit künstlicher Intelligenz (KI) im Bereich unstrukturierter deutscher Texte aus dem kardiologischen Umfeld zu nennen.

Klaus Tschira Stiftung gemeinnützige GmbH

KTS

Das Klaus Tschira Institute für Computational Cardiology wurde im September 2015 mit Förderung der Klaus Tschira Stiftung gegründet und wird von Prof. Dr. Christoph Dieterich geleitet. Wir befassen uns in der Bioinformatik mit der Verarbeitung genetischer Informationen von der DNA zu Proteinen. Dies wurde häufig als geradliniger Weg angesehen, auf dem die RNA nur ein Zwischenprodukt darstellt. Dieses Bild wird der Rolle der RNA allerdings nicht gerecht; vielmehr ist sie ein interaktiver und dynamischer Informationsträger, der eine Vielzahl von Funktionen erfüllt. Stabilität und Translationseffizienz der RNA werden sowohl von ihrer Sekundärstruktur als auch von Interaktionen mit RNA-Bindeproteinen und nichtkodierenden RNAs wie beispielweise microRNA oder lncRNA (lange, nichtkodierende RNA) gesteuert. Co- und posttranskriptionelle Prozesse, wie RNA Modifikationen, können RNA Moleküle zudem auf Basenpaarebene verändern und so auch noch nach der Transkription Einfluss auf die finale Proteinsequenz nehmen. Mit der wiederentdeckten Klasse der zirkulären RNAs (circRNAs) hat zudem eine weitere, noch weitgehend unerforschte Gruppe von RNA-Molekülen Aufnahme in den Kreis der nichtkodierenden RNAs gefunden. Das Zusammenspiel all dieser Teile in einem großen Interaktionsnetzwerk wird heute unter dem Begriff der post-transkriptionalen Genregulation zusammengefasst und steuert zahlreiche Abläufe in unseren Zellen. Klassischerweise stehen spezifische Fragestellungen oder Beobachtungen aus der RNA-Biomedizin am Anfang unserer Arbeit.

Eine mögliche Fragestellung wäre beispielsweise:

„Herzmuskelzellen wachsen sowohl durch Fitnesstraining als auch durch krankhafte Einflüsse, beispielsweise Bluthochdruck. Warum aber unterscheiden sich die Langzeiteffekte auf molekularer und medizinischer Ebene deutlich?"

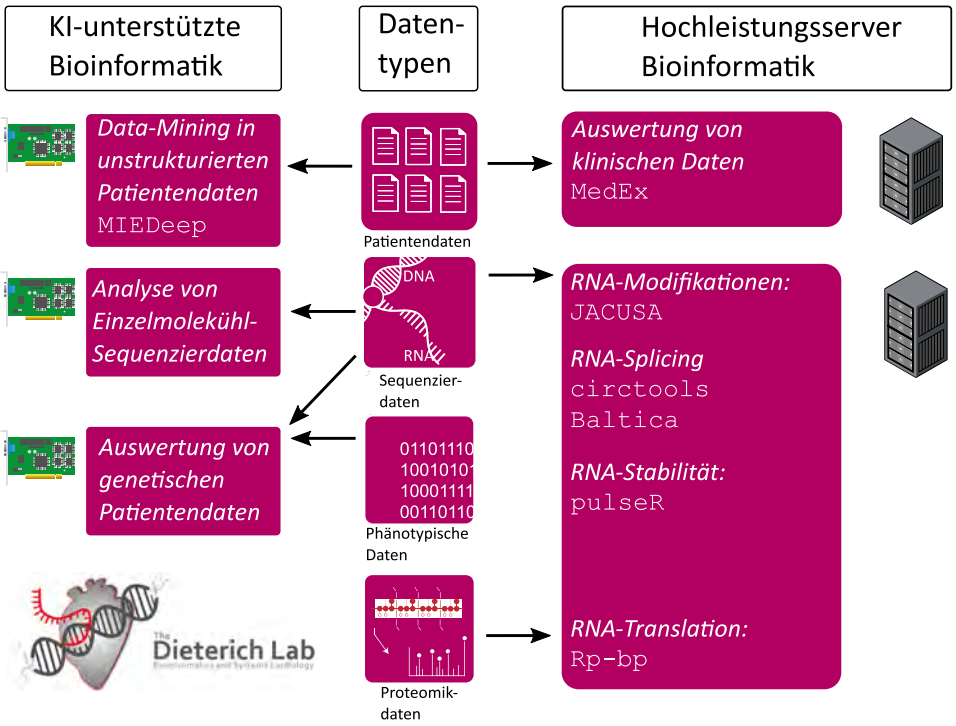


Abbildung 1: Softwareübersicht. Neue Methoden, die auf künstlicher Intelligenz basieren halten Einzug in verschiedene Felder der Bioinformatik und medizinischen Informatik (links), auch wenn derzeit noch klassische Bioinformatiktools einen Großteil der Workflows dominieren (rechts) (Quelle: Tobias Jakobi).

Quelloffene Softwarewerkzeuge für die wissenschaftliche Gemeinschaft

In der Regel erstellen wir mit unseren experimentellen Partnern gemeinsam Hypothesen, die wir dann sowohl durch etablierte bioinformatische und statistische Methoden als auch durch selbstentwickelte Software und Verfahren überprüfen. Neuentwickelte Softwarewerkzeuge werden quelloffen für die wissenschaftliche Gemeinschaft bereitgestellt und stetig weiterentwickelt.

Die Arbeitsgruppe hat so beispielsweise eine Software entwickelt, die in der Lage ist, modifizierte RNA Basenpaare aus Sequenzierungsdaten zu erkennen (Piechotta *et. al,* 2017). Weitere spezialisierte Softwarelösung für RNA Spleißen sind Baltica und insbesondere für zirkuläre RNAs circtools (Jakobi *et al.,* 2019). Die Software wurde implementiert, um den gesamten Workflow von Qualitätsanalyse der Rohdaten, über Detektion und Rekonstruktion von zirkulären RNAs, bis hin zum Design molekulargenetischer Primersequenzen für Validierungsexperimente abzudecken.

Für viele regulatorische Funktionen ist die Stabilität von RNA ein kritischer Faktor. In vielen Fällen wird die Verfügbarkeit der RNA Blaupause rasch durch Zerfallsprozesse oder Neusynthese kontextabhängig reguliert. Mit PulseR und weitergehenden theoretischen Arbeiten wurde in der Arbeitsgruppe ein Werkzeug für die Analyse der RNA-Stoffwechselkinetik aus RNA-Sequenzierungsdaten entwickelt (Uvarovskii *et al.,* 2019).

In anderen Fällen ist es jedoch wichtig, welche RNAs tatsächlich in Proteine translatiert werden und wie sich die Translation der Proteine im Vergleich zur Transkription der RNA verhält. Die Erstellung von Ribosomenprofilen mittels Hochdurchsatz-Sequenzierung (Ribo-seq) ist eine vielversprechende neue Technik zur Charakterisierung der Ribosomenverteilung auf RNA mit Basenpaar-Auflösung. Das Ribosom ist für die Übersetzung der mRNA in Proteine verantwortlich, so dass Informationen über ihre Belegung eine detaillierte Ansicht der Ribosomendichte und -position bieten, die u.a. zur Entdeckung neuer translatierter offener Leseraster (ORFs) verwendet werden könnte. Ein Bayes'scher Ansatz zur Vorhersage von ORFs aus Ribosomenprofilen wurde in der Software Rp-Bp implementiert (Malone *et al.,* 2017).

Abbildung 2: Gruppenbild. Von links nach rechts: Maja Bencun, Aljoscha Kindermann, Thiago Britto Borges, Isabel Naarmann-de Vries, Etienne Boileau, Tami Liebfried, Jessica Eschenbach, Christoph Dieterich, Magdalena Smieszek, Phillip Richter-Pechanski, Qi Wang und Tobias Jakobi (Foto: Tobias Jakobi)

Systemkardiologie benötigt adäquate und spezialisierte Hardware

Die quantitative Systemkardiologie zeichnet sich durch immense Datenmengen aus, die auf gewöhnlichen Arbeitsplatzrechnern nicht mehr handhabbar sind. Die Arbeitsgruppe unterhält zu diesem Zweck ein eigenes Netzwerk von Hochleistungsrechnern, die in der Lage sind, auch umfangreiche experimentelle Datensätze in kurzer Zeit zu analysieren. Der Rechnercluster besteht derzeit aus 26 dedizierten Rechenknoten mit einem Arbeitsspeicher von bis zu einem Terabyte, der zum Beispiel für die Genomassemblierung oder die parallele Analyse großer OMICS Datensätze benötigt wird.

Darüber hinaus wurde der Rechnerverbund mit einem dedizierten Server ausgerüstet, der NVIDIA GPUs (*Graphics Processing Units*) beherbergt. Diese Spezialhardware stammt von 3D-Grafikkarten für Computerspiele ab, welche in den letzten Jahren immer leistungsfähiger wurden und durch ihre hochparallele Architektur prädestiniert sind, Aufgaben des maschinellen Lernens und der künstlichen Intelligenz zu verarbeiten. Das Spezialsystem wird für eine Vielzahl von Aufgaben eingesetzt, die von der Extraktion der Sequenz von Basenpaaren aus Sequenzierungsrohdaten, über Text Mining in medizinischen Dokumenten bis hin zur Analyse von Patientengenomen aus molekulargenetischen Daten reichen.

Softwarelösungen für strukturierte und nichtstrukturierte Patientendaten

In der klinischen Praxis fallen routinemäßig große Mengen an Daten aus den verschiedensten Bereichen an. Unsere Software, der Medical Data Explorer (MedEx) (Kindermann *et al.,* 2019), ist eine intuitive, webbasierte Lösung, mit Möglichkeiten zum einfachen Datenimport. Wir verbinden eine moderne dynamische Webschnittstelle mit einer In-Memory-Datenbanklösung für eine nahezu Echtzeit-Reaktionsfähigkeit. MedEx bietet verschiedene Visualisierungsoptionen, um einen einfachen Überblick über die geladenen Daten zu erhalten, um Hypothesen zu generieren und elementare Analysen durchzuführen.

In der Medizin werden viele behandlungsrelevante Informationen nach wie vor in Form von unstrukturierten Texten in deutscher Sprache erfasst. Ein typisches Beispiel ist der Arztbrief, der als Transferdokument für die Kommunikation zwischen Ärzten gedacht ist. Unser Projekt MIEdeep (*Medical Information*

Extraction using deep learning) möchte diese Datenquelle für die Gewinnung von Informationen nutzbar machen. Hierfür kommen innovative Ansätze aus den Bereichen der tiefen neuronalen Netze (*Deep Learning*) und der maschinellen Sprachverarbeitung (NLP) zum Einsatz. Wir verknüpfen hierbei Ansätze maschinellen Lernens für die Datenaufbereitung, Erstellung von Trainingsdaten und Informationsextraktion mit einer modernen grafischen Benutzeroberfläche, die für den Einsatz in einem klinischen Umfeld geeignet ist.

Referenzen:

Piechotta, M., Wyler, E., Ohler, U., Landthaler, M., Dieterich, C. (2017). JACUSA: site-specific identification of RNA editing events from replicate sequencing data. *BMC Bioinformatics*, 18, 7, doi:10.1186/s12859-016-1432-8.

Jakobi, T., Uvarovskii, A., Dieterich, C. (2019). circtools – a one-stop software solution for circular RNA research. *Bioinformatics*, 35, 2326–2328, doi:10.1093/bioinformatics/bty948.

Uvarovskii, A., Vries, I.S.N., Dieterich, C. (2019). On the optimal design of metabolic RNA labeling experiments. *PLOS Computational Biology*, 15, e1007252, doi:10.1371/journal.pcbi.1007252.

Malone, B., Atanassov, I., Aeschimann, F., Li, X., Großhans, H., Dieterich, C. (2017). Bayesian prediction of RNA translation from ribosome profiling. *Nucleic Acids Res*, 45, 2960–2972, doi:10.1093/nar/gkw1350.

Kindermann, A., Stepanova, E., Hund, H., Geis, N., Malone, B., Dieterich, C. (2019). MedEx - Data Analytics for Medical Domain Experts in Real-Time. *Stud Health Technol Inform*, 267, 142–149, doi:10.3233/SHTI190818.

Kontakt:

Prof. Dr. rer. nat. Christoph Dieterich
Professur für Bioinformatik und Systemkardiologie,
Universitätsklinikum Heidelberg,
Klinik für Kardiologie,
Angiologie und Pneumologie &
Klaus Tschira Institute for Computational Cardiology
christoph.dieterich@uni-heidelberg.de

https://dieterichlab.org

Dr. rer. nat. Tobias Jakobi,
Postdoctoral Fellow,
Universitätsklinikum Heidelberg,
Klinik für Kardiologie,
Angiologie und Pneumologie &
Klaus Tschira Institute for Computational Cardiology
tobias.jakobi@med.uni-heidelberg.de

https://tobias.jako.bi

das virtuelle gehirn

Interview mit Petra Ritter, der BIH-Johanna-Quandt-Professorin für Gehirnsimulation am Berlin Institute of Health (BIH) und an der Charité

In einem modernen Büro am Robert-Koch Platz am Campus Charité Mitte arbeitet Petra Ritter. Sie ist BIH-Johanna-Quandt-Professorin für Gehirnsimulation am Berlin Institute of Health (BIH) und an der Charité. Petra Ritter leitet ein internationales Projekt namens „The Virtual Brain“, das virtuelle Gehirn. Ihr Ziel ist es, mithilfe von Computersimulationen das Gehirn besser zu verstehen. Mit ihr sprach Stefanie Seltmann.

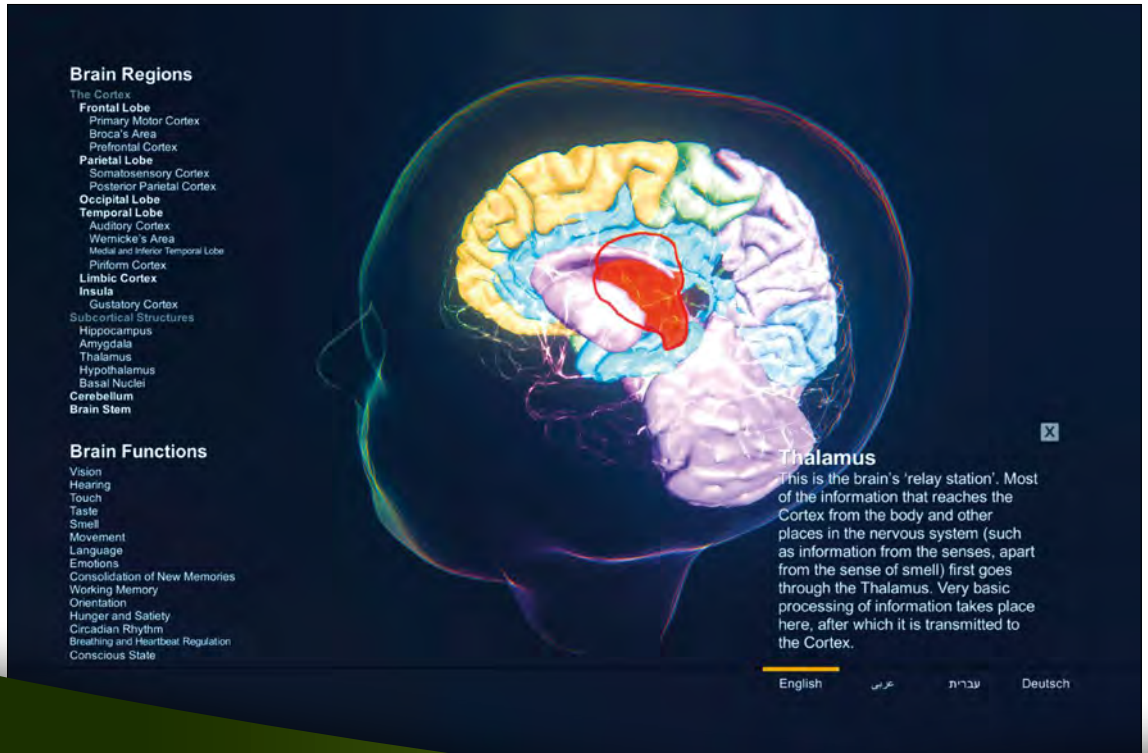
gesundhyte.de: *Frau Ritter, was verbirgt sich hinter „The Virtual Brain“?*

Prof. Dr. med. Petra Ritter: Hinter dem Virtual Brain verbirgt sich eine Gehirnsimulationsplattform, die von unserem internationalen Konsortium entwickelt wurde und seit 2012 öffentlich verfügbar ist. Das heißt, jeder Wissenschaftler und jede Wissenschaftlerin kann sie benutzen, und man kann damit personenbezogene Gehirne konstruieren und am Computer simulieren.

Normalerweise erforscht man ja das Gehirn mit EEG, MRT oder einer PET-Untersuchung und kann damit auf die Aktivität im Gehirn oder sogar einzelner Nervenzellen Rückschlüsse ziehen. Und diese Informationen speisen Sie in den Computer ein?

Ja, wir integrieren diese Informationen von einzelnen Personen in mathematische Modelle vom Gehirn. Dazu vereinfachen wir zuerst einmal sehr, wir wollen ein Modell bauen für ein Gehirn, das bestimmte Merkmale hat, die für eine Fragestellung, die uns gerade interessiert, entscheidend sind.

Ein digitaler interaktiver Gehirnatlas steht in vielen Sprachen unter https://www.brainsimulation.org/atlasweb/ zur Verfügung



Quelle: https://www.brainsimulation.org/bsw/zwei/research-visualization

Prof. Dr. med. Petra Ritter (Foto: David Ausserhofer)

Ein Gehirn hat einhundert Milliarden Nervenzellen, die untereinander verschaltet sind mit noch viel mehr Synapsen. Die können Sie nicht alle abbilden...

Nein, mit dieser Menge von interagierenden Elementen käme kein Supercomputer der Welt und auch nicht alle zusammen zurecht. Daher müssen wir reduzieren. Viele Nervenzellen sind synchronisiert, sie sind miteinander verbunden. So wie bei einem Vogelschwarm nicht jeder Vogel in eine andere Richtung fliegt, sondern man kann den Pfad eines jeden Vogels mit dem Pfad des Gesamtschwarms bestimmen. Und so machen wir das auch für die Nervenzellen. Wir bündeln sie und können so die Komplexität des Gehirns deutlich reduzieren.

„Mit dieser Menge von interagierenden Elementen käme kein Supercomputer der Welt und auch nicht alle zusammen zurecht."

Nennen wir doch ein Beispiel, bei dem Sie das schon einmal durchgespielt haben. Was sieht man bei Alzheimer-Patienten mit Ihrem Modell?

Beispielsweise kann man mit PET sehen, dass sich das sogenannte Beta-Amyloid bei Patienten mit Alzheimererkrankung in vielen Bereichen des Gehirns ablagert. Wir wissen außerdem von Tierstudien und von zellulären Studien, dass in der Nähe dieser Proteinablagerungen die Funktion der hemmenden Nervenzellen gestört ist. Und beide Informationen bauen wir in unser Modell ein. Das heißt, wir haben die räumliche Verteilung des Proteins von den PET-Daten, und wir haben dazu ein kleines mathematisches Modell, das das Protein übersetzt in eine verminderte Aktivität der hemmenden Zellen in dem entsprechenden Areal. Und wenn wir das zusammenfügen und eine Simulation von diesem Patienten oder der Patientin starten, sehen wir, dass alleine durch die Berücksichtigung der Proteinablagerungen wir in der Simulation veränderte EEG, also elektromagnetische Signale bekommen, es kommt zu einer Verlangsamung. Und diese Verlangsamung sehen wir auch in den echten gemessenen Daten.

Jetzt würde man sich natürlich wünschen, gemäß dem BIH Motto „Aus Forschung wird Gesundheit", dass Sie vielleicht sogar simulieren könnten, welcher Eingriff oder welche medikamentöse Behandlung am besten helfen könnte?

Es gibt bereits eine multizentrische klinische Studie in Frankreich zur Epilepsie. Dreißig Prozent der Patienten und Patientinnen sprechen nicht auf Medikamente an, und ihre Option ist dann ein neurochirurgischer Eingriff. Um den gut zu planen, platziert man Elektroden im Gehirn. Weil man natürlich nicht

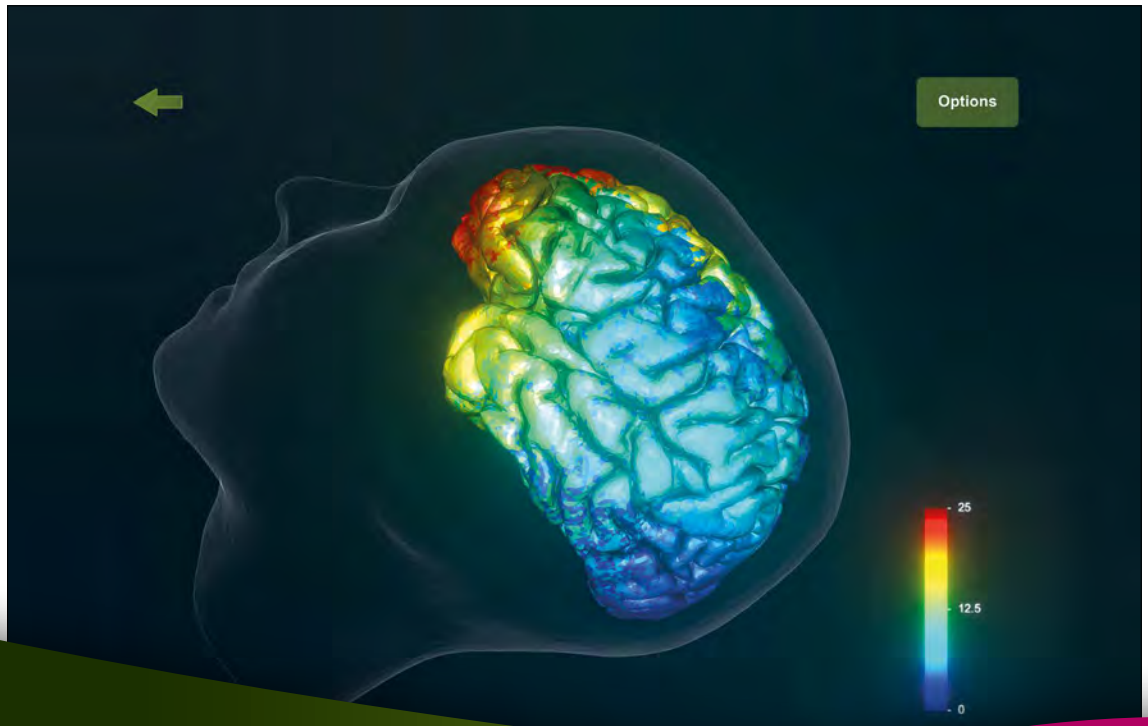
überall Elektroden platzieren kann, gibt es Lücken. Und das Modell kann mit allen vorhandenen Informationen simulieren, wie die Aktivität an den Stellen ist, von denen wir keine Informationen haben. Und da sind erste Hinweise publiziert worden, dass die Vorhersagen des Modells bezüglich der Zone, in der die krankhafte Aktivität im Gehirn startet und sich dann über verschiedene Regionen des Gehirns ausbreitet, besser sein könnten als die Schlussfolgerungen der Neurologen, wenn sie sich die Daten einfach nur anschauen. Und nun startet eine Studie in Frankreich mit vierhundert Patienten und Patientinnen. Die Hälfte wird unter Hinzunahme der Informationen aus dem Virtual Brain operiert, und die andere Hälfte unter Hinzunahme der klassischen Informationen, das heißt, das menschliche Expertenteam schaut sich die Daten an und zieht daraus Schlussfolgerungen. In knapp vier Jahren können wir sagen, ob das Computermodell tatsächlich bessere Vorhersagen liefert und nach der Operation die Anzahl der epileptischen Anfälle stärker reduziert ist als bei dem herkömmlichen Ansatz.

Könnte man denn an einem solchen Gehirnmodell auch Vorhersagen machen, wie Medikamente wirken?

„Die Hoffnung ist, dass man so in Zukunft am Computer bereits untersuchen kann, welches Medikament ganz spezifisch für einen bestimmten Patienten oder eine Patientin am besten wirksam ist."

Das haben wir am Alzheimer-Gehirn sogar schon demonstriert. Häufig wird das Medikament Memantin eingesetzt, das in vielen Fällen zumindest temporär eine Verbesserung der kognitiven Leistungsfähigkeit bei Patientinnen und Patienten hervorruft. Und da konnten wir am Computer zeigen, dass tatsächlich

Dieses Bild zeigt ein auf die kortikale Oberfläche des Gehirns projiziertes EEG



Quelle: https://www.brainsimulation.org/bsw/zwei/research-visualization

durch die virtuelle Gabe dieses Medikaments die Gehirne von den Alzheimer-Patienten eine normalisierte Funktion aufweisen. Die Hoffnung ist, dass man so in Zukunft am Computer bereits untersuchen kann, welches Medikament ganz spezifisch für einen bestimmten Patienten oder eine Patientin am besten wirksam ist.

Könnte man möglicherweise sogar eine erste klinische Studie an Ihrem Modell durchführen?

Ja, wir hoffen tatsächlich, dass in Zukunft bevor man bestimmte Interventionen an Menschen ausprobiert oder an Tieren diese Gehirnmodelle am Computer der erste Schritt sind und dass man so bereits beispielsweise Nebenwirkungen erkennen kann.

Inwieweit spielt die künstliche Intelligenz bei Ihrem Projekt eine Rolle?

Eine ganz wesentliche Rolle. Einerseits lernt die künstliche Intelligenz von unseren Modellen biologischer Netzwerke und deren Fähigkeit, Kognition zu erzeugen, und andererseits benutzen wir die künstliche Intelligenz, um bessere Modelle zu bauen und bessere Vorhersagen treffen zu können, beispielsweise für Patienten.

Die künstliche Intelligenz hilft der menschlichen Intelligenz dabei, die natürliche Intelligenz zu verstehen?

Das könnte man durchaus so sagen, ja.

Vielen Dank, Frau Ritter, für dieses Gespräch.

Gerne

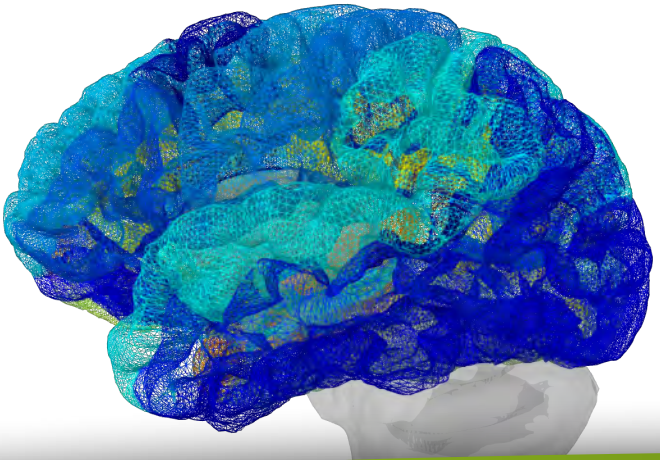
Das Gespräch führte Stefanie Seltmann.

Eine App für mobile Geräte wird ebenfalls im Labor von Frau Professor Ritter entwickelt (Quelle: https://www.brainsimulation.org/bsw/zwei/research-visualization#).

Kontakt:

Prof. Dr. med. Petra Ritter
Direktorin der Sektion Gehirnsimulation (CCM)
Charité – Universitätsmedizin Berlin
petra.ritter@charite.de

www.thevirtualbrain.org
https://brainsimulation.charite.de
www.brainsimulation.org
https://virtualbraincloud-2020.eu



Quelle: https://www.thevirtualbrain.org/tvb/zwei/brainsimulator-screens

WWW.THEVIRTUALBRAIN.ORG

modellaustausch für die regulatorische genomik MERGE

Förderung der gemeinsamen Nutzung und Wiederverwendung von Vorhersagemodellen in der Genomik durch Software-Standardisierung

von Julien Gagneur, Oliver Stegle, Michael J. Ziller

Die künstliche Intelligenz verändert unsere Möglichkeiten zur Entschlüsselung von menschlichen Genomen grundlegend. Neuartige Modelle in der regulatorischen Genomik, mithilfe derer sich die Auswirkungen von DNA-Variationen auf zelluläre Prozesse vorhersagen lassen, sind für die Aufklärung der Mechanismen von krankheitsverursachenden Varianten und Mutationen in Tumoren entscheidend. Um diese neuen Technologien vollumfänglich nutzen zu können, mangelt es jedoch an einer kohärenten Software-Infrastruktur, an Arbeitsabläufen und Schnittstellen. Diesem Defizit begegnet **MERGE** mit der Einrichtung einer Plattform für den Austausch, das Benchmarking und die Erweiterung von regulatorischen Genomik-Modellen. Mithilfe dieser Modelle sollen die molekularen Grundlagen von seltenen Erkrankungen, Krebs, psychischen und anderen Krankheiten erforscht werden.

Regulatorische Genomik

Die biomedizinische Gemeinschaft hat in den vergangenen zehn Jahren einen enormen Bestand an genomischen und molekularen Daten gesammelt, die als Grundlage für eine personalisierte Genominterpretation dienen. Hierzu zählen die Datenbestände internationaler Projekte wie ENCODE, Roadmap Epigenome, Blueprint, TCGA, FANTOM5 und GTEx, bei denen umfassende Karten des genetischen, epigenetischen (z. B. Ziller *et al.,* 2015) und transkriptionellen Zustands für ein breites Spektrum von Zelltypen, Individuen und Krankheiten erstellt wurden. In der Summe bieten diese umfassenden Ressourcen ein erhebliches Potenzial zur Entschlüsselung der funktionellen Rolle krankheitsbedingter genetischer Veränderungen in kodierenden und nicht-kodierenden Regionen des Genoms. Während Varianten in kodierenden Sequenzen häufig direkt zu einem Gewinn oder Verlust von Genfunktion führen, ist die Rolle von nicht-kodierenden genetischen Varianten nur sehr schwer zu entschlüsseln. Insbesondere können diese Varianten auf verschiedene Ebenen der Genregulation einwirken und Veränderungen der dreidimensionalen Genomarchitektur, der Transkriptionsebenen, des RNA-Splicing, der RNA-Stabilität oder der Translationsfrequenz mit sich bringen.

Bisher wurden verschiedene Strategien angewandt, um die Rolle nicht-kodierender Sequenzen bei der Genregulation zu erforschen. Eine dieser Strategien besteht darin, genetische Varianten zu identifizieren, die mit genregulatorischen Veränderungen assoziiert sind. Hierbei werden große Kohorten mit abgestimmten molekularen Profilen (z. B. Genexpressionsprofilen) und genetischen Daten eingesetzt. Solche assoziationsbasierten Ansätze liefern jedoch nur begrenzt mechanistische Erkenntnisse über die molekularen Grundlagen spezifischer genetischer Varianten. Außerdem erlauben sie aufgrund von Leistungseinschränkungen keine Einblicke in seltene genetische Varianten, wie sie beispielsweise in einzelnen Tumorzellen oder bei seltenen Krankheiten auftreten können. Eine zweite Strategie basiert auf biophysikalischen Modellen genregulatorischer Mechanismen, die beispielsweise die Bindungsaffinität von Transkriptionsfaktoren an die DNA-Sequenz modellieren. Die biophysikalischen Modelle reichen jedoch nicht aus, um die tatsächliche Komplexität der genregulatorischen Mechanismen *in vivo* realistisch zu erfassen.

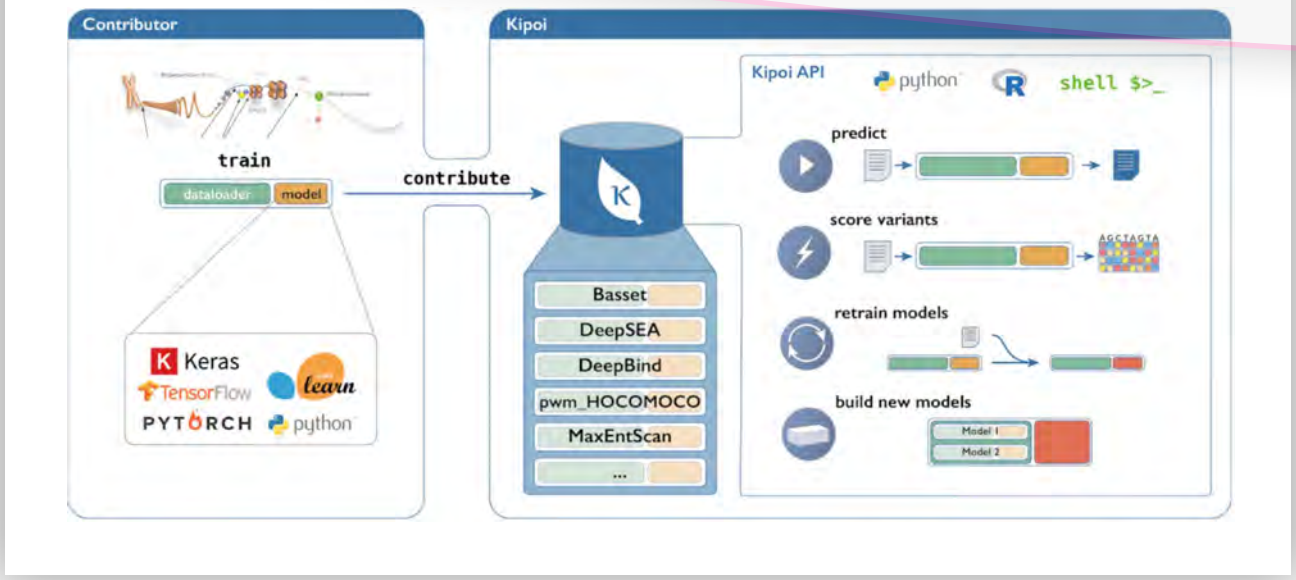


Abbildung 1: Kipoi-Konzept. Über die standardisierte Schnittstelle für regulatorische Genomik-Modelle können Modelle gemeinsam genutzt bzw. geteilt und für neue Daten und die Modellkomposition verwendet werden. Kipoi erleichtert die nachgeschaltete Analyse von Modellen, einschließlich Vorhersagen. Das Modell-Repository ist unter https://kipoi.org/ verfügbar (Quelle: Roman Kreuzhuber, Kipoi team).

Die Deep-Learning-Revolution

Eine dritte Strategie, die konzeptionell zwischen biophysikalischer Modellierung und assoziationsbasierten Ansätzen liegt, setzt Methoden der künstlichen Intelligenz ein. Bioinformatiker haben gezeigt, dass sich Datenrepräsentationen, die bei der Bild- und Sprachverarbeitung genutzt werden, auch auf Genomdaten anwenden lassen. Damit können Methoden des Deep Learning – ein relativ neuer Bereich der Künstlichen Intelligenz – genutzt werden, um Beziehungen zwischen Genotyp und molekularen Phänotypen zu modellieren. Bei Deep-Learning-Modellen handelt es sich um Schichten aus flexiblen Abbildungen, die durchgehend ineinandergreifen, um aus einer Eingabe (z. B. einer DNA-Sequenz) eine Ausgabe (z. B. die RNA-Häufigkeit) vorherzusagen. Im Gegensatz zu genetischen Assoziationsmodellen lassen sich Deep-Learning-Modelle auf genetische Varianten anwenden, die in den Trainingsdaten nicht erkannt wurden. So können sie zum einen auf seltene Varianten und de-novo-Mutationen generalisieren, was im Zusammenhang mit seltenen Krankheiten wichtig ist, und zum anderen auf somatische Mutationen, die nur bei bestimmten Tumorarten auftreten. Darüber hinaus bauen Deep-Learning-Modelle, anders als biophysikalische Modelle, nicht auf starke Modellannahmen, sondern lernen die Zusammenhänge direkt von den beobachteten Daten. Neue, leistungsstarke Software- und Hardware-Infrastrukturen (Deep-Learning-Frameworks wie Pytorch, Tensorflow und Keras bzw. Graphical Processing Units) bieten außerdem die Möglichkeit, komplexe Modelle mit riesigen Datenmengen zu trainieren, wie sie unter anderem von großen internationalen Genomik-Projekten zur Verfügung gestellt werden. Deep-Learning-Modelle sind die modernsten Vorhersagemodelle für eine Vielzahl genregulatorischer Bereiche geworden (Eraslan *et al.,* 2019), darunter Bindungsstellen für Transkriptionsfaktoren, Chromatinmodifikationen (z. B. Angermueller *et al.,* 2017), DNA-Kontakt-Maps, Genexpression, Spleißen (z. B. Cheng *et al.,* 2019), RNA-Abbau und Translation.

Teilen wir die Modelle!

Obwohl Modelle für zentrale Aufgaben in der regulatorischen Genomik in zahlreichen Veröffentlichungen beschrieben werden, bleibt ihr Potenzial unausgeschöpft. Dies liegt zum großen Teil daran, dass es bisher keinen kohärenten Rahmen für die gemeinsame Nutzung und den Austausch von Modellen in der wissenschaftlichen Community gibt. Verschiedene Faktoren haben dazu beigetragen, dass Modelle nur eingeschränkt ausgetauscht wurden:

- **Es fehlen einheitliche Modellschnittstellen für den Umgang mit genomischen Daten.**
- **Die Frameworks für maschinelles Lernen sind heterogen, die Modelle sind abhängig von spezifischer Software.**
- **Maschinelle Lernmodelle müssen interpretierbar sein, damit aus ihnen biologische Erkenntnisse gewonnen werden können.**
- **Die Methoden sind nicht ausreichen vergleichbar und lassen kein objektives Benchmarking zu.**
- **Die Wiederverwendung und Anwendung bestehender Modelle auf neue Daten ist zu komplex.**
- **Die Eintrittsbarriere für Bioinformatiker, die keine Experten für maschinelles Lernen sind, ist zu hoch.**

Mit der Entwicklung des Programmierstandards und Modell-Archivs Kipoi (Avsec *et al.,* 2019) begegnen wir den genannten Problemen. Kipoi ist ein so genannter „Modellzoo“: ein Repository (Archiv) mit Hunderten von ausgebildeten Modellen (Abbildung 1). Unsere Lösung richtet sich an Entwickler, die ihre Modelle hinterlegen und aus den vorhandenen Bausteinen neue Modelle ableiten können. Das Hauptaugenmerk liegt jedoch auf den Nutzern. Kipoi-Modelle werden mit Code zur Vorverarbeitung und zum Laden von Eingabedaten in den wichtigsten Dateiformaten geliefert. So lassen sich bestehende Modelle in wenigen Codezeilen, über Python, R oder über die Befehlszeile auf neue Daten anwenden. Die Modelle können auch als Grundlage genutzt werden und auf neue Datensätze und Aufgaben abgestimmt werden. Da Kipoi-Modelle standardisiert sind, lassen sich außerdem Modelle aus heterogenen Ursprüngen auf einfache Weise vergleichen und neue Modelle auf der Grundlage bestehender Modelle erstellen. Damit gestattet der Standard eine echte kompositorische Modellierung. Auch generische Routinen für die nachgelagerte Analyse und die Interpretation von Kipoi-Modellen werden durch diese Modellstandards ermöglicht. Kipoi unterstützt bereits die wichtigsten Frameworks für maschinelles Lernen und wendet das FAIR-Sharing-Prinzip („*Findable, Accessible, Interoperable, Reusable*“), das für Daten definiert wurde, auf trainierte Machine-Learning-Modelle an.

Die MERGE-Arbeitspakete

MERGE ist ein Gemeinschaftsprojekt der Gruppen von Julien Gagneur (TUM) und Oliver Stegle (DKFZ), den Initiatoren von Kipoi, sowie von Michael Ziller vom MPI für Psychiatrie. Das MERGE Projekt wird den Funktionsumfang und die Leistungsfähigkeit von Kipoi und die darin enthaltenen Modelle erweitern (Abbildung 2). Neu zu entwickelnde Funktionen werden es ermöglichen Kipoi Modelle zukünftig auch in Cloud-Umgebungen wie der BMBF-geförderten de.NBI-Cloud zu nutzen. Parallel zu Verbesserungen der Infrastruktur wird die Interpretierbarkeit der Modelle weiter verbessert. In diesem Zusammenhang werden wir das Modell-Repository mit Modellen für Transkriptionsverstärker, Spleißen und Translation erweitern. Diese werden an Hochdurchsatz-Assays für genetische Störungen sowie an experimentelle Daten angepasst, die natürliche genetische Variationen untersuchen. Darüber hinaus wird der Kipoi-Modellzoo anwendbar auf die Interpretation von Krebsgenomen sowie von genetischen Varianten, die im Zusammenhang mit häufigen und seltenen Krankheiten auftreten. Die MERGE-Teammitglieder arbeiten in laufenden Kooperationsprojekten auf diesen drei Krankheitsgebieten und unterstützen die kontinuierliche Zusammenarbeit im Rahmen von MERGE am Deutschen Krebsforschungszentrum in Heidelberg, am MPI für Psychiatrie, am Deutschen Herzzentrum und am Institut für Humangenetik der TUM.

Ausblick

Die in MERGE entwickelten Modelle und Ansätze werden entscheidend dazu beitragen, das exponentielle Wachstum der Humangenomdaten, das in den kommenden Jahren erwartet wird, voll auszunutzen. Projekte wie die „1+ Million Genomes Initiative“ der EU (https://www.pubaffairsbruxelles.eu/germany-joins-the-1-million-genomes-initiative-eu-commission-press/), der auch Deutschland beigetreten ist, könnten von diesen Fortschritten und zukünftigen Entwicklungen profitieren. Von zentraler Bedeutung wird es dabei sein, parallel zu den neuartigen KI-Technologien und Werkzeugen auch den Zugang zu Dateninfrastrukturen zu fördern, die diese wichtigen Datenressourcen enthalten. Das „Deutsche Humangenom-Phenom-Archiv“ (http://www.ghga.de) ist ein Beispiel für eine neu entstehende Infrastruktur für Humangenomdaten, die derzeit im Rahmen der Nationalen Forschungsdateninfrastruktur (NFDI) aufgebaut wird. Fortschritte aus Projekten wie MERGE werden auch in die nationale Bioinformatik-Infrastruktur-Initiative de.NBI (https://www.denbi.de/) eingebunden. Die Integration von KI-Methoden und Werkzeugen mit Dateninfrastrukturen stellt eine im weiteren Sinne wichtige Herausforderung der Zukunft dar.

Steckbrief Forschungsprojekt:

Projekt: MERGE – Model Exchange for Regulatory Genomics (Modellaustausch für die regulatorische Genomik)
Fördermaßnahme: CompLS – Computational Life Sciences (https://www.gesundheitsforschung-bmbf.de/de/compls-computational-life-sciences-9161.php)
Partner: Julien Gagneur (Technische Universität München, München), Oliver Stegle (Deutsches Krebsforschungszentrum, Heidelberg), Michael Ziller (Max-Planck-Institut für Psychiatrie, München)

Referenzen:

Angermueller, C., Lee, H.J., Reik, W., und Stegle, O. (2017). DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning. Genome Biol. *18*, 67.

Avsec, Ž., Kreuzhuber, R., Israeli, J., Xu, N., Cheng, J., Shrikumar, A., Banerjee, A., Kim, D.S., Beier, T., Urban, L., *et al.* (2019). The Kipoi repository accelerates community exchange and reuse of predictive models for genomics. Nat. Biotechnol. *37*, 592–600.

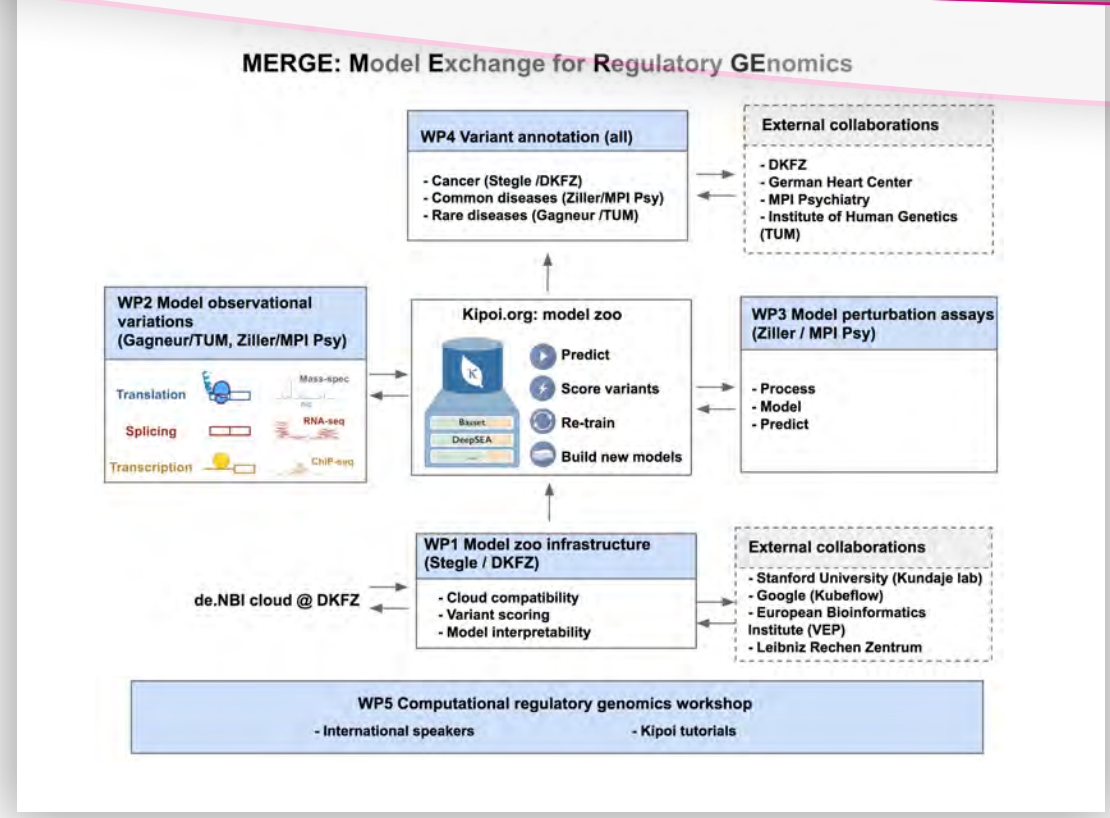


Abbildung 2: Das MERGE-Projekt: Ansatz und Überblick. Im Mittelpunkt der Planung stehen Erweiterungen von Kipoi, um (1) Cloud-Komponenten in Kipoi zu integrieren, (2) die Funktionalität von Kipoi auszubauen, um die Modellierung von molekularen Auswirkungen natürlicher Variationen aus Beobachtungsdaten ebenso zu ermöglichen wie (3) die effiziente Nutzung von Perturbation Assays und (4) die Verwendung von Modellen für die Interpretation menschlicher genetischer Variation. MERGE ist in ein starkes Netzwerk von Kooperationspartnern eingebunden (Quelle: Julien Gagneur).

Cheng, J., Nguyen, T.Y.D., Cygan, K.J., Çelik, M.H., Fairbrother, W.G., Avsec, Ž., und Gagneur, J. (2019). MMSplice: modular modeling improves the predictions of genetic variant effects on splicing. Genome Biol. *20*, 48.

Eraslan, G., Avsec, Ž., Gagneur, J., und Theis, F.J. (2019). Deep learning: new computational modelling techniques for genomics. Nat. Rev. Genet. *20*, 389–403.

Ziller M.J., Reuven E., Yaffe Y., Donaghey J., Pop R., Mallard W., Issner R., Gifford C.A., Goren A., Xing J., Gu H., Cacchiarelli D., Tsankov A., Epstein C., Rinn J.L., Mikkelsen T.S., Kohlbacher O., Gnirke A., Bernstein B.E., Elkabetz Y., Meissner A. (2015), Dissecting neural differentiation regulatory networks through epigenetic footprinting. Nature *518*, 355–359.

Kontakt:

Prof. Julien Gagneur
Lehrstuhlinhaber
Lehrstuhl für Computational Molecular Medicine
Technische Universität München
gagneur@in.tum.de

https://www.in.tum.de/gagneurlab

Dr. Oliver Stegle
Abteilungsleiter
Abteilung Bioinformatik der Genomik und Systemgenetik
Deutsches Krebsforschungszentrum Heidelberg & Europäisches Molecular Biologie Laboratorium Heidelberg
o.stegle@dkfz.de

https://www.dkfz.de/en/bioinformatik-genomik-systemgenetik

Prof. Dr. Michael Ziller
Functional Genomics in Psychiatry
WWU Münster
und Forschungsruppenleiter
Genomik komplexer Erkrankungen
Max-Planck-Institut für Psychiatrie
München
michael_ziller@psych.mpg.de

https://www.psych.mpg.de/2164378/ziller

Fotos: privat / Gagneur, privat / Stegle, privat / Ziller

„bench to bedside“: systempharmakologie in der praxis

Wie computerbasierte Methoden die digitale Transformation der Pharmaforschung und Therapie vorantreiben

Firmenporträt der esqLABS GmbH

von Stephan Schaller

Die Erwartung an die Pharmaindustrie kostengünstigere Behandlungen für komplexe Krankheitsbilder anzubieten wird zunehmend zur Herausforderung. Lösungen hierfür bietet das „Fast Prototyping“, sowie die „Personalisierte Medizin“, die es ermöglichen, Forschungs- und Entwicklungs- (F&E-) Zeiten zu verkürzen und optimale Therapien für einzelne Patienten zu finden. Eine zentrale Voraussetzung hierfür ist eine verstärkte Integration der Datenverarbeitung und -analyse in der pharmazeutischen F&E, als auch in der klinischen Praxis. Softwarelösungen zur Analyse der (klinischen) Daten sollen hierbei helfen schneller informierte Entscheidungen zu treffen.

esqLABS
WE EMPOWER HEALTH CARE

„Fast Prototyping“ in der Pharmazeutischen F&E

Die Erforschung neuer Medikamente ist ein langwieriger, wissenschaftlich hoch komplexer und kostenintensiver Prozess.

Aufgrund des rasanten Anstiegs an Datenmengen, die in der F&E erhoben werden, ist seit einigen Jahren in der Pharmaindustrie eine steigende Anerkennung für – und ein stark steigender Bedarf an – Computermodelle(n) zur Datenanalyse und -verarbeitung zu verzeichnen (EFPIA MID3 Workgroup *et al.,* 2016). Derzeit existieren zwar hochspezialisierte Einzellösungen für die verschiedenen Abschnitte entlang der Wertschöpfungskette (Abbildung 1, unten), allerdings sind dadurch Wissenstransfer und Transparenz entlang der Wertschöpfungskette nicht gewährleistet.

Doch gerade der Know-how- und Wissenstransfer über die verschiedenen Forschungs- und Entwicklungsphasen hinweg sowie die erfolgreiche Demonstration der Wirksamkeit eines Medikaments in klinischen Studien sind die zentralen Herausforderungen bei der F&E in der Pharmaindustrie (Goldblatt and Lee, 2010).

esqLABS hat hierbei den Wert in der Nutzung von physiologiebasierter quantitativer System-Pharmakologie (PB-QSP) Modellplattformen erkannt, welche die Medikamentenbewegung und Wirkung im Körper erfassen. Die Architektur dieser Modelle ermöglicht es eine Vielzahl unterschiedlicher Daten aus Zellkulturen, Tieren, einzelner Patienten und Patienten-Populationen, die in der Pharmaforschung erhoben werden zu integrieren (Abbildung 1). Die Datenintegration in einer einzelnen Plattform fördert den Wissenstransfer und -erhalt entlang des F&E Prozesses, und führt zur Qualitäts- und Effizienzsteigerung, und macht die Plattformen zuverlässiger in ihrer Vorhersage von klinischen Anwendungen als bisherige Technologien (Jones *et al.,* 2006).

Mit diesen Modellen werden von esqLABS durch Integration von Forschungsdaten über translationale Modellierungsansätze

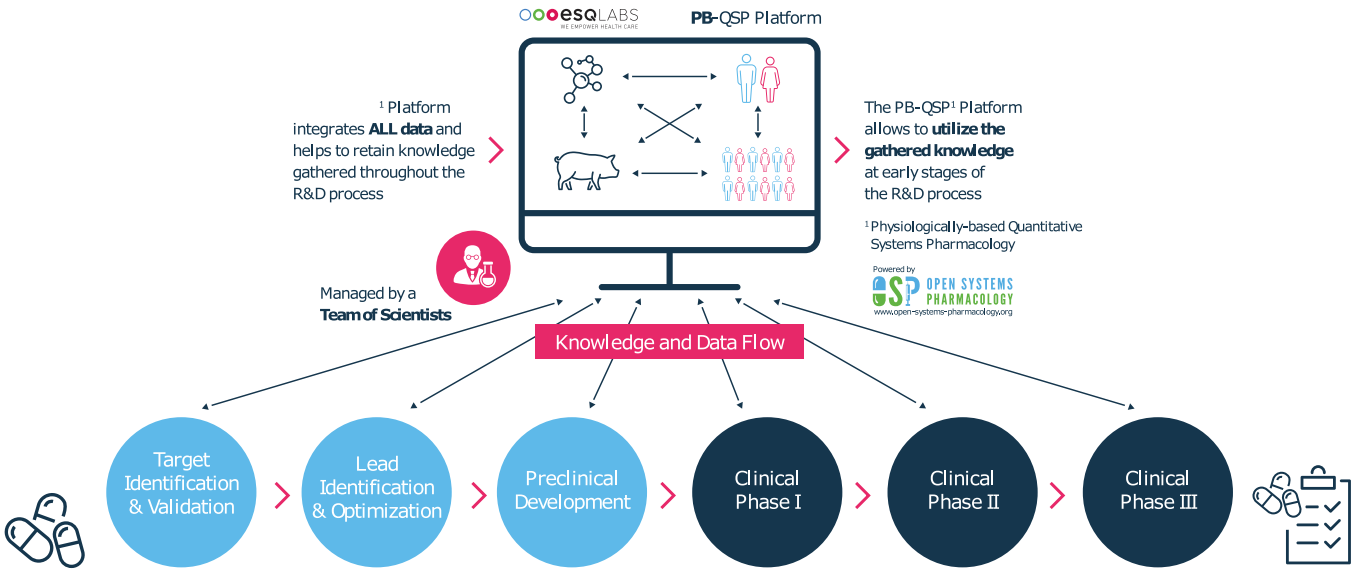


Abbildung 1: Einsatz von Computermodellen im Forschungs- & Entwicklungsprozess in der Pharmaindustrie. Dem „state-of-the-art"-Ansatz des „Flickenteppichs" aus QSAR, Systembiologie, PK/PD, herkömmlichen QSP Plattformen und popPK/PD Modellen wird der „eine Plattform für Alles"-Ansatz mit PBPK/PD QSP Plattformen von esqLABS gegenübergestellt (Quelle: esqLABS GmbH).

die unterschiedlichsten Extrapolationsszenarien in der Medikamentenentwicklung untersucht. Hierzu gehören die Prognose der Medikamentenwirkung im Menschen basierend auf *in-vitro* und Tierversuchen, die Vorhersage von Medikamenteninteraktionen, sowie den Auswirkungen von Alter, Krankheit und Erbgut auf die Verteilung und Wirkung von Medikamenten.

Die in der Plattform enthaltenen Computermodelle werden in Kooperation mit Pharmafirmen auf ihre spezifischen F&E-Prozesse angepasst und helfen durch Simulationen bei der Entscheidungsfindung in der Medikamentenentwicklung, verkürzen F&E-Zyklen und bergen dadurch erhebliche Kosteneinsparpotenziale.

Personalisierte Medizin

Therapeutische Entscheidungsfindungssysteme (TCDS) im Krankenhaus nutzen Krankheits-Computermodelle für die Berechnung der optimalen Medikamentendosis für einzelne Patienten. Dies ermöglicht den Paradigmenwechsel von der populationsbasierten Medizin zur personalisierten Medizin.

Die von esqLABS entwickelten PB-QSP Plattformen und darauf aufgebaute Lösungen werden es ermöglichen, die Menge und Wirkung von Medikamenten personalisiert für den Patienten vorab zu berechnen. Dies hilft Fehler in der Kommunikation und in der Dosisberechnung zu vermeiden, Druck vom medizinischen Personal zu nehmen, und die Effizienz der Abläufe zu erhöhen.

Ein großer Vorteil wird bei der Therapie mit Medikamenten von geringer therapeutischer Breite erwartet, da bei diesen Medikamenten schon kleine Abweichungen in der Dosierung lebensbedrohliche Folgen haben könnten (Blix *et al.*, 2010; FDA).

Die Plattform unterstützt somit die präzise Dosierung von komplexen Medikationen. Die Individualisierung der Behandlung für den einzelnen Patienten führt damit auch bei der Gesamtheit der Patienten zu einer besseren Behandlung (Stern *et al.*, 2016). Zusätzlich zum direkten Nutzen der Therapie können durch TCDS Systeme auch Folgekosten durch Fehldosierungen, wie z. B. einer Überdosis, und deren Folgebehandlung, ggf. einer notwendigen Hospitalisierung vermieden werden (*Blix et al.*, 2010).

Simulationsmodelle für Forschung und Therapie

Zur Umsetzung des „Fast Prototyping" sowie der personalisierten Medizin ist es notwendig komplexe biologische Zusammenhänge in mathematischen Modellen zu kontextualisieren. esqLABS entwickelt hierfür Lösungen, die auch zukünftige Anforderungen in der (Gesundheits-) Wissenschaft und Wirtschaft

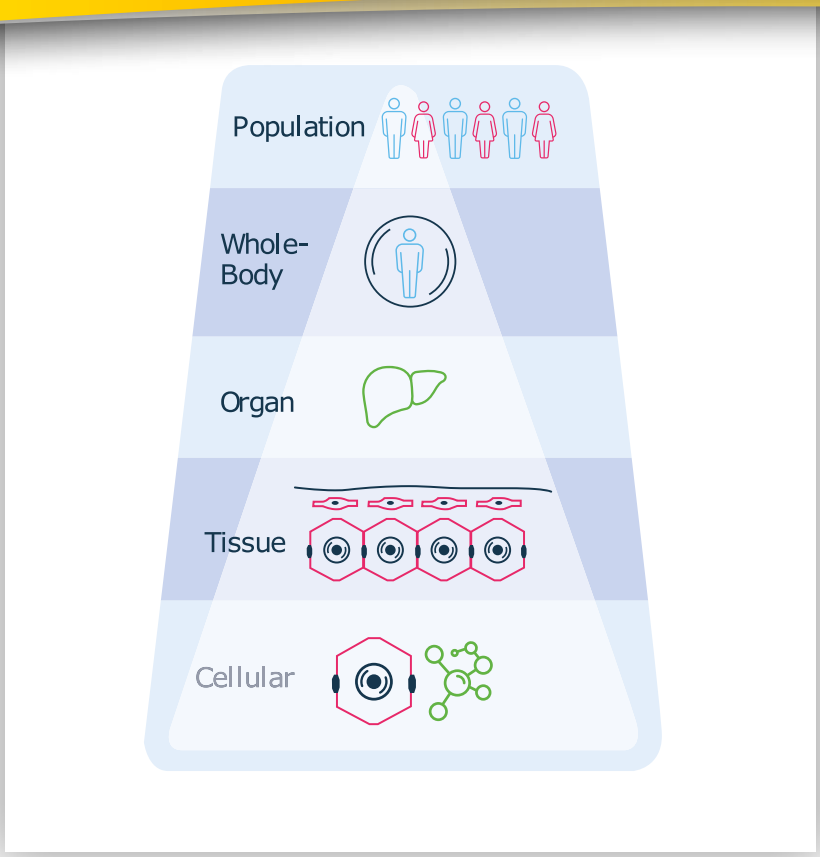


Abbildung 2: Multi-Skalen Modellierung und -Simulation. Die Softwareplattform Open Systems Pharmacology Suite wurde zur Modellierung und Simulation biologischer Prozesse mit Schwerpunkt auf Pharmakokinetik, Pharmakodynamik und Krankheitsprogression (einschließlich biochemischer Reaktionsnetzwerke) entwickelt. Dadurch ermöglicht die Plattform die Kombination mehrerer organisatorischer und physiologischer Skalen, von zellulären Prozessen, bis hin zu Populationen (Quelle: esqLABS GmbH).

berücksichtigen. Die esqLABS GmbH baut bei der Entwicklung ihrer Dienstleistungen auf der open-source Technologie der Open-Systems-Pharmacology Software-Plattform (www.open-systems-pharmacology.org) auf und entwickelt Lösungen um diese in der personalisierten Medizin einzusetzen.

Ein Beispiel hierfür ist die Weiterentwicklung der „Open Systems Pharmacology Suite“ (OSPS), einer open-source Software, die zur Integration und Analyse komplexer biomedizinischer Daten eingesetzt wird. Die Integration komplexer Krankheitsprozesse bis hin zu zellulären Interaktionen sowie die Simulation verschiedenster Anwendungsszenarien, auch auf Populationsebene (Abbildung 2) können dazu führen, dass Modell-Handhabung und -Qualitätssicherung zu einer Herausforderung werden. Diese Herausforderung soll durch ein Modularisierungskonzept sowie durch eine automatisierte Qualifizierungsroutine gelöst werden.

esqLABS bündelt mit ihren Arbeiten alle F&E-Abschnitte auf einer Plattform mit einem einheitlichen durchgängigen Computermodell. Im Vergleich zu bisherigen Ansätzen ermöglicht dies zuverlässigere und schnellere Berechnungen für effizientere Entscheidungen in der Pharmaforschung, sowie auch bei der Anwendung von Medikamenten.

Firmenprofil:

Gegründet: Juni 2017

Anzahl Mitarbeiter: 8

Produkte: Software, Beratung und Services im Bereich der computerbasierten Modellierung und Simulation von Medikamentenwirkung und Krankheitsprozessen:

- **PBPK-basierte Analysen der systemischen Verteilung (Pharmakokinetik (PK)) von Medikamenten**
  - Altersabhängigkeit der PK (Pediatrische Untersuchung)
  - Abhängigkeit der PK von Co-Medikationen (DDI)
  - Pharmakogenomische Abhängigkeit der PK
  - Abhängigkeit der PK von Organschädigungen (Leber/Niere)
- **Quantitative systempharmakologische (QSP) Analyse der Medikamentenwirkung (Pharmakodynamik (PD)) auf unterschiedliche Krankheiten:**
  - Analyse des Medikamenten-Wirkungsprofils
  - Analyse optimaler individueller Medikation

Computerbasierte Methoden treiben die digitale Transformation der Pharmaforschung und Therapie voran (Quelle: esqLABS GmbH).

Referenzen:

Blix, H.S., Viktil, K.K., Moger, T.A., and Reikvam, A. (2010). Drugs with narrow therapeutic index as indicators in the risk management of hospitalised patients. Pharm. Pract. 8, 50–55.

Coorevits, P., Sundgren, M., Klein, G.O., Bahr, A., Claerhout, B., Daniel, C., Dugas, M., Dupont, D., Schmidt, A., Singleton, P., *et al.* (2013). Electronic health records: new opportunities for clinical research. J. Intern. Med. 274, 547–560.

EFPIA MID3 Workgroup, Marshall, S.F., Burghaus, R., Cosson, V., Cheung, S.Y.A., Chenel, M., DellaPasqua, O., Frey, N., Hamrén, B., Harnisch, L., *et al.* (2016). Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharmacomet. Syst. Pharmacol. 5, 93–122.

FDA, C. for D.E. and Generic Drug User Fee Amendments of 2012 - Narrow Therapeutic Index Drugs.

Goldblatt, E.M., and Lee, W.-H. (2010). From bench to bedside: the growing use of translational research in cancer medicine. Am. J. Transl. Res. 2, 1–18.

Jones, H.M., Parrott, N., Jorga, K., and Lavé, T. (2006). A novel strategy for physiologically based predictions of human pharmacokinetics. Clin. Pharmacokinet. 45, 511–542.

Stern, A.M., Schurdak, M.E., Bahar, I., Berg, J.M., and Taylor, D.L. (2016). A Perspective on Implementing a Quantitative Systems Pharmacology Platform for Drug Discovery and the Advancement of Personalized Medicine. J. Biomol. Screen. 21, 521–534.

Kontakt:

Dr. Stephan Schaller
Geschäftsführer
esqLABS GmbH
Saterland, Deutschland
Stephan.schaller@esqlabs.com

www.esqlabs.com

Foto: privat / Schaller

der roboter im OP

Die Roboter-assistierte Chirurgie (RAC): vom Gimmick zum Standard

von Klaus-Peter Jünemann

1961 installierte General Motors die ersten Roboter in der Autoindustrie. Heute werden praktisch alle Industriestraßen von Roboterarmen dominiert, weil wiederkehrende Prozesse und Arbeitsabläufe von Maschinen besser und präziser bewältigt werden als von Menschen. Roboter arbeiten heute auch als „Cobots“ in Interaktion mit dem Menschen. Digitalisierung und Mensch-Maschine-Interaktion sind die Schlüsseltechnologien des 21. Jahrhunderts. In der Industrie, der Telekommunikation, im Handel – in fast allen Lebensbereichen sind diese Technologien bereits selbstverständlich. Aber: brauchen wir den Roboter auch im OP? Können wir unser Leben einer Maschine anvertrauen?

Warum brauchen wir Roboter-assistierte Chirurgie

Es gibt sie schon, die sogenannte „Roboter-assistierte Chirurgie“ (RAC). Fragt sich: sind die heutigen Chirurgen, die sich dieser Technologie verschrieben haben, zu faul oder zu schlecht ausgebildet, um konventionell offen oder laparoskopisch zu operieren? Sind Roboter im OP nur ein „Nice-to-have“, ein Gimmick für den Chirurgen? Was hat der/die Patient/in davon? Warum sollte man sich mit Hilfe eines Roboters operieren lassen?

Zum besseren Verständnis: Roboter in der Medizin sind Telemanipulatoren nach dem Master-Slave-Prinzip, die Bewegungen der chirurgischen Instrumente werden vom Arzt gesteuert. Die Lernkurve ist im Vergleich zur herkömmlichen Laparoskopie steil d. h. die RAC kann deutlich schneller erlernt werden.

Prof. Dr. Klaus-Peter Jünemann
Direktor der Klinik für Urologie und Kinderurologie, Sprecher des Kurt-Semm-Zentrums für laparoskopische und roboterassistierte Chirurgie am UKSH, Campus Kiel (Foto: Felix Prell)

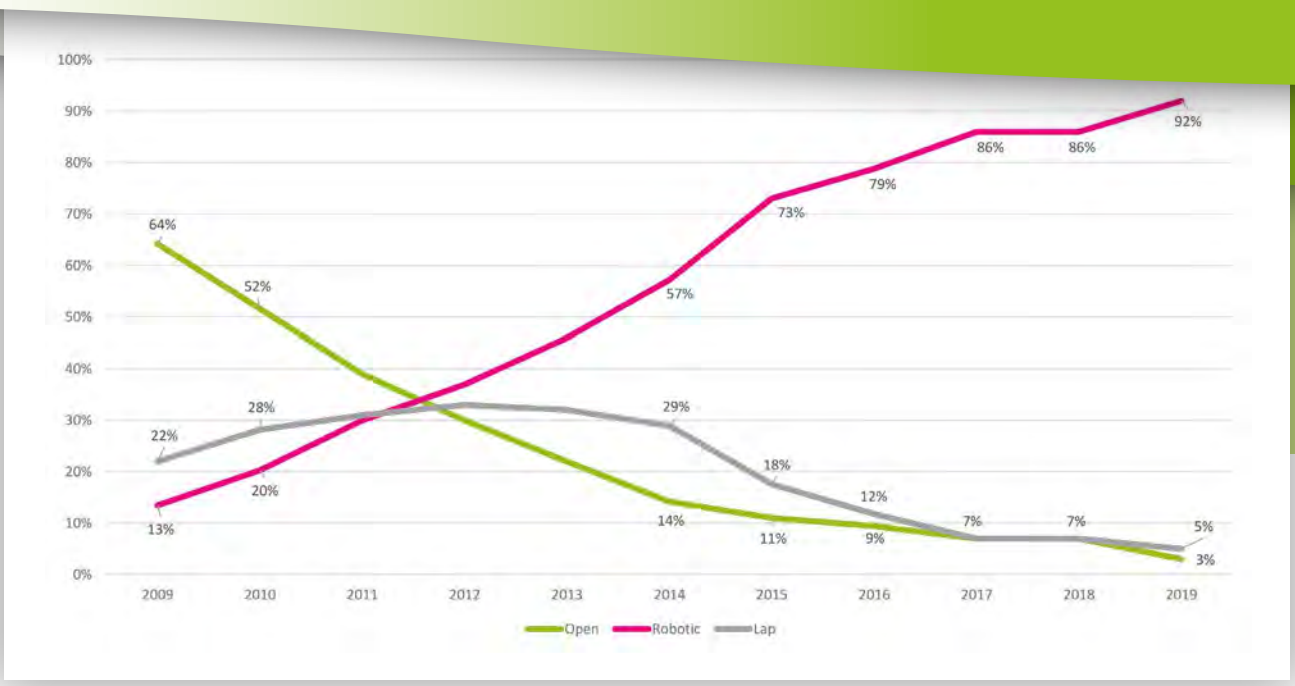


Abbildung 1: Hypothetical Evolution of Prostatectomies in the UK 2009 to 2019 (Quelle: Simmonds, 2020).

Anders als der gekrümmt stehende Laparoskopiker sitzt der Operateur bei Roboter-assistierten Eingriffen bequem an einer Konsole und steuert die endoskopischen Instrumente tremorfrei über Controller bzw. Fingerschlaufen. Wie das menschliche Handgelenk besitzen die sog. „Endowrist®"-Instrumente sieben Freiheitsgrade. Das insufflierte (mit CO_2-Gas gefüllte) OP-Gebiet bietet optimalen Bewegungsraum. Der Chirurg sieht das OP-Gebiet in 10-facher Vergrößerung, hochauflösend und in 3D und kann sensible Strukturen wie Blutgefäße und Nervenbahnen optimal schonen.

Die RAC kombiniert die Vorteile des maximalen Bewegungsraums der offenen Chirurgie mit den minimalinvasiven Zugangswegen der laparoskopischen Chirurgie. Bisherige Studien zeigen, dass die onkologischen Ergebnisse (z. B. in Nieren-, Prostata-, Blasenchirurgie, Kolon-, Pankreas- und Ösophaguschirurgie) mindestens gleichwertig sind.

Der Hauptvorteil für den Patienten/die Patientin liegt in den geringen Komplikationsraten (minus 50% gegenüber offenen Verfahren) und der schnelleren Heilung: kaum Wundheilungsstörungen durch minimal-invasive Zugänge und schneller wieder fit sein, das überzeugt auch Patientinnen und Patienten. Dazu kommt, dass das Risiko für Intensivaufenthalte und Reinterventionen sinkt. Ein tendenzieller Vorteil der RAC zeichnet sich auch in den funktionellen Ergebnissen ab, so z. B. bei der radikalen Prostatektomie im Hinblick auf den Erhalt von Kontinenz und Erektionsfähigkeit. Die Roboter-assistierte radikale Prostatektomie verdrängt nicht nur die offene, sondern auch die laparoskopische Alternative, wie an einem Beispiel aus dem Vereinigten Königreich zu sehen ist, wo 2019 bereits 92 % aller Prostatektomien Roboter-assistiert durchgeführt wurden (Abbildung 1).

Dass sich diese Entwicklung weg von der klassischen hin zur Roboter-assistierten Chirurgie nicht nur in der Urologie wiederfindet, bestätigt eine kürzlich erschienene Studie für die Allgemeinchirurgie im Raum Michigan an 169.000 Patienten. Während die ersten fünf Jahre seit Implementierung der Roboter-assistierten Chirurgie dadurch gekennzeichnet waren, dass nicht nur die Roboter-assistierten Prozeduren, sondern zunächst auch die laparoskopischen zu Lasten der offenen Chirurgie zugenommen hatten, so bestätigen die beiden letzten Jahre ebenfalls einen Abfall der laparoskopischen Eingriffe zugunsten Roboter-assistierter Verfahren (Abbildung 3).

Warum die konventionelle Chirurgie aussterben wird

Der Verdrängungswettbewerb durch die Roboter-assistierte Chirurgie wird sich weiter fortsetzen und erfasst schon heute neben Urologie und Allgemeinchirurgie und Gynäkologie auch die Fachdisziplinen Thorax-, Kiefer- und Kinderchirurgie. Die RAC ist eine disruptive Innovation, die die konventionelle Chirurgie fast vollständig ablösen wird. Dieser Prozess ist unaufhaltsam und zwangsläufig. Anders als konventionelle Verfahren ist die RAC ein digitales Verfahren, d. h. sie ist innovations- und damit auch zukunftsfähig. So können zukünftig Vorbefunde in das operative Bild implementiert werden (Augmented Reality), farbige Hervorhebungen von Tumorarealen und sensiblen Strukturen werden dem Operateur den Weg weisen, intelligente Sicherheitskonzepte werden festlegen, wo die Instrumente arbeiten dürfen und welche Strukturen geschont werden müssen. Zusätzlich bietet die digitale OP-Aufzeichnung die Möglichkeit, Künstliche Intelligenz und sog. Expert-Systeme zu implementieren. Weitere potenzielle Verbesserungen liegen in der Weiterentwicklung der chirurgischen Kamerasysteme und der Visualisierung des OP-Gebiets.

Abbildung 2: State of the Art – da Vinci Chirurgie mit dem Xi-System (Quelle: F. Prell/G. Böhler).

Durch die Digitalisierung der Chirurgie eröffnen sich fast unbegrenzte Möglichkeiten der Optimierung und Erweiterung bis hin zu ersten Initiativen in Richtung „Autonomes Operieren". Auch die Ausbildung der Chirurgen wird sich verändern und digitaler werden: weg von Tiermodellen und Kadaverstudien hin zu maßgeschneiderten perfundierten Organsystemen und immer detaillierter werdenden Simulationen. Die bestimmenden Faktoren des Wandels hin zu einer digitalen Chirurgie werden von außen aus den neuen Disziplinen der Künstlichen Intelligenz bzw. der Virtuellen und Mixed Reality kommen und die Möglichkeiten des chirurgischen Handelns in eine neue Dimension überführen.

Warum die RAC in Deutschland eher unterdurchschnittlich vertreten ist

Mit den in Deutschland bisher implementierten 150 da Vinci-Systemen werden pro Jahr bis zu 250 Eingriffe pro System durchgeführt. Das sind nur 37.500 Roboter-assistierte Operationen von den standardmäßig in Frage kommenden rund 280.000 Eingriffen p. a. in Urologie, Gynäkologie und Allgemeinchirurgie, also nur ca. 13 % der möglichen Eingriffe. Ähnliches gilt für die Verbreitung der Systeme. In den USA, dem Ursprungsland des bisher einzigen Herstellers robotischer Chirurgiesysteme, kommt ein System auf ca. 95.000 Menschen, in Deutschland sind es ca. 547.000. In einigen anderen europäischen Ländern ist die RAC bereits deutlich stärker verbreitet als in Deutschland (z. B. in Schweden mit 280.000 Einwohnern pro System).

Das Potential der Roboter-assistierten Chirurgie ist in Deutschland bisher erst unterdurchschnittlich ausgeschöpft; die Gründe dafür liegen an erster Stelle in der aktuellen Kostenunterdeckung roboter-assistierter Eingriffe durch die geltenden Fallpauschalen. Zudem ist die US-Firma Intuitive mit den sehr hochentwickelten da Vinci-Systemen bisher der einzige etablierte Anbieter am Markt. Aktuell stehen zahlreiche Hersteller von ähnlichen Chirurgie-Systemen unmittelbar vor der Markteinführung. Die Roboterchirurgie wird auch europäisch, dafür stehen Avatera-medical (D), DistalMotion (CH) Cambridge Medical Robotics (UK), und Transenterix (I). Die neuen Systeme stehen im Wettbewerb miteinander und werden die Innovationen auf diesem Gebiet befeuern, aber gleichzeitig auch unterschiedliche Nischen besetzen.

Warum die Roboter-assistierte Chirurgie bei der Pandemiebewältigung helfen kann

Die aktuelle Covid-19-Krise verdeutlicht eindrücklich, wie vulnerabel die Gesundheitssysteme sind und in welcher unmittelbaren Abhängigkeit die Wirtschaftssysteme dazu stehen. So hat die COVID-19 Pandemie auch gravierende Auswirkungen auf die klinische Versorgung und die Operative Medizin gehabt. Zum Vorhalten ausreichender Kapazitäten für COVID-19-Fälle wurden elektive Eingriffe auf unbestimmte Zeit verschoben und Betten wie Personal im großen Stil aus dem Regelbetrieb abgezogen. Diese und zusätzliche Bettenkapazitäten wurden für viel Geld als Leerbetten vorgehalten, unabhängig davon, ob sie gebraucht wurden oder nicht. Gleichzeitig hat die Pandemie das Infektionsrisiko für klinisches Personal deutlich erhöht und zu einem zusätzlichen Bedarf an Operationen durch nachzuholende Eingriffe und die Behandlung von Corona-Spätfolgen an Lunge und Herz geführt.

Die Bewältigung dieser und zukünftiger Pandemien erfordert eine Effizienzsteigerung in der klinischen Medizin, die es erlaubt bei gleichbleibender Bettenkapazität den Regelbetrieb aufrecht zu erhalten. Durch die drastische Verkürzung der Liegezeit und der Komplikationen, kann die RAC hier entscheidende Impulse setzen, da die stationären Betten kürzer und Intensivbetten seltener in Anspruch genommen werden müssen, - einhergehend mit den entsprechenden Vorteilen der Personaleinsparung und einer schnelleren Genesung für die Patienten. Gleichzeitig bietet die Roboter-gestützte Chirurgie den Vorteil, dass berührungsfrei über Konsolen operiert wird, im Sinne einer „non-touch surgery".

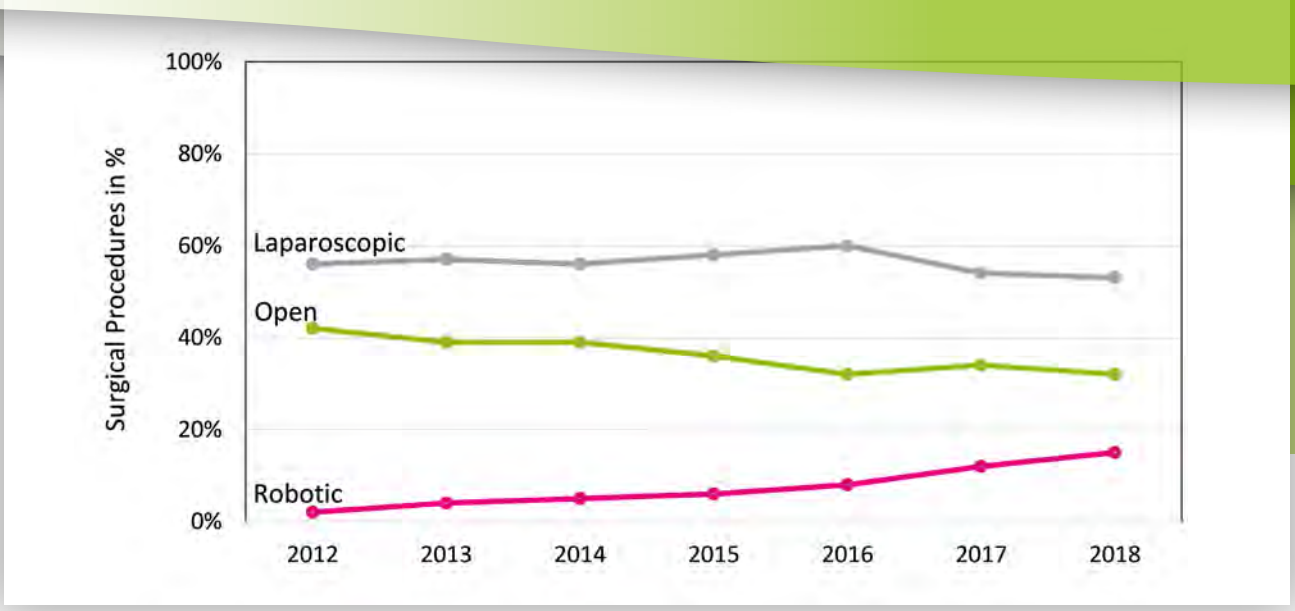


Abbildung 3: Temporal Trends in the Proportional Use of Robotic, Laparoscopic and Open Surgery (Quelle: Sheetz *et al.*, 2020).

Die kleinen Einstichstellen verringern zudem die Verbreitung von möglicherweise infektiösen Aerosolen aus dem Patientenkörper. Somit ist die Roboterchirurgie nicht nur der Schlüssel zu einer nachhaltigen Effizienzsteigerung im Gesundheitswesen, sie verringert auch das Infektionsrisiko im OP. Auf diese Weise könnte sich die COVID-19-Krise als Katalysator für die schnellere Verbreitung Roboter-assistierter Verfahren erweisen.

Warum in der Roboter-assistierten Chirurgie eine Chance für Deutschland liegt

Deutschland kann sich zukünftig nicht mehr blind auf seinen internationalen Ruf als „Land der Autoindustrie" verlassen. Gerade in der Coronakrise wurde deutlich, dass Deutschland international und auch zu Recht für seine Stärken in der Medizintechnik bewundert wird. Um sich als medizintechnischer Innovationsstandort zu behaupten, muss Deutschland eine führende Rolle in der Robotik, der Künstlichen Intelligenz und vor allem auch in der Medizintechnik übernehmen.

Die Roboter-assistierte Chirurgie ist eine Schlüsseltechnologie innerhalb der Medizintechnik und wird sich auch auf diagnostische Maßnahmen, wie z. B. Roboter-geführte Biopsien, ausdehnen. Ohne diese Technologie können innovative Verfahren aus dem Bereich von KI und Augmented Reality ihren Nutzen in Chirurgie und Diagnostik nicht entfalten. Die Möglichkeiten, operative Therapien – z. B. von Krebserkrankungen – mit Hilfe der Digitalisierung schonender und präziser zu machen, sind schier endlos. Wir stehen erst am Anfang. Die High-End-Medizintechnik kann gerade nach den Erfahrungen der Coronakrise einen Prestigegewinn verzeichnen und hat das Potential zum neuen Flaggschiff der deutschen Wirtschaft zu werden; – dazu benötigen wir eine nationale Strategie, die diese neuen Technologien nachhaltig vorantreibt.

Referenzen:

Stolzenburg, J.U., Kyriazis, I., Fahlenbrach, C., Gilfrich, C., Günster, C., Jeschke, E., Popken, G., Weißbach, L., von Zastrow, C., Leicht, H. (2016). National trends and differences in morbidity among surgical approaches for radical prostatectomy in Germany. World J Urol 34(11), 1515-1520. Epub 2016 Mar 24.

Cheng, C.L., und Rezac, C. (2018). The role of robotics in colorectal surgery. BMJ; 360: j5304. doi: 10.1136/bmj.j530.

Simmonds, Ch., BS, PGCE (2020): The Growth of Robotic Surgery. mimic Newsletter, Feb. 7, 2020

Sheetz K.H., Claflin J., Dimick, J.B. (2020). Trends in the Adoption of Robotic Surgery for Common Surgical Procedures. JAMA Netw Open. 2020 Jan3;3(1):e1918911.doi:10.1001/ jamanetworkopen.2019.18911.

Statistisches Bundesamt, Genesis-Online Datenbank, https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/Publikationen/Downloads-Krankenhaeuser/operationen-prozeduren-5231401187015.html; Datenabfrage vom 16.04.2020 (Zahlen zu den häufigsten operativen Eingriffen in der Weichteilchirurgie)

Kontakt:

Almut Kalz
Koordinatorin Kurt-Semm-Zentrum und Kurt-Semm-Akademie am Universitätsklinikum Schleswig-Holstein, Campus Kiel
almut.kalz@uksh.de
www.kurt-semm-zentrum.de

Prof. Dr. K.-P. Jünemann
Sprecher des Kurt-Semm-Zentrums für laparoskopische und roboter-assistierte Chirurgie
Direktor der Klinik für Urologie und Kinderurologie am UKSH, Campus Kiel

Bundesministerium
für Bildung
und Forschung

Meldungen aus dem BMBF

Gefährliche Doppelrolle des Immunsystems

Die körpereigene Abwehr ist die stärkste Waffe gegen Viren, aber wenn sie sich gegen gesunde Zellen wendet, besteht große Gefahr. Ein Forschungsteam aus Berlin untersucht mithilfe aufwendiger bioinformatischer Analysen die Rolle des Immunsystems. Für einen Großteil der mit dem neuen Corona-Virus Infizierten verläuft die Erkrankung nicht viel schlimmer als eine herkömmliche Erkältung. Sie haben milde Symptome und nach zwei Wochen sind die meisten wieder gesund. Bei einem kleineren Anteil von rund zehn bis 15 Prozent ist der Krankheitsverlauf jedoch dramatisch und es kann zu lebensbedrohlichen Komplikationen kommen. Was diese beiden Patientengruppen voneinander unterscheidet und welche Rolle das körpereigene Immunsystem dabei spielt, will das Forschungsteam um den Mathematiker Professor Roland Eils vom Berlin Institute of Health herausfinden. Dabei unterstützt sie das Bundesforschungsministerium im Rahmen des Deutschen Netzwerks für Bioinformatik-Infrastruktur de.NBI.

„Wir haben frühzeitig vermutet, dass das Immunsystem der Infizierten über den Verlauf der Krankheit entscheidet“, erklärt Eils. Um diesen Verdacht genauer zu untersuchen, haben die Forscherinnen und Forscher auf sogenannte Einzelzellanalysen gesetzt. Denn nicht nur jeder Mensch, sondern auch jede Zelle reagiert anders auf eine Infektion. Die Zellreaktion untersuchen die Wissenschaftlerinnen und Wissenschaftler, indem sie sich die Regulation der einzelnen Gene ansehen. Mithilfe der Genregulation reagiert die Zelle auf äußere Einflüsse. So kann die Zelle etwa einzelne Gene verstärkt ablesen und dadurch mehr Rezeptoren für die Interaktion mit anderen Zellen bilden.

Die Analyse einzelner Zellen ist extrem aufwendig. Aus dem Nasen-Rachen-Abstrich der Testpersonen müssen die Forscherinnen und Forscher zunächst tausende Zellen isolieren. Für jede dieser Zellen wird dann die Aktivität tausender Gene bestimmt. Dabei kommen unvorstellbar große Datenmengen zusammen, die die Forscherinnen und Forscher nur noch mit sehr leistungsfähigen Computern und speziell entwickelten bioinformatischen Werkzeugen analysieren können.

Unheilvolle Allianz mit dem Virus

Mittlerweile ist bekannt, dass das neuartige Corona-Virus den sogenannten ACE-2-Rezeptor als Eingangspforte benutzt, um in die Zellen der Nasen- und Rachenschleimhaut einzudringen und diese zu infizieren. „Unsere Untersuchungen haben gezeigt, dass dieser Rezeptor nur in einem Bruchteil der Zellen überhaupt vorhanden ist“, erklärt Eils. Die Viren können also zunächst nur wenige Zellen infizieren. Das Forschungsteam wollte verstehen, warum das neuartige Coronavirus dennoch so effizient dabei ist, eine Erkrankung auszulösen. Sie haben sich daher auch den weiteren Krankheitsverlauf angesehen.

Die infizierten Zellen rufen nach dem Viruseintritt die körpereigne Abwehr zu Hilfe. Die Immunzellen schütten daraufhin einen bestimmten Faktor aus – das sogenannte Interferon Gamma. Neben einer verstärkten Immunreaktion führt dieser Faktor allerdings auch dazu, dass die Gewebezellen den ACE-2-Rezeptor vermehrt bilden. „Das Immunsystem geht hier eine unheilvolle Allianz mit dem Virus ein“, sagt Eils. Im Vergleich mit gesunden Testpersonen ist der ACE-2-Rezeptor bei Erkrankten zwei bis dreimal so oft vorhanden. Dadurch gibt es viel mehr Eingangspforten

Die infizierten Zellen rufen nach dem Viruseintritt die körpereigne Abwehr zu Hilfe.

Quelle: Adobe Stock / jijomathai

für das Virus in den Körper und im weiteren Krankheitsverlauf können die Viren so deutlich mehr Zellen infizieren.

Zerstörung gesunder Lungenzellen

Das ist allerdings nicht die einzige gefährliche Rolle des Immunsystems. Bei schwer erkrankten Patientinnen und Patienten konnte das Forschungsteam beobachten, dass es bei diesen zusätzlich zu einer überschießenden Entzündungsreaktion kommt. „Die Immunzellen zerstören dann nicht mehr nur infizierte Zellen, sondern auch gesunde Lungenzellen," sagt Eils. „Patientinnen und Patienten, bei denen es zu so einer überschießenden Reaktion des Immunsystems kommt, haben massive Atemwegsprobleme und müssen in der Regel künstlich beatmet werden". Warum das Immunsystem bei einigen Menschen so heftig reagiert und die Krankheit – anstatt sie zu bekämpfen – am Ende sogar noch schlimmer macht, können die Forscherinnen und Forscher bislang noch nicht sagen. „Um das zu verstehen, brauchen wir deutlich mehr Daten", so Eils.

Das bessere Verständnis der überschießenden Immunantwort bei schweren Krankheitsverläufen bietet jedoch schon jetzt einige interessante Behandlungsansätze. „Noch sind es Hypothesen, aber es wäre naheliegend, die Interaktion zwischen den Gewebe- und Immunzellen gezielt zu stören", erklärt Eils. Die Medikamente, die das Forschungsteam dafür vorschlägt, sind teilweise schon im Zulassungsprozess oder bereits für andere virale Erkrankungen zugelassen. Die Wissenschaftlerinnen und Wissenschaftler haben ihre Forschungsergebnisse bewusst so veröffentlicht, dass die Daten weltweit und kostenfrei zur Verfügung stehen. Das Forschungsteam hofft, dass klinische Partner auch aus anderen Ländern ihre Ergebnisse aufgreifen werden und sie zu einer effizienten Behandlung der schweren Krankheitsverläufe beitragen können.

Weitere Informationen unter:

www.bmbf.de/de/gefaehrliche-doppelrolle-des-immunsystems-12604.html

Eine intelligente Datennutzung scheitert noch oft an uneinheitlichen Datenformaten.

Quelle: Adobe Stock / ipopba

Daten helfen heilen

Drei Bundesministerien – ein Ziel: Mit der im September 2020 veröffentlichten Roadmap zur Innovationsinitiative „Daten für Gesundheit" stellen das BMBF, das BMG und das BMWi gemeinsam die Weichen für die digitale Medizin der Zukunft.

Bei jeder Untersuchung und bei jeder Behandlung fallen medizinische Daten an. Sie sind für die Forschung von enormer Bedeutung, denn in ihnen liegt der Schlüssel für neue Erkenntnisse über die Entstehung von Erkrankungen. Zugleich liefern sie wichtige Ansatzpunkte für eine bessere Gesundheitsversorgung – für schnellere und präzisere Diagnosen und für individuell auf die Patientinnen und Patienten abgestimmte Behandlungsmöglichkeiten.

Voraussetzungen für die intelligente Datennutzung schaffen

Eine intelligente Datennutzung scheitert heute oft an uneinheitlichen Datenformaten und -standards. So liegen Arztbriefe häufig als Freitexte vor und Laborwerte als Tabellen mit unterschiedlichen Einheiten. Selbst simple Blutdruckwerte können innerhalb einer Klinik unterschiedlich dokumentiert werden. Solche Daten kann auch die intelligenteste Software nicht sinnvoll auswerten. Daher gilt es, die digital erfassten Gesundheitsdaten künftig nach denselben Regeln und unter Einhaltung internationaler Standards zu dokumentieren.

Diese Harmonisierung und die Vernetzung von Daten voranzutreiben sind zentrale Ziele der Innovationsinitiative „Daten für Gesundheit: Roadmap für eine bessere Patientenversorgung durch Gesundheitsforschung und Digitalisierung“, die das Bundesministerium für Bildung und Forschung, das Bundesministerium für Gesundheit und das Bundesministerium für Wirtschaft und Energie ins Leben gerufen haben. Mit einer gemeinsamen Strategie werden die drei Ministerien mit den Akteuren aus der Gesundheitsversorgung, -forschung und -wirtschaft dieses zukunftweisende Thema voranbringen.

Die fünf prioritären Handelsfelder im Überblick

1. Strukturen für die digitale Vernetzung von Gesundheitsversorgung und Gesundheitsforschung auf- und ausbauen,
2. die Verfügbarkeit und Qualität von gesundheitsrelevanten Daten verbessern,
3. die Entwicklung innovativer Lösungen zur Verbesserung von Datensicherheit und Datenverknüpfung vorantreiben,
4. den Weg in die datenunterstützte Medizin gemeinsam gehen,
5. zukünftige Anwendungsperspektiven frühzeitig erschließen.

Eine weitere zentrale Voraussetzung für den Erfolg der Initiative ist die Bereitschaft der Menschen, Daten für die Forschung bereitzustellen. Diese Bereitschaft hängt von dem Vertrauen der Menschen in jene Personen und Prozesse ab, die ihre Daten erheben, verarbeiten und auswerten. Ebenso wichtig ist die Frage der Datenhoheit: Wer entscheidet eigentlich, welche Daten wofür genutzt werden können? Die Antwort ist klar: Die Datenhoheit haben die Bürgerinnen und Bürger.

Sie bestimmen selbst, ob und in welcher Form ihre Daten für heutige und zukünftige Forschungsfragen verwendet werden dürfen. Dafür wurde bereits ein einheitlicher Mustertext für die Patienteneinwilligung entwickelt, auf den sich alle Universitätskliniken in Deutschland zusammen mit den Datenschutzaufsichtsbehörden des Bundes und der Länder verständigt haben. Um die Akzeptanz und das Vertrauen der Patientinnen und Patienten zu gewinnen, werden zudem umfangreiche Informationsmaterialien – auch zu den Aspekten des Datenschutzes und der Datensicherheit – erstellt.

Deutschland als Standort für Gesundheit stärken

In ihrer Hightech-Strategie 2025 hat die Bundesregierung die bessere Vernetzung von Forschung und Versorgung zu einer ihrer zwölf übergreifenden Missionen erklärt. Ziel ist, dass Deutschland weltweit eine führende Position bei der Entwicklung und Anwendung digitaler Gesundheitsinnovationen einnehmen soll. Die Initiative „Daten für Gesundheit“ der drei beteiligten Bundesministerien ist ein wichtiger Bestandteil dieser Mission: Sie soll dazu beitragen, frühzeitig neue Technologien und Anwendungsbereiche zu erschließen.

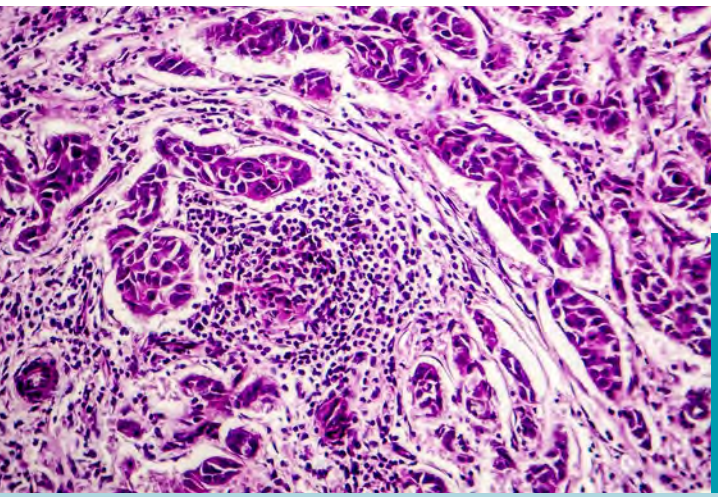
Weitere Informationen unter:

www.bmbf.de/de/daten-helfen-heilen-12503.html

Künstliche Intelligenz bringt Ordnung ins molekulare Chaos

In jedem Tumor stecken Informationen, die für die Wahl der richtigen Therapie entscheidend sind. Bisher ist es jedoch schwierig, an sie heranzukommen. Ein neues KI-Tool soll das Problem lösen und der Krebsforschung damit zehn Jahre Zeit sparen.

„Tumor-Matsche“, so nennen die Forscherinnen und Forscher das Material, das ihnen so wichtige Informationen liefert. „Matsche“ ist in diesem Sinne nicht negativ gemeint, sondern steht bildlich dafür, dass diese Informationen in einer bislang schwer überschaubaren Masse vermengt sind. Der Bioinformatiker Professor



In jedem Tumor stecken Informationen, die für die Wahl der richtigen Therapie entscheidend sind.

Quelle: Adobe Stock/Kateryna_Kon

Dr. Rainer Spang von der Universität Regensburg und sein Forschungsteam haben das Ziel, diese Informationen mithilfe von Künstlicher Intelligenz zu sortieren und auszuwerten. Wenn ihnen dies gelingt, würde das für die Krebsforschung einen enormen Zeitgewinn bedeuten. Das Bundesministerium für Bildung und Forschung unterstützt sie dabei im Rahmen der Fördermaßnahme „CompLS – Computational Life Sciences“.

Seit 20 Jahren messen Krebsforscherinnen und -forscher, welche Gene im jeweiligen Tumor in welcher Intensität aktiv sind. Dies hilft dabei, die richtige Therapieentscheidung zu treffen. Da jeder Tumor einzigartig ist, spricht er auch nur auf bestimmte Medikamente an. Um an die Moleküle der Zellen im Tumor zu gelangen, müssen diese aufgebrochen werden. Dadurch entsteht ein Molekül-Gemisch, das die Regensburger Forscher „Tumor-Matsche“ genannt haben. „Da die Tumore jedoch nicht nur aus Tumorzellen bestehen, werden diese beim Aufschließen mit Teilen von Bindegewebs- und Immunzellen vermischt“, erklärt Spang. „Das macht es schwierig, die gewonnenen molekularen Informationen den richtigen Zellen zuzuordnen.“

KI-Tool soll Therapie-Vorhersagen deutlich verbessern

Wenn ein Gewebe etwa die Information enthält, dass Gene für Zellteilung hochreguliert sind, stellt sich die Frage, woher dieses Signal kommt. Stammt es aus den Tumorzellen, so wächst der Tumor. Kommt das Signal aber aus den Immunzellen, so bekämpft das Immunsystem den Tumor. Für die behandelnden Ärztinnen und Ärzte ist dies ein gewaltiger Unterschied: Je nach Szenario müssten sie völlig andere Therapiewege gehen. Hier setzt das Forschungsteam um Spang an. Die Wissenschaftlerinnen und Wissenschaftler entwickeln ein KI-Tool, das die Messdaten den jeweiligen Zellen zuordnet und damit die Therapie-Vorhersagen deutlich verbessern könnte.

Seit kurzer Zeit gibt es zwar auch Verfahren, die Messungen für einzelne Tumorzellen möglich machen. Hier fehlen jedoch die Erfahrungswerte. „Von den so untersuchten Fällen weiß man noch nicht, wie sich die Krankheiten entwickeln werden“, sagt Spang. Das bedeutet, aus einzelnen Messdaten kann noch keine Therapieempfehlung abgeleitet werden. Zur Analyse der „Tumor-Matsche“ werden dagegen seit Jahrzehnten Vergleichswerte gesammelt, die Rückschlüsse für bestimmte Patientengruppen zulassen. „Man braucht Tausende von Patientinnen und Patienten, um verlässliche Zahlen zu bekommen“, erklärt Spang. „Und man muss abwarten, wie sich die Krankheit im Laufe der Zeit entwickelt.“ Bis die Einzelzell-Datensätze so weit sind, werden noch bis zu 15 Jahre vergehen, schätzt Spang.

„Mindestens zehn Jahre Zeit gewinnen“

Deshalb wollen die Regensburger Bioinformatiker es mit ihrem Tool schaffen, Informationen zu einzelnen Zellen aus den Messdaten der „Tumor-Matsche“ zu erhalten. Denn die entscheidenden Signale sind auch im unsortierten Gewebe enthalten, sie sind lediglich überdeckt. Sie aus dem molekularen Chaos herauszufiltern, ist die Aufgabe der KI. „Wenn wir das schaffen, dann können wir zum Wohle der Patientinnen und Patienten mindestens zehn Jahre Zeit gewinnen, bis die neuen Methoden soweit sind.“

Die neuen aufgeschlüsselten Daten der Regensburger Forscherinnen und Forscher könnten darüber hinaus Anknüpfungspunkte für neue Therapien und Medikamente liefern. Als Beispiel nennt Spang die personalisierten Immuntherapien, die im Kampf gegen Krebs immer häufiger zum Einsatz kommen. In der Praxis könnte das so aussehen: Das KI-Tool hat festgestellt, dass bestimmte Immunzellen den Tumor zwar erkannt haben, aber nicht angreifen. Ursache hierfür ist eine Blockade. Die Ärztin oder der Arzt könnte dann eine Immuntherapie verordnen. Diese würde die Blockade auflösen und die Immunabwehr erfolgreich in Gang setzen.

Weitere Informationen finden Sie unter:

www.bmbf.de/de/kuenstliche-intelligenz-bringt-ordnung-ins-molekulare-chaos-9796.html

BMBF-Newsletter:

Das Wichtigste der letzten Wochen aus dem BMBF im Überblick (erscheint monatlich).
www.bmbf.de/newsletter

Das BMBF hält Sie auch über Twitter, Facebook und Instagram auf dem Laufenden:

twitter.com/BMBF_Bund

www.facebook.com/bmbf.de

www.instagram.com/bmbf.bund

bioinformatik: junge frauen brauchen vorbilder

Interview mit Bioinformatikerin Janine Felden
PANGAEA Gruppenleiterin am Alfred-Wegener Institut Bremerhaven

Im deutschen Netzwerk für Bioinformatik sind Gruppenleiterinnen in der Minderheit. Janine Felden ist eine von ihnen. Im Interview mit gesundhyte.de spricht sie über ihren Weg als Quereinsteigerin in die Bioinformatik und was sich ändern sollte, damit mehr Frauen in der Bioinformatik arbeiten. Ein Gebiet, das sie bis heute fasziniert und zum Staunen bringt.

gesundhyte.de: *Im Gegensatz zur Biologie sind Frauen in der Bioinformatik bislang in der Minderheit. Was schreckt die Frauen ab?*

Dr. Janine Felden: Bei Bioinformatik denken manche Leute leider noch immer an Menschen, die allein im abgedunkelten Kellerraum vor ihren Rechnern sitzen. Ich denke, dieses Bild schreckt gerade junge Frauen ab, die gerne sozial interagieren möchten. Bei einem grundsätzlichen Interesse für Biologie entscheiden sie sich dann doch lieber erstmal für die eher anwendungsorientierte Arbeit im Labor. Dabei hat Bioinformatik auch ganz viel mit Teamarbeit zu tun. Man ist sogar angewiesen auf einen regen Austausch mit vielen engagierten und kommunikativen Leuten. Bioinformatikerinnen und Bioinformatiker sind alles andere als unsoziale Einsiedler. Im Verlauf eines Biologie-Studiums kann zum Glück auch immer noch das Interesse für Bioinformatik geweckt werden. Manchmal ist es eher die Frage, wie man den Zugang findet – es gibt viele Quereinsteigerinnen wie mich.

Wie sind Sie zur Bioinformatik gekommen?

Über viele Umwege. Ich habe zunächst Biologie studiert, weil mich Meeresforschung fasziniert hat und dort vor allem das mikrobielle Leben. Mikroorganismen steuern letztendlich das gesamte Leben auf unserem Planeten. Die Covid-19-Pandemie hat uns das nochmal eindrücklich gezeigt. Mikroorganismen kann man im Gegensatz zu Pflanzen und Tieren nicht allein an ihren äußeren Merkmalen unterscheiden. Erst über eine Analyse ihrer DNA mit Werkzeugen der Bioinformatik lassen sich die Mikroorganismen und ihre Eigenschaften eindeutig identifizieren. So bin ich erstmalig mit Bioinformatik in Berührung gekommen. Ich wollte verstehen: Wer ist da draußen? Und was machen die? Dafür brauchen wir die Bioinformatik. Sie stellt entsprechende Werkzeuge zur Verfügung und bildet die Grundlage für zahlreiche Labormethoden.

> „Faszinierend ist für mich vor allem die Bandbreite an Fragen, die man mit bioinformatischen Werkzeugen beantworten kann."

Wie sieht Ihr Arbeitsalltag aus?

Im Gegensatz zum Klischee umfasst mein Arbeitstag viel Interaktion mit realen Menschen. In meinem Umfeld sind es eigentlich auch meistens gemischte Arbeitsgruppen. Männer und Frauen gehen durchaus unterschiedlich an Dinge heran, aber am Ende sind die Ergebnisse qualitativ gleichwertig. Für die Zusammenarbeit sind unterschiedliche Sichtweisen sogar förderlich. Meine Arbeitsgruppe kümmert sich im Rahmen der Datenbank PANGAEA hauptsächlich um das Veröffentlichen und Bereitstellen von wissenschaftlichen Daten. So können sie von allen gefunden, genutzt und wiederverwendet werden. Unsere tägliche Arbeit umfasst jedoch nicht nur das Organisieren und Integrieren von Daten. Wir befinden uns als Service-Plattform des deutschen Netzwerks für Bioinformatik-Infrastruktur ständig im Austausch mit unseren Nutzerinnen und Nutzern und begleiten die wissenschaftlichen Arbeiten von der Idee bis zur Veröffentlichung. Dafür unterstützen wir die Forscherinnen und Forscher bei ihrem Datenmanagement und ermöglichen eine Verknüpfung von DNA- und Umweltdaten.

Dr. Janine Felden (Foto: Privat / Felden).

Was fasziniert Sie an der Bioinformatik?

Faszinierend ist für mich vor allem die Bandbreite an Fragen, die man mit bioinformatischen Werkzeugen beantworten kann. Die Bioinformatik hat uns dabei erstmal gezeigt, wie vielfältig unsere Umwelt ist. Mit nur einer einzigen Probe aus dem Meer können wir viele Fragen auf einmal beantworten z. B. Wer ist da? Was machen die? Es hat mich immer fasziniert, die größeren Zusammenhänge in und zwischen verschiedenen Ökosystemen zu verstehen. Aber auch in der Medizin gibt es spannende Themen. Krebsmedikamente werden heute zum Teil an der individuellen DNA-Sequenz der Erkrankten ausgerichtet und erhöhen so die Überlebenschance der Patientinnen und Patienten. Das wäre ohne Bioinformatik nicht möglich. Das finde ich unglaublich faszinierend. Außerdem ist es schön zu sehen, dass die Bioinformatik mittlerweile ein Impulsgeber für andere Wissenschaften ist. Die Verarbeitung von großen Datenmengen – auch „Big Data" genannt – ist ein gutes Beispiel dafür. Neben der Klimaforschung, ist die Bioinformatik führend auf diesem Gebiet und kann jetzt als Impulsgeber für andere Disziplinen fungieren.

Wie bewerten Sie die Zukunftsaussichten der Bioinformatik?

In unserer Informationsgesellschaft ist Bioinformatik eine Ausbildung, die auch langfristig eine gute Perspektive bietet. Das gilt nicht nur für die akademische Welt, auch in der freien Wirtschaft gibt es zahlreiche und spannende Aufgabenfelder. Man ist nicht zu spezialisiert und kann erlernte Abläufe auch auf andere Wissensgebiete anwenden. Um die exponentiell ansteigenden Mengen an Daten zu bewältigen und nutzbar zu machen, brauchen wir gut ausgebildeten Nachwuchs – unabhängig vom Geschlecht.

Was muss sich ändern, damit mehr Frauen in der Bioinformatik arbeiten?

Ein paar Vorurteile müssen wohl noch abgebaut werden. Ich denke, hier sind Vorbilder ganz wichtig. Junge Frauen sollen

„Um die exponentiell ansteigenden Mengen an Daten zu bewältigen und nutzbar zu machen, brauchen wir gut ausgebildeten Nachwuchs – unabhängig vom Geschlecht."

sehen, dass es durchaus Frauen in diesem Bereich gibt und diese auch sehr erfolgreich sein können. Ich denke, es ist wichtig, dass die Studienfächer durchlässig sind. Frauen die zunächst im Labor arbeiten, sollten die Möglichkeit haben, ihren Schwerpunkt auf die Bioinformatik zu verlegen, wenn sie merken, dass nicht nur die Arbeit im Labor spannende Fragestellungen und Antworten bereithält. Die Vereinbarkeit von Beruf und Familie ist in der digitalen, bioinformatischen Welt sogar einfacher, da man sich die Zeit flexibler einteilen kann als im Labor.

Manche Universitäten bieten schon spezielle Frauenstudiengänge in Informatik an, um mehr Frauen zum Studium zu bewegen. Was halten Sie davon?

An sich bin ich ein Fan von einem gut ausgeglichenen Verhältnis zwischen Männern und Frauen. Ein rein weibliches Umfeld

(Foto: iStock © metamorworks)

entspricht einfach nicht der Realität. Warum sollte es dann im Studium so sein? Ich sehe allerdings auch, dass es wichtig ist, die Attraktivität der Bioinformatik für Frauen zu erhöhen und mögliche Hemmnisse zu senken. Die Interessen sind bei heranwachsenden Mädchen häufig anders gelagert als bei Jungen. Für den Einstieg wäre es daher sicherlich gut, wenn Angebote auf die jeweiligen Vorkenntnisse eingehen. Einstiegskurse für Frauen könnten also durchaus sinnvoll sein, allerdings sollten sie nicht zu einem Stigma werden. Und warum sollten nicht auch junge Männer ohne Vorwissen davon profitieren? Obwohl die Jugendlichen heute eine höhere Technikaffinität besitzen als noch vor 20 Jahren, verbringen nicht alle ihre Teenager-Jahre damit, kleine Roboter zu programmieren. Hier kann ein niederschwelliger Einstieg helfen, Interesse zu wecken. Vielleicht hätte ich es dann auch studiert?

Das Gespräch führten Melanie Bergs und Gesa Terstiege.

Kontakt:

Dr. Janine Felden
Alfred-Wegener-Institut, Bremerhaven
janine.felden@awi.de

www.marum.de/Dr.-janine-felden.html

Datenbank PANGAEA

PANGAEA – Data Publisher for Earth & Environmental Science – ist ein digitales Bibliothekssystem für Daten aus der Erdsystemforschung und den Umweltwissenschaften, das vom Zentrum für Marine Umweltwissenschaften MARUM an der Universität Bremen und dem Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung gemeinschaftlich betrieben wird. Das professionelle Datenmanagement ist als Service-Plattform ein Teil des deutschen Netzwerks für Bioinformatik-Infrastruktur (de.NBI), das das BMBF noch bis Ende 2021 mit 81 Millionen unterstützt.

ein motor für die analyse von COVID-19-forschungsdaten

Das Deutsche Netzwerk für Bioinformatik-Infrastruktur überzeugt durch Auswerteprogramme und eine Cloud-basierte Rechnerstruktur

von Alfred Pühler, Irena Maus, Vera Ortseifen und Andreas Tauch

Die Coronavirus-Pandemie hält die Welt seit Monaten in Atem. Forscher aus den Lebenswissenschaften versuchen, Informationen über das Virus SARS-CoV-2 zu sammeln, die COVID-19-Erkrankung aufzuklären, aber auch Impfstoffe sowie Medikamente zur Bekämpfung der Epidemie zu entwickeln. In umfangreichen Versuchsreihen werden Forschungsdaten erhoben, die jedoch erst mit geeigneten Computerprogrammen analysiert werden müssen. Hier ist das Deutsche Netzwerk für Bioinformatik-Infrastruktur (de.NBI) gefragt. Es wird das Potenzial dieses Netzwerks bei der Bekämpfung von COVID-19 aufgezeigt sowie die Beteiligung der de.NBI-Mitglieder an COVID-19 Projekten in deutschen und europäischen Initiativen vorgestellt.

Das de.NBI-Netzwerk zur Analyse lebenswissenschaftlicher Forschungsdaten

Das Deutsche Netzwerk für Bioinformatik-Infrastruktur (www.denbi.de) wurde 2015 vom Bundesministerium für Bildung und Forschung (BMBF) etabliert, um Forschenden in den Lebenswissenschaften eine Infrastruktur zur Analyse von umfangreichen Datenmengen zur Verfügung zu stellen. Das Netzwerk zählt zurzeit mehr als 300 wissenschaftliche Mitglieder und besteht aus 40 Projekten, die in acht Servicezentren angesiedelt sind. Die acht Servicezentren sind thematisch ausgerichtet und decken verschiedene Teildisziplinen der Lebenswissenschaften ab, u. a. auch alle Aspekte, die für eine datenbasierte Medizin benötigt werden.

Die im de.NBI-Netzwerk aufgebaute Infrastruktur umfasst die Bereiche Service, Training und Compute (Abbildung 1).

Im Servicebereich steht eine Vielzahl von Analyseprogrammen zur Auswertung von lebenswissenschaftlichen Daten zur Verfügung. Eng verbunden mit dem Servicebereich ist der Trainingsbereich, der de.NBI-Nutzende im Umgang mit Analyseprogrammen sowie mit erzielten wissenschaftlichen Ergebnissen schult. Im Servicebereich können mehr als 100 Analyseprogramme und Workflows genutzt werden. Auf dem Trainingssektor wurden bisher rund 70 Trainingskurse pro Jahr angeboten, womit das de.NBI-Netzwerk bis heute rund 6.000 Nutzende schulen konnte.

Auch auf dem Compute-Sektor spielt das de.NBI-Netzwerk durch die Etablierung einer eigenen de.NBI-Cloud eine bemerkenswerte Rolle. Diese de.NBI-Cloud, die an sechs Standorten in Berlin, Bielefeld, Freiburg, Gießen, Heidelberg und Tübingen angesiedelt ist, hat durch eine umfangreiche BMBF-Förderung eine beachtenswerte Größe erreicht und erlaubt die Analyse von großen in den Lebenswissenschaften auftretenden Datenmengen. Allen Forschenden aus den Lebenswissenschaften wird die Nutzung der de.NBI-Cloud gebührenfrei zur Verfügung gestellt (siehe https://www.denbi.de/cloud). Erfahren Sie mehr über die de.NBI Cloud im Folgeartikel auf Seite 80.

Zusammenfassend stellt das de.NBI-Netzwerk eine wichtige Infrastruktur auf nationaler Ebene zur Verfügung und ist darüber hinaus über den deutschen Knoten ELIXIR Germany eng mit ELIXIR Europe verbunden, einer europäischen Organisation zur Entwicklung einer länderübergreifenden Bioinformatik-Infrastruktur.

Abbildung 1: Die Hauptaktivitäten des de.NBI-Netzwerks umfassen Service und Training sowie das Angebot einer Cloud-basierten Rechner-Einrichtung. Zusätzlich ist das de.NBI-Netzwerk mit seinen Aktivitäten in den europäischen Verbund ELIXIR eingebunden (Quelle: de.NBI-Geschäftsstelle).

Die Analyse von COVID-19-Forschungsdaten – Nagelprobe für das de.NBI-Netzwerk

Die aktuelle Corona-Pandemie stellt eine große Herausforderung für unsere Gesellschaft dar und erfordert deshalb besondere Aufmerksamkeit durch die etablierten Wissenschaftsstrukturen. Es gilt, die molekularbiologische Forschung an Coronaviren, insbesondere am SARS-CoV-2, voranzutreiben, epidemiologische Aspekte des Infektionsgeschehens aufzuklären sowie den Verlauf von COVID-19-Erkrankungen zu erforschen. Zusätzlich müssen Anstrengungen unternommen werden, um Medikamente zur Behandlung von Erkrankten zu entwickeln und geeignete Impfstoffe zu produzieren. All diese Forschungsgebiete generieren große Datenmengen, die ihren Wert erst nach einer sorgfältigen bioinformatischen Auswertung offenbaren. Die Auswertung großer Datenmengen ist einer der Schwerpunkte von de.NBI und so erweist sich de.NBI in der Corona-Pandemie als ein wahrer Glücksfall. Die vorhandenen de.NBI-Ressourcen konnten effektiv und zeitnah für die Analyse von COVID-19-Forschungsdaten eingesetzt werden. Damit hat de.NBI die Nagelprobe bestanden, kurzfristig auf den aktuellen Bedarf an Datenanalyse reagieren zu können. Insgesamt wurden 29 COVID-19-Forschungsprojekte mit de.NBI-Beteiligung ermittelt und auf der de.NBI-Webseite veröffentlicht (https://www.denbi.de/covid-19).

Beiträge des de.NBI-Netzwerks zu COVID-19-Forschungsprojekten

Die im de.NBI-Netzwerk durchgeführten Aktivitäten zur Unterstützung von COVID-19-Forschungsprojekten lassen sich in sechs Forschungsrubriken (Abbildung 2) einordnen. Die 29 Forschungsprojekte mit COVID-19-Bezug werden von einer Vielzahl von Forschungseinrichtungen des de.NBI-Netzwerks durchgeführt (Tabelle 1).

In der Rubrik **„Sequenzvarianten und Ausbreitung von SARS-CoV-2“** liegen drei Projekte vor, die sich mit der Sequenzerstellung und -interpretation von Virusgenomen beschäftigen. Am EMBL in Heidelberg wird der Genomanalyse-Server GEAR eingesetzt, während Forscher der Charité und der Universität Heidelberg das Onlinetool MapMyCorona entwickelt haben. In Bielefeld werden mithilfe der Nanopore-Sequenziermethode SARS-CoV-2-Genome entschlüsselt und mittels der de.NBI-Cloud analysiert.

Die Rubrik **„Wechselwirkung von SARS-CoV-2 mit menschlichen Zellen“** beinhaltet Projekte, die die Interaktion des Virus mit Zellen aus menschlichem Gewebe oder dem Immunsystem zum Inhalt haben. Als bevorzugte Methode wird die Einzelzellanalyse von menschlichen Zellen eingesetzt. In dieser Rubrik ist ein großes gemeinsames Projekt der beiden Standorte Heidelberg und Tübingen angesiedelt, welches darauf

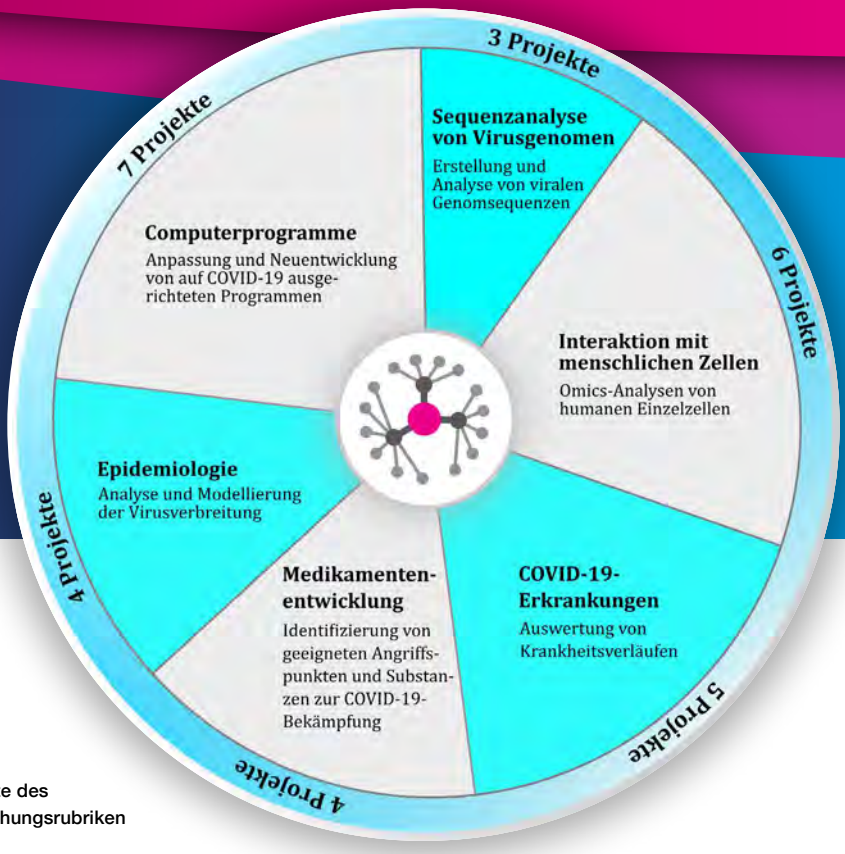


Abbildung 2: Die Einordnung der 29 Forschungsprojekte des de.NBI-Netzwerks mit COVID-19-Bezug in sechs Forschungsrubriken (Quelle: de.NBI-Geschäftsstelle).

abzielt, die Zusammenhänge zwischen SARS-CoV-2-Genomik und unterschiedlichen Infektionsverläufen unter Nutzung der de.NBI-Cloud zu analysieren. Aus Gießen sind drei Projekte zu nennen, die sich mit Lungenerkrankungen, speziell mit SARS-CoV-2-Infektionen, beschäftigen. Der Omics-Explorer Magellan, entwickelt an der Charité, bietet Tools zur Visualisierung von aktuellen COVID-19 Forschungsdaten an und unterstützt damit weitergehende Analysen. Ein weiteres Projekt aus Berlin beschäftigt sich mit der Zellantwort von infiziertem Gewebe und analysiert die gewonnenen Omics-Daten mit verschiedenen de.NBI-Programmen. Schließlich wird aus Berlin noch ein Projekt eingebracht, das nach Infektion von *In-vitro*-Zellgewebe mit SARS-CoV-2 das Transkriptionsgeschehen in Einzelzellen analysiert.

In der Rubrik **„Analyse der COVID-19-Erkrankung“** wurden drei Projekte aus Bochum aufgenommen, die Informationen zum COVID-19-Krankheitsverlauf sammeln und mit speziellen Informatik-Methoden, unter anderem auch auf dem Gebiet der künstlichen Intelligenz, analysieren. Im Weiteren ist in dieser Rubrik auch ein Projekt aus Rostock eingebunden, das die Erstellung von disease maps vorsieht und mittels Textmining die Auswahl von Medikamenten erleichtern soll. In Kiel wird schließlich mit genomweiten Assoziierungsstudien nach Auffälligkeiten in Genomen von COVID-19-Erkrankten gesucht.

Besondere Bedeutung kommt der Rubrik **„Entwicklung von Medikamenten zur Bekämpfung von SARS-CoV-2-Infektionen“** zu. Am DKFZ in Heidelberg wird nach Substanzen mit antiviralen Eigenschaften gesucht, indem geprüft wird, ob diese die Transkription des Virusgenoms hemmen. An der Universität Heidelberg wird mittels kleiner, interferierender RNA-Fragmente (siRNA) getestet, ob diese die Virusvermehrung unterbinden können. Beide Gruppen setzen für ihre Untersuchungen Hochdurchsatz-Mikroskopieverfahren ein. An der Universität Freiburg wird mittels *In-silico*-Methoden die Suche nach möglichen Medikamenten betrieben. In Hamburg wird ebenfalls mittels *In-silico*-Methoden nach Substanzen gesucht, die an SARS-CoV-2-Proteine binden können.

Die Rubrik **„Unterstützung klassischer epidemiologischer Studien“** spielt beim Verständnis der Virusausbreitung eine bedeutende Rolle. In dieser Rubrik werden SARS-CoV-2-Ausbreitungsverläufe modelliert. Das de.NBI-Netzwerk ist an dieser Forschungskategorie mit Projekten aus Jena, Rostock und Heidelberg beteiligt. Das Projekt aus Jena hat zum Ziel, eine Modellierung und Vorhersage des Epidemieverlaufs durchzuführen. Auch das Projekt aus Rostock nutzt mathematische Modellierung, um den Verlauf der Epidemie vorherzusagen. Der Ansatz aus Heidelberg nutzt die Software COPASI zur Modellierung von SARS-CoV-2-Infektionen.

Tabelle 1: Die 29 COVID-19-Forschungsprojekte des de.NBI-Netzwerks

Rubrik	Anzahl der Projekte	Forschungseinrichtung
Sequenzvarianten und Ausbreitung von SARS-CoV-2	3	• The European Molecular Biology Laborarory (EMBL), Heidelberg • Charité, Berlin und Universität Heidelberg • Universität Bielefeld
Wechselwirkung von SARS-CoV-2 mit menschlichen Zellen	6	• Justus-Liebig-Universität Gießen • Charité, Berlin und Universität Heidelberg • Universität Heidelberg • Eberhard Karls Universität Tübingen • Max-Delbrück-Centrum für Molekulare Medizin, Berlin
Analyse der COVID-19-Erkrankung	5	• Ruhr-Universität Bochum • Universität Rostock • Christian-Albrechts-Universität zu Kiel
Entwicklung von Medikamenten zur Bekämpfung von SARS-CoV-2-Infektionen	4	• Deutsches Krebsforschungszentrum (DKFZ), Heidelberg • Universität Heidelberg • Albert-Ludwigs-Universität Freiburg • Universität Hamburg
Unterstützung klassischer epidemiologischer Studien	4	• Fritz-Lipmann-Institut, Jena • Universität Rostock • Universität Heidelberg
Entwicklung von Werkzeugen zur Analyse von COVID-19-relevanten Daten	7	• Ruhr-Universität Bochum • Albert-Ludwigs-Universität Freiburg • Heidelberger Institut für Theoretische Studien • Universität Bielefeld • Eberhard Karls Universität Tübingen

In der letzten Rubrik **„Entwicklung von Werkzeugen zur Analyse von COVID-19-relevanten Daten“** sind Projekte angesiedelt, die sich insgesamt der Entwicklung von Computerprogrammen zur Analyse von COVID-19-bezogenen Daten widmen. Zu diesen Werkzeugen gehören auch Programme, die den Betrieb der de.NBI-Cloud und den Einsatz der Galaxy-Plattform unterstützen. Insbesondere ermöglicht die Galaxy-Plattform aus Freiburg in Zeiten der Corona-Forschung eine schnelle und unkomplizierte Analyse von SARS-CoV-2-Daten (https://galaxyproject.eu/). Die prominenteste Nutzung erfährt jedoch die de.NBI-Cloud, da mit ihr äußerst effektiv und schnell die Analyse von anfallenden COVID-19-Daten durchgeführt werden kann.

Die Beteiligung des de.NBI-Netzwerks an europäischen COVID-19-Initiativen

Neben nationalen Aktivitäten ist das de.NBI-Netzwerk über den deutschen ELIXIR-Knoten ELIXIR Germany zusammen mit internationalen Partnern an europäischen COVID-19-Initiativen beteiligt. Dabei ist besonders die europäische COVID-19-Datenplattform zu nennen, die im Rahmen einer Initiative des Europäischen Bioinformatik-Instituts EMBL-EBI, von ELIXIR und der Europäischen Kommission initiiert wurde. Einen wesentlichen Teil dieser Plattform stellt das COVID-19-Datenportal (*COVID-19 Data Portal*) dar, welches am 20. April 2020 vom EMBL-EBI ins Leben gerufen wurde (https://www.covid19dataportal.org/). Das Ziel dieser neuen Infrastruktur ist die Zusammenführung relevanter COVID-19-Datensätze und Analysetools sowie dessen kontinuierliche Aktualisierung.

Dabei bietet das COVID-19-Datenportal Forschenden aus aller Welt unter dem Motto „Beschleunigung der Forschung durch Datenaustausch“ die Möglichkeit, ihr Wissen und ihre Daten im Zusammenhang mit dem neuartigen Coronavirus SARS-CoV-2 schnell, offen und nachhaltig mit der Gemeinschaft zu teilen und zu analysieren. Dabei können Forschenden nicht nur reine Sequenzen, sondern u. a. auch Struktur-, Expressions-, klinische und epidemiologische Daten sowie eine umfangreiche Literatursammlung der Wissenschaft zur Verfügung gestellt werden.

Zu den derzeitigen nationalen Partnern, die intensiv mit dem EMBL-EBI zusammenarbeiten, um die Infrastruktur der Plattform bereitzustellen, gehören neben der Technischen Universität Dänemark und der ungarischen Eötvös-Loránd-Universität auch das Universitätsklinikum Heidelberg. Außerdem leisten de.NBI-Mitglieder im Zusammenhang mit ELIXIR Europe einen wesentlichen Beitrag zum Aufbau sowie zur Instandhaltung des europäischen COVID-19-Datenportals.

Kontakt:

Prof. Dr. Alfred Pühler
Koordinator Deutsches Netzwerk
für Bioinformatik-Infrastruktur
Centrum für Biotechnologie
Universität Bielefeld
puehler@cebitec.uni-bielefeld.de

www.denbi.de

Dr. Irena Maus
Kommunikation und Öffentlichkeitsarbeit -
de.NBI Geschäftsstelle
irena.maus@cebitec.uni-bielefeld.de

Dr. Vera Ortseifen
Wissenschaftliche Mitarbeiterin -
de.NBI Geschäftsstelle
vera@cebitec.uni-bielefeld.de

apl. Prof. Dr. Andreas Tauch
Leiter der
de.NBI-Geschäftsstelle
tauch@cebitec.uni-bielefeld.de

Fotos: privat / Pühler, privat / Maus, privat / Ortseifen, privat / Tauch

de.NBI cloud als akademische lösung für lebenswissenschaftler

Cloud Computing bietet flexible skalierbare Rechen- und Speicherressourcen für Big Data Anwendungen

von Christian Lawerenz und Alexander Sczyrba

Die Fortschritte moderner Technologien stellen für die Lebenswissenschaften immer größere Herausforderungen dar. So entstehen etwa bei Genomanalysen oder auch bildgebenden Verfahren enorm große Datenmengen, die mit jeder neuen Generation von Laborgeräten immer schneller wachsen. Während selbst kleine Forschungslabore heute relativ einfach „Big Data“ erzeugen können, wird die Auswertung der Daten immer mehr zum Flaschenhals. Das Problem sind die lokal limitierten Rechenressourcen und die notwendige aufwendige Verwaltung von einer Vielzahl von Computersystemen. Cloud Computing bietet einen neuen Ansatz: hoch skalierbare Software-Lösungen werden auf performanter Infrastruktur ausgeführt, die je nach Bedarf dynamisch angemietet wird.

Eine Cloud für Alle

Das Bundesministerium für Bildung und Forschung (BMBF) hat zur Lösung des Big Data-Problems in den Lebenswissenschaften das Deutsche Netzwerk für Bioinformatik-Infrastruktur (de.NBI) ins Leben gerufen. Um den Biowissenschaften Rechen- und Speicherkapazitäten und wichtige de.NBI-Dienste zugänglich zu machen, wurde eine an de.NBI angebundene Cloud Infrastruktur etabliert.

Diese de.NBI Cloud hat als akademische und nicht-kommerzielle Cloud-Lösung zum Ziel, Lebenswissenschaftlern in Deutschland Rechen- und Speicherressourcen kostenlos zur Verfügung zu stellen. Durch die Bereitstellung relevanter Methoden und zugehöriger Referenzdaten des jeweiligen Forschungsbereichs, können eigene Datensätze angemessen analysiert werden. Dabei ist die Infrastruktur der de.NBI Cloud auf die speziellen Anforderungen von unterschiedlichsten Analysen in den Lebenswissenschaften zugeschnitten, beispielsweise durch GPU Cluster für Anwendungen im Bereich des Machine Learnings oder der Bildverarbeitung, oder durch spezielle HighMemory Rechner für besonders speicherintensive Anwendungen. Für den Wissenschaftler spielt es dabei keine Rolle mehr, an welchem Standort die Analysen tatsächlich durchgeführt werden. Durch die Nutzung von Cloud-Computing-Umgebungen können sich so erhebliche Kosteneinsparungen ergeben, eine Investition in lokale Computer-Hardware, oft nur sporadisch von einzelnen genutzt, ist nicht mehr nötig.

Zusätzlich ermöglicht der Cloudansatz die einfache Weiterverwendung der Analysetools. Über die eingesetzten Virtualisierungsmethoden kann die Software oft ohne großen Aufwand auf andere Systeme ausgerollt werden. Das de.NBI Cloud Personal bietet hier fachliche Expertise für die Einbettung kleiner Projekte zur Förderung des wissenschaftlichen Nachwuchses bis hin zu internationalen Großprojekten mit deutscher Beteiligung an. Zu diesen Projekten zählen unter anderem das ELIXIR-Netzwerk, eine zwischenstaatliche Organisation, die Life-Science-Ressourcen aus ganz Europa zusammenbringt, oder die European Open Science Cloud (EOSC).

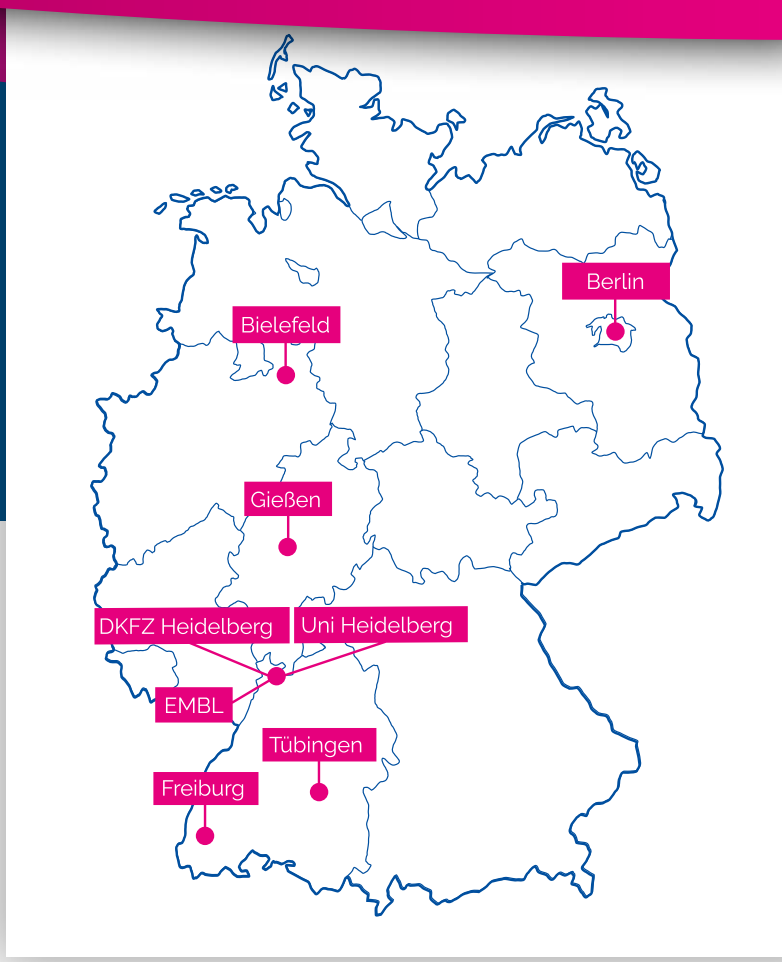


Abbildung 1: Verteilung der einzelnen de.NBI Cloud Standorte innerhalb Deutschlands (Quelle: https://cloud.denbi.de/cloud-federation).

de.NBI Cloud als Föderation deutscher Cloudstandorte

Die de.NBI Cloud ist eine Föderation mit Standorten an den Universitäten in Bielefeld, Freiburg, Gießen, Heidelberg und Tübingen, weiteren Installationen am Deutschen Krebsforschungszentrum (DKFZ), am Berliner Institut für Gesundheitsforschung (BIH) der Charité und einer geplanten Installation des assoziierten de.NBI Cloud Partners am European Molecular Biology Laboratory (EMBL in Heidelberg). Von Beginn an wurde ein Cloud-Verbundkonzept entwickelt, das alle Standorte in eine einzige Cloud-Plattform integriert. Alle Cloud-Sites sind in die Authentifizierungs- und Autorisierungsinfrastruktur (AAI) von ELIXIR integriert und bieten den Cloud-Benutzern somit einen einfachen Zugriff auf die de.NBI Cloud über das webbasierte de.NBI Cloud Portal. So kann das Benutzerkonto der eigenen Universität für den Zugang zum Portal und allen Services der de.NBI Cloud verwendet werden. Das de.NBI Cloud Portal (https://cloud.denbi.de) ist der zentrale Zugang für die Nutzer der de.NBI Cloud. Über das Portal können sich Nutzer für die Cloud registrieren, neue Projekte beantragen, sowie bestehende Projekte verwalten.

Alle de.NBI Cloud Standorte nutzen die OpenStack Software-Architektur für Cloud Computing. Diese steuert große Pools von Rechen-, Speicher- und Netzwerkressourcen in einem Rechenzentrum, die über ein Dashboard oder über die OpenStack-API verwaltet werden. OpenStack arbeitet mit gängigen Unternehmens- und Open Source-Technologien, ist daher ideal für die heterogenen Infrastrukturen der Cloud Standorte und ist somit ein entscheidender Baustein zur Sicherstellung der Interoperabilität.

Die de.NBI Cloud ist die größte deutsche akademische Cloud für nicht-kommerzielle und wissenschaftliche Zwecke. Mehr als 22.000 Rechenkerne, 38 Petabyte-Datei- und Objektspeicher und 220 Terabyte RAM stehen Wissenschaftlern aus den Lebenswissenschaften zurzeit zur Verfügung. Um auch zukünftige Anforderungen abdecken zu können, wird das Einbinden neuer Hardware stetig neu geplant und umgesetzt.

Wer nutzt die de.NBI Cloud?

Insgesamt sind derzeit über 1000 Cloud-Entwickler in 350 Cloud-Projekten registriert (siehe Abbildung 2). Bei diesen Projekten handelt es sich um Dienste, die jeweils von einer Handvoll von Entwicklern in die Cloud gebracht, aber oft von einer Vielzahl von Endbenutzern genutzt werden. Ein Beispiel dafür ist das Galaxy-Projekt am Standort Freiburg, das weltweit insgesamt 15.000 Nutzer unterstützt. Weitere Beispiele sind der SILVAngs Service für die Analyse von ribosomaler RNA mit ca. 200 Nutzern pro Monat oder der EggNOG Mapper für die Annotation von DNA Sequenzen, der pro Monat über 4500 Anfragen von über 600 Nutzern erhält.

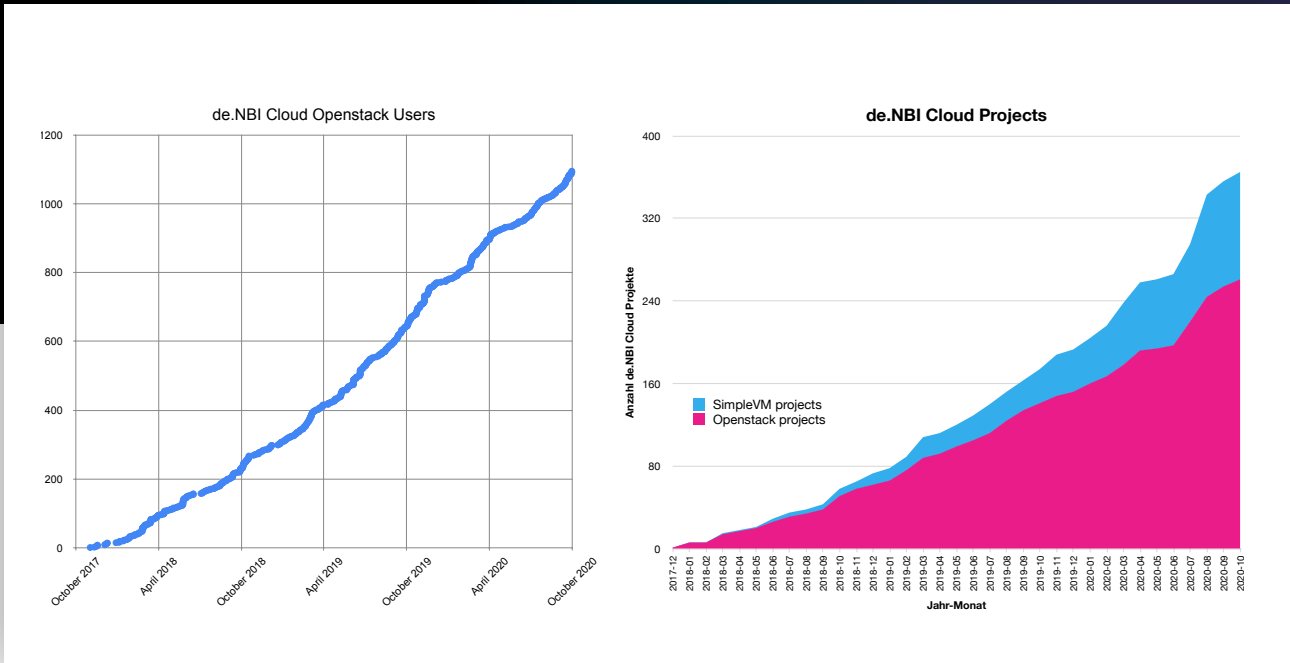


Abbildung 2: Entwicklung der Nutzerzahlen und Projekte der de.NBI Cloud (Quelle: Peter Belmann, https://cloud.denbi.de/cloud-numbers).

Die de.NBI Cloud wird von Wissenschaftlern unabhängig von ihren Vorkenntnissen im Bereich Cloud Computing verwendet. Darüber hinaus werden verschiedene Benutzerrollen angesprochen, die die jeweiligen IT-Erfahrungen widerspiegeln. Benutzer mit wenig IT Erfahrungen greifen eher über das zentralisierte de.NBI Cloud Portal auf einzelne virtuelle Maschinen über sogenannte „SimpleVM“ Projekte zu. So können spezialisierte Workflows besonders benutzerfreundlich zur Verfügung gestellt werden. Hochdurchsatzanalysen mit mehreren hundert Compute-Cores erfordern erfahrene Bioinformatiker mit direktem Zugriff auf die Cloud-Infrastruktur über die OpenStack-API. Und natürlich nutzen viele der de.NBI Servicezentren die Cloud Infrastruktur als Compute-Backend für ihre Serviceangebote. Entsprechend den unterschiedlichen Vorkenntnissen werden die Cloud-Benutzer in einer Vielzahl von Tutorials, Workshops, Sommerschulen, Hackathons bis hin zu individuellen technischen Besprechungen geschult. Neben der technischen Hilfestellung steuern die de.NBI Standorte ihre jeweilige besondere wissenschaftliche Expertise zu einer Vielzahl von Cloudprojekten bei. Themen wie Humangenomik, mikrobielle (Meta-) Genomik, Pflanzengenomik, Proteomik, Systembiologie oder integrative Bioinformatik sind an den verschiedenen Standorten schwerpunktmäßig vertreten.

Im besonderen Fokus für die Zukunft: Datenschutz und IT-Sicherheit

An den de.NBI Cloud-Standorten werden umfangreiche IT-Sicherheitskonzepte umgesetzt, um ein hohes Maß an Datenschutz und Datensicherheit zu gewährleisten. Eine der zentralen Anforderungen an die de.NBI Cloud-Infrastruktur ist eine solide Sicherheitsarchitektur. Die Umsetzung dieser Sicherheitskonzepte und Datenschutzregelungen soll zukünftig die Verarbeitung personenbezogener Daten in der de.NBI Cloud ermöglichen. Um diese notwendigen Anforderungen nachhaltig und nachweisbar umzusetzen und ein möglichst hohes Maß an IT-Sicherheit zu gewährleisten, streben die de.NBI Cloud Standorte an, Bewertungen des jeweiligen Cloudsystems vorzunehmen und sich entweder nach IS0 27001 und 2/27017 zertifizieren zu lassen oder ein Prüfsiegel nach dem Cloud Computing Compliance Controls Catalogue (C5) des Bundesamtes für Sicherheit in der Informationstechnik (BSI) zu erhalten. Somit wird der de.NBI Cloud Nutzer den Nachweis einer vertrauenswürdigen Cloudinfrastruktur mit den notwendigen Maßnahmen hinsichtlich Integrität und Vertraulichkeit der Daten erhalten.

Cloud Standorte

de.NBI Geschäftsstelle – Cloud Governance (Peter Belmann)
Justus-Liebig-Universität Gießen (Prof. Dr. Alexander Goesmann)
Universität Freiburg (Dr. Björn Grüning)
Universität Tübingen (Dr. Jens Krüger)
Universität Heidelberg / DKFZ (Christian Lawerenz)
EMBL Heidelberg (Dr. Jan Korbel)
Universität Bielefeld (Prof. Dr. Alexander Sczyrba)
Charité – Universitätsmedizin Berlin (Harald Wagener)

Kontakt:

Prof. Dr. Alexander Sczyrba
Universität Bielefeld
Bielefelder Institut für Bioinformatik-Infrastruktur
asczyrba@cebitec.uni-bielefeld.de

https://ekvv.uni-bielefeld.de/pers_publ/publ/PersonDetail.jsp?personId=17894

Christian Lawerenz
Steinbeis-Transferzentrum Health Data
christian.lawerenz@stw.de

www.steinbeis.de/su/2191

Abbildung 3: Schematische Darstellung für die Antragstellung eines Projektes in der de.NBI Cloud (Quelle: Alexander Sczyrba und Susanne Konermann).

Fotos: privat / Sczyrba, privat / Lawerenz

datenintegrationszentrum – drehscheibe für daten in der medizinischen forschung und versorgung

Datenintegration und ihre Voraussetzungen[1]

von Björn Schreiweis, Danny Ammon, Martin Sedlmayr, Fady Albashiti und Thomas Wendt

- In der Medizininformatik-Initiative etablieren fast alle deutschen Universitätsklinika Datenintegrationszentren (DIZ).
- In den DIZ werden Daten aus Krankenversorgung und Forschung integriert und für klinische Forschungsprojekte verfügbar gemacht.
- Der Einsatz von Interoperabilitätsstandards für die Erschließung und Verarbeitung der Daten spielt eine tragende Rolle.
- Die DIZ unterstützen weiterhin Prozesse zur Machbarkeitsprüfung, Beantragung, Begutachtung und Durchführung klinischer Forschungsprojekte.

Das Datenintegrationszentrum

Die standortübergreifende Nutzung von Daten stellt bis heute in vielen Forschungsprojekten eine Hürde dar. Kliniker und medizinische Forscher wünschen sich schon lange zuverlässige Werkzeuge, die ihnen helfen, Fragen mit Hilfe von Analysen der Routinedaten aus dem Gesundheitswesen zu beantworten. Deshalb werden in der Medizininformatik-Initiative (MII) Datenintegrationszentren (DIZ) aufgebaut – in nahezu allen deutschen Universitätskliniken bzw. deren Universitäten (Schreiweis *et al.*, 2019). In den DIZ werden die technischen und organisatorischen Voraussetzungen geschaffen, um medizinische Daten so zusammenzuführen und bereitzustellen, dass sie für die Forschung und deren Rückkopplung in die Versorgung optimal genutzt werden können.

Was kann ein Datenintegrationszentrum?

Beispiele, in denen DIZ unterstützen können:

- Sie fragen sich, wie viele geeignete Patienten voraussichtlich im kommenden Jahr für die Durchführung einer Studie zu erwarten sind?
- Sie möchten wissen, ob die Nutzung der Patientendaten des Klinikums in Ihrem Projekt erlaubt ist?
- Sie benötigen für Ihre Auswertungen Daten aus der Intensivmedizin oder aus der Mikrobiologie?
- Sie möchten Daten Ihres Klinikums mit Daten anderer Einrichtungen zusammenführen und auswerten?
- Sie möchten klinische Patientendaten für ein Forschungsprojekt nach einem anerkannten Verfahren pseudonymisieren?
- Sie möchten Daten aus Forschungsergebnissen unmittelbar in der Patientenversorgung anwenden?

Neben forschungsmotivierten Fragen können die Dienste von DIZ auch für die Krankenversorgung, das Qualitätsmanagement, für Planungsfragestellungen oder andere Integrationsaufgaben genutzt werden. Statt Daten mühsam aus verschiedenen Quellsystemen stets aufs Neue zu extrahieren und über diverse Wege freigeben zu lassen, wird das DIZ zum zentralen Ansprechpartner für klinische Daten und Services um diese Daten.

[1] Der Artikel ist eine aktualisierte Version des Artikels „Das Datenintegrationszentrum – Ausgangspunkt für die datengetriebene medizinische Forschung und Versorgung“ aus der mdi-Themenheftausgabe 4/2019 zur Medizininformatik-Initiative herausgegeben von BVMI e.V. und DVMD e.V. (https://www.bvmi.de/mdi)

Abbildung 1: Standorte der Datenintegrationszentren der Medizininformatik-Initiative (Quelle: TMF e.V.).

Woran arbeiten Datenintegrationszentren?

Derzeit werden die technischen und organisatorischen Strukturen sowie die Prozesse der DIZ aufgebaut. Dazu gehören u. a.

- Komponenten, mit denen Daten aufbereitet und in standardisierter Form zur Verfügung gestellt werden können, inkl. Zugriffsmanagement und Pseudonymisierung,
- Einwilligungsmanagement,
- Vorlagen für Nutzungsanträge und -verträge
- Data Use and Access Commitees (UAC).

In jedem Konsortium der MII wurden Use Cases definiert, mit denen die Leistungsfähigkeit der DIZ gezeigt wird. Beispiele sind: Diagnostik und Versorgung des akuten Lungenversagens und der Blutstrominfektionen, sowie ein generischer Ansatz zur Identifikation von Phänotypen (SMITH); Entwicklung von Klassifikations-/Prädiktionsmodellen für Patienten mit COPD und Hirntumoren (MIRACUM); Visualisierung von lokalen Infektionsclustern, aber auch Aufbau einrichtungsübergreifender molekularer Tumorboards (HiGHmed); Verbesserungen in der Versorgung der Multiplen Sklerose, des Morbus Parkinson und bestimmter onkologischer Erkrankungen (DIFUTURE) (Hemmer *et al.*, 2019).

Wie ist die Bereitstellung von Daten für die Forschung organisiert?

Forscher stellen zunächst Anträge, in denen u. a. das Forschungsthema und die benötigten Daten beschrieben sind. Diese Anträge beurteilt an jedem angefragten Standort ein UAC u. a. nach bestimmten organisatorischen und rechtlichen Kriterien, aber auch hinsichtlich der Verfügbarkeit der Daten. Nach positivem Votum schließt das DIZ einen Nutzungsvertrag, der Forscher bspw. zum sorgfältigen Umgang mit den Daten sowie zur Rückübermittlung der Ergebnisse verpflichtet.

Neben dem Austausch der Daten werden – auch nach Antrag – dezentrale Analysen ermöglicht. Die Zentralisierung von Daten ist in Zeiten von »Big Data« weder immer sinnvoll, noch datenschutzrechtlich immer möglich. Werden Analysen aber lokal an den Standorten durchgeführt und nur die (Teil-) Ergebnisse zentral zurückgegeben, reduzieren sich Risiken.

Anträge, Verträge und Prozesse werden auf nationaler Ebene harmonisiert. Als Anlaufpunkt für Forscher wird eine Zentrale Antrags- und Registerstelle aufgebaut, z. B. zur Weiterleitung von Anträge an die DIZ.

Interoperabilität

Die Schnittstellen zwischen den DIZ und zu den Nutzern werden in MII-Arbeitsgruppen konsentiert. Damit können technische Konzepte der DIZ teilweise unterschiedlich sein, wahren aber die Interoperabilität auf dem Weg zu einer nationalen Dateninfrastruktur. Dafür steht u. a. der gemeinsam entwickelte Kerndatensatz (Ammon *et al.*, 2019). Damit wird festgelegt, wie Daten inhaltlich und technisch aufbereitet sein müssen, so dass sie für weitere Analysen verwendet werden können.

Integration als Multi-Themen-Aufgabe

Bei der Kooperation der Einrichtungen der MII bleiben Unterschiede in Sichtweisen, Interessen, Verpflichtungen nicht aus. Die DIZ arbeiten mit am gemeinsamen Verständnis über Möglichkeiten und Grenzen der Nutzung von Patientendaten in einrichtungsübergreifenden Projekten.

Auch ganz konkret ist für jedes Projekt zu klären, welche Daten bereitgestellt werden sollen, mit welchen Formaten und Terminologien sowie über welche Schnittstellen.

Was ist der aktuelle Stand?

Der DIZ-Aufbau wird über fünf Jahre gefördert. In den ersten drei Jahren konnten bereits Projekte durch die Bereitstellung von Daten unterstützt werden. Insbesondere in den Use Cases wurden sehr gute Fortschritte erzielt, technisch und medizinisch-fachlich. Aktuell werden aufgrund der SARS-COV 2 Pandemie insbesondere die Themen Infektionskontrolle (Smart Infection Control System (SmICS)) und Algorithmic Surveillance kritisch Kranker auf alle DIZ ausgeweitet.

Darüber hinaus zeigt die MII, dass DIZ auch interkonsortial Daten analysieren können (vgl. Demonstrator-Studie (Ganslandt *et al.*, 2019)). Ausgehend von der Demonstrator-Studie werden zwei weiterführende Projekte – CORD_MI (Collaboration on Rare Diseases) zur Verbesserung der Dokumentation seltener Erkrankungen in den klinischen Systemen und POLAR_MI (Polypharmazie – Arzneimittelnebenwirkungen – Risiken) zur Minimierung von Risiken der Polypharmazie – durchgeführt.

Über die Use Cases hinaus sind viele DIZ lokal aktiv. Nutzungsanfragen wurden nach den individuellen Möglichkeiten der DIZ beantwortet, auch wenn die Routinestrukturen teilweise noch aufgebaut werden.

Die bisherigen Arbeiten zeigen auch Grenzen auf. Nicht alle eingesetzten klinischen Anwendungssysteme sind in der Lage, Daten auf geeignete Weise bereitzustellen. So entstehen hohe Aufwände für die Aufbereitung. Die Harmonisierung der Nutzungsanträge und -verträge sowie der breiten Patienteneinwilligung erfordert viele Abstimmungen, auch mit Partnern außerhalb der MII. Gemeinsam mit den Datenschutzbehörden der

Steckbrief Medizininformatik-Initiative (MII)

Ziel der Medizininformatik-Initiative (MII) ist die Verbesserung von Forschungsmöglichkeiten und Patientenversorgung durch innovative IT-Lösungen. Diese sollen den Austausch und die Nutzung von Daten aus Krankenversorgung, klinischer und biomedizinischer Forschung über die Grenzen von Institutionen und Standorten hinweg ermöglichen. Das Bundesministerium für Bildung und Forschung (BMBF) fördert die MII bis 2022 mit rund 160 Millionen Euro. In den vier Konsortien DIFUTURE, HiGHmed, MIRACUM und SMITH arbeiten alle Einrichtungen der Universitätsmedizin in Deutschland an über 30 Standorten gemeinsam mit Forschungseinrichtungen, Unternehmen, Krankenkassen sowie Patientenvertreterinnen und -vertretern daran, die Rahmenbedingungen zu entwickeln, damit Erkenntnisse aus der Forschung direkt die Patientinnen und Patienten erreichen können. Datenschutz und Datensicherheit haben dabei höchste Priorität.

Für die nationale Abstimmung der Entwicklungen innerhalb der MII ist eine Koordinationsstelle zuständig, die die Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (TMF) gemeinsam mit dem Medizinischen Fakultätentag (MFT) und dem Verband der Universitätsklinika Deutschlands e.V. (VUD) in Berlin betreibt.

Weitere Informationen:
https://www.medizininformatik-initiative.de und https://medizininformatik-karte.de

Länder, den Ethikkommissionen und den Biobanken wurde eine national einheitliche Einwilligungserklärung zur Datennutzung verabschiedet (Medizininformatik-Initiative, 2020).

Ausblick

Der DIZ-Aufbau wird mit großem Enthusiasmus vorangetrieben. Die DIZ schaffen nicht nur Services für die datengetriebene und vernetzte medizinische Forschung, sondern bündeln das Know-How über den Lebenszyklus von Daten. Sie unterstützen bald auch Studien, bei denen z.B. Künstliche Intelligenz zum Einsatz kommt, aber auch Prozesse des Qualitätsmanagements oder des Controllings.

Perspektivisch sollen DIZ auch über die Universitätskliniken hinauswachsen und sich mit außeruniversitären Leistungserbringern verknüpfen, wie es einige bereits tun.

DIZ sind zentrale Ansprechpartner, wenn es um klinische Daten geht. Fragen Sie bei Ihrer Universitätsmedizin nach!

In den DIZ werden die Voraussetzungen geschaffen, um medizinische Daten so zusammenzuführen und bereitzustellen, dass sie für die Forschung optimal genutzt werden können (Foto: shutterstock / Panchenko Vladimir).

Referenzen:

Schreiweis, B., Ammon, D., Sedlmayr, M., Albashiti, F., Wendt, T. (2019). Das Datenintegrationszentrum – Ausgangspunkt für die datengetriebene medizinische Forschung und Versorgung. mdi 4 2019, 21, 106-110.

Hemmer, B., Börries, M., Christoph, J., Marx, G., Maaßen, O., Schuppert, A., Scheithauer, S. (2019). Die klinischen Anwendungsbeispiele (Use Cases) der vier MII-Konsortien. mdi 4 2019, 21, 98-102.

Ammon, D., Bietenbeck, A., Boeker, M., Ganslandt, T., Heckmann, S., Heitmann, K., Sax, U., Schepers, J., Semler, S.C., Thun, S., Zautke, A. (2019). Der Kerndatensatz der Medizininformatik-Initiative – Interoperable Spezifikation am Beispiel der Laborbefunde mittels LOINC und FHIR. mdi 4 2019, 21, 113-117.

Ganslandt, T., Schaaf, J., Schepers, J., Storf, H., Balzer, F., Haferkamp, S., Lodahl, R., Prasser, F., Sax, U., Stenzhorn, H. and Prokosch, H.U. (2019). Experiences from the National Demonstrator Study within the German Medical Informatics Initiative. In: AMIA 2019 Annual Symposium, 16.–20. November 2019, Washington, D.C.

Medizininformatik-Initiative (2020). Medizininformatik-Initiative erhält grünes Licht für bundesweite Patienteneinwilligung. Pressemitteilung. 27.04.2020. https://www.medizininformatik-initiative.de/de/medizininformatik-initiative-erhaelt-gruenes-licht-fuer-bundesweite-patienteneinwilligung (letzter Zugriff: 17.08.2020)

Kontakt:

Dr. sc. hum. Björn Schreiweis
Senior Researcher
Christian-Albrechts-Universität zu Kiel &
Universitätsklinikum Schleswig-Holstein
bjoern.schreiweis@uksh.de

Danny Ammon
Leiter des Datenintegrationszentrums
Universitätsklinikum Jena
danny.ammon@med.uni-jena.de

Prof. Dr. rer. nat. Martin Sedlmayr
Professor für Medizinische Informatik
Technische Universität Dresden
martin.sedlmayr@tu-dresden.de

Dr. sc. hum. Fady Albashiti
CEO des Zentrums für Medizinische Datenintegration und –analyse
Klinikum der Universität München
fady.albashiti@med.uni-muenchen.de

Dr. rer. med. Thomas Wendt
Leiter des Datenintegrationszentrums
Universitätsklinikum Leipzig
thomas.wendt@medizin.uni-leipzig.de

Fotos: privat / Schreiweis, privat / Ammon, privat / Sedlmayr, privat / Albashiti, privat / Wendt

einzug der massenspektrometrie in die systemmedizin

Vier Forschungskerne auf der Suche nach neuen Biomarkern

von Jeroen Krijgsveld, Ursula Klingmüller, Carsten Müller-Tidow, Bernhard Küster, Daniel Teupser, Ulrich Keilholz, Frederick Klauschen, Markus Ralser, Matthias Selbach, Philipp Wild und Stefan Tenzer

Im Rahmen der Hightech-Strategie 2025 der Bundesregierung und der Initiative „Forschungskerne für Massenspektrometrie in der Systemmedizin“ fördert das Bundesministerium für Bildung und Forschung (BMBF) die Entwicklung neuer Analysewerkzeuge für die Gesundheitsforschung. An vier Standorten deutschlandweit sollen diese interdisziplinären Forschungsprojekte Möglichkeiten zur klinischen Nutzung massenspektrometrischer Methoden etablieren.

Die ganzheitliche Betrachtungsweise der Systemmedizin und die Etablierung der Massenspektrometrie als eine Schlüsseltechnologie zur Quantifizierung krankheitsrelevanter Biomoleküle (Proteine, Lipide und Metabolite) eröffnet ein großes Innovationspotential: Da meist mehrere Faktoren bei der Entstehung von Krankheiten und ihrem weiteren Verlauf von Bedeutung sind, ist es essentiell, diese in ihrer Gesamtheit zu erfassen und quantitativ zu analysieren. Daraus resultierende Erkenntnisse können vielfältig genutzt werden: Zum einen tragen sie zu einem tieferen Verständnis der Krankheitsentstehung bei. Zum anderen können sie durch ihre Implementierung in der Patientenversorgung zahlreiche Aspekte wie (frühzeitige) Diagnosen, Prognoseerstellungen sowie Therapieempfehlungen verbessern. Nicht zuletzt könnten so neue zielgerichtete, personalisierte Therapieansätze entwickelt werden.

Systemmedizinische Ansätze erfordern ein hohes Maß an interdisziplinärer Zusammenarbeit zwischen Analytikern, Medizinern und (Bio-) Informatikern. Massenspektrometrische Methoden sind in der medizinischen Diagnostik bislang noch stark unterrepräsentiert und ihre Möglichkeiten bei weitem nicht ausgeschöpft. Dies ist unter anderem auf das Fehlen standardisierter Abläufe und technischer Limitationen bisheriger Massenspektrometer zurückzuführen. Zum Aufbau eines massenspektrometrischen Netzwerks in der Systemmedizin in Deutschland fördert das BMBF nun vier Forschungskerne in Berlin, Heidelberg, Mainz und München mit einem Gesamtvolumen von 26,6 Millionen Euro in den nächsten drei Jahren.

Das gemeinsame Ziel der Forschungskerne ist es, die enge Vernetzung der Bereiche Massenspektrometrie, Medizin und Informatik zu fördern und deren Zusammenspiel im klinischen Alltag zu etablieren, um somit eine breitere Anwendung der Systemmedizin zu erreichen (Abbildung 1). Durch die massenspektrometrische Analyse diverser klinischer Proben hoffen die beteiligten Forscher, molekulare Signaturen zu entdecken, die als neue Biomarker für Diagnose, Prognose, Therapieentscheidung oder Therapieansprechen klinisch verwertbar gemacht werden können. Hierzu sollen standortübergreifend „best practice workflows“ etabliert werden. Neben den notwendigen medizinischen, logistischen und technischen Infrastrukturen bezüglich Probensammlung, -extraktion und -analyse werden die Forschungskerne auch die erarbeiteten neuen methodischen Ansätze hinsichtlich Probendurchsatz, Robustheit und Reproduzierbarkeit optimieren und standardisieren (Abbildung 2).

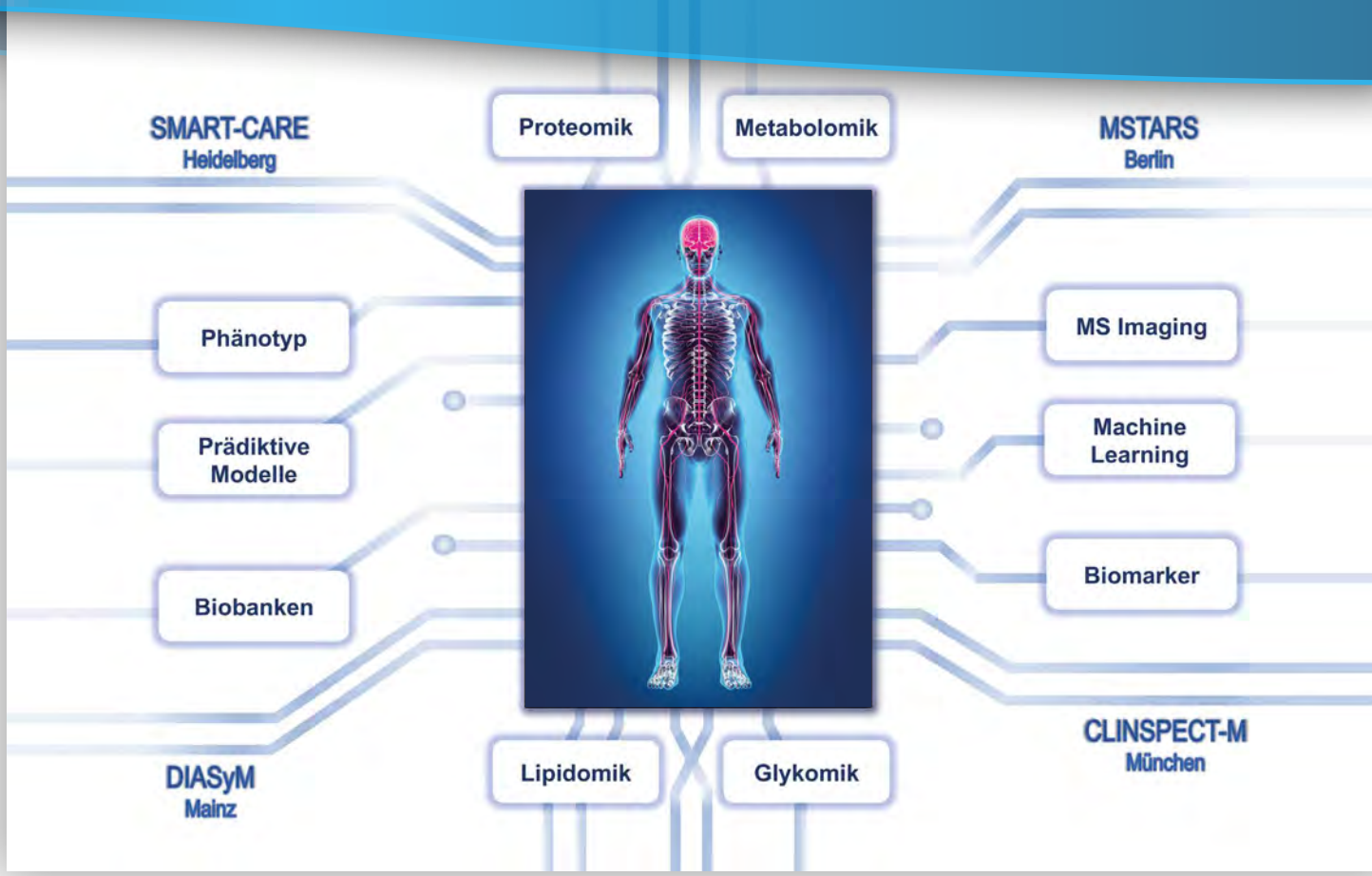


Abbildung 1: Die vier neuen Forschungskerne innerhalb der MSCoreSys Initiative befassen sich mit der Entwicklung von massenspektrometrischen Methoden zur Etablierung und Weiterentwicklung systemmedizinischer Ansätze zur Aufklärung pathophysiologischer Mechanismen (Quelle: Philipp Wild und yodiyim - fotolia.com).

Die vier Forschungskerne beschäftigen sich mit unterschiedlichen Anwendungsfeldern und Schwerpunkten, die im Folgenden vorgestellt werden:

DIASyM in Mainz

Der Mainzer Forschungskern DIASyM *(Data-Independent Acquisition-based Systems Medicine: Mass spectrometry for high-throughput deep phenotyping of the heart failure syndrome)* wird gemeinschaftlich von Stefan Tenzer, Koordinator der massenspektrometrischen Technologie-Plattform und Methodenforschung im neuen Forschungskern, und Philipp Wild, der die Systemmedizin koordiniert, geleitet. DIASyM bündelt die Expertise von Wissenschaftlern und Medizinern der Universitätsmedizin und der Johannes Gutenberg-Universität Mainz. Europaweit leiden ca. 15 Millionen Menschen an Herzinsuffizienz, die eine der Hauptursachen für Krankenhausaufenthalte für Patientinnen und Patienten über 65 Jahren darstellt. Durch die begrenzte, symptomatische Behandlung schreitet die Erkrankung konstant fort und resultiert letztlich in einer verkürzten Lebenserwartung. Zusammen mit nur wenigen Behandlungsoptionen führt dies zu einer erheblichen Belastung des Gesundheitswesens.

Die pathophysiologischen Grundlagen der Erkrankung sind nur wenig erforscht. Innerhalb des Mainzer Forschungskonsortiums werden optimierte datenunabhängige massenspektrometrische Messverfahren für die Hochdurchsatz-Phänotypisierung des Herzinsuffizienz-Syndroms etabliert. DIASyM nutzt und analysiert massenspektrometrische Informationen aus der Proteomik, Lipidomik und Metabolomik in Verbindung mit einer Vielzahl klinischer Parameter unter Nutzung modernster Methoden der künstlichen Intelligenz wie beispielsweise des tiefen oder probabilistischen maschinellen Lernens. Durch die Anwendung eines systemmedizinisch orientierten Ansatzes versprechen sich die Wissenschaftlerinnen und Wissenschaftler ein tieferes Verständnis der zugrundeliegenden biologischen Prozesse der Erkrankung. Dies ist die Grundlage für die Entwicklung punktgenauer Therapien, die möglichen Spätfolgen der Herzinsuffizienz früh entgegenwirken und damit die Lebensqualität und die Lebenserwartung der Patientinnen und Patienten deutlich verbessern können.

CLINSPECT-M in München

Der Münchener Forschungskern CLINSPECT-M *(Clinical Mass Spectrometry Center Munich)* wird von Bernhard Küster und Daniel Teupser koordiniert und vereint die Expertisen der

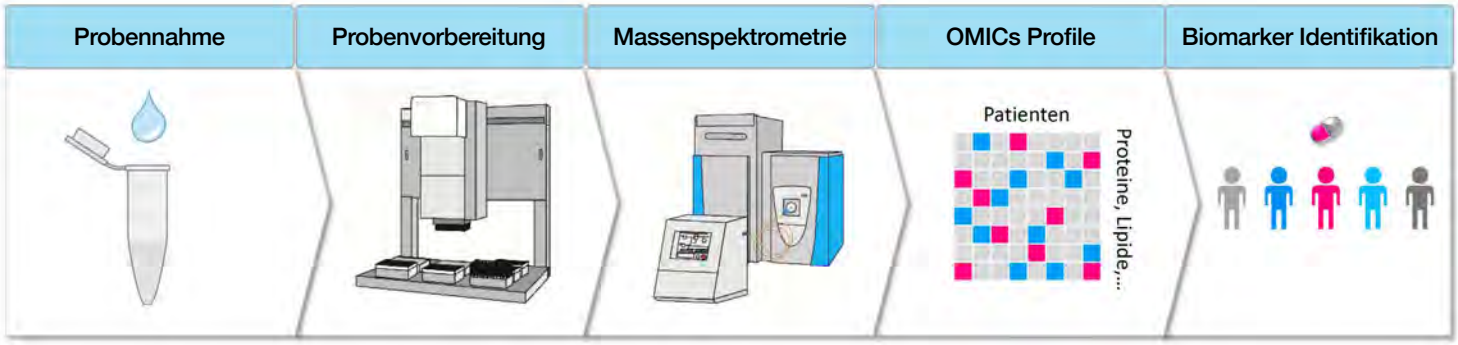


Abbildung 2: Schematischer Workflow von der Probennahme bis zur (möglichen) Biomarkeridentifizierung (Quelle: Stephanie Heinzlmeir).

Technischen Universität München, der Ludwig-Maximilians-Universität München, beider Universitätsklinika, des Helmholtz Zentrums München und des Max-Planck Instituts für Biochemie. Neben geplanten technischen Entwicklungen in den Bereichen Proteomik, Medizin und Bioinformatik, erhält dieser interdisziplinäre Verbund einen biologisch-medizinischen Fokus durch die Forschung an Erkrankungen des Nervensystems, insbesondere Multiple Sklerose, Alzheimer, Schlaganfall und Krebs.

Trotz der hohen Inzidenz und gesellschaftlichen Relevanz bleiben bislang viele Aspekte dieser Volkskrankheiten unverstanden: Wodurch unterscheidet sich die Multiple Sklerose von anderen neuro-inflammatorischen Krankheiten? Warum sprechen manche Alzheimer Patienten besser auf eine medikamentöse Therapie an als andere? Wie können wir ischämische Schlaganfall-Patienten von sogenannten „stroke mimics“ unterscheiden? Kann das proteomische Profil von Tumorpatienten helfen, effektivere personalisierte Therapien zu empfehlen? Diese und weitere Fragen sollen im Rahmen von CLINSPECT-M mit Hilfe modernster massenspektrometrischer und bioinformatischer Ansätze adressiert werden und neue Erkenntnisse zu pathologischen Prozessen im Gehirn als auch deren Behandlung liefern. Die klinische Anwendbarkeit der Proteomik in diesen Krankheitsfeldern befindet sich in unterschiedlichen Entwicklungsstufen. CLINSPECT-M wird somit exemplarisch kurz-, mittel- und langfristige Möglichkeiten aufzeigen, um die Technologie zunehmend im klinischen Alltag zu etablieren.

MSTARS in Berlin

Der Berliner Forschungskern MSTARS *(Multimodal clinical mass spectrometry to target treatment resistance)* wird gleichberechtigt von den vier Koordinatoren Ulrich Keilholz (Charité Comprehensive Cancer Center), Frederick Klauschen (Charité, Institut für Pathologie), Markus Ralser (Charité, Institut für Biochemie) und Matthias Selbach (Max-Delbrück-Centrum für Molekulare Medizin in der Helmholtz Gemeinschaft, MDC) geleitet. Die meisten chronischen Erkrankungen weisen eine komplexe Pathophysiologie auf, wobei die zugrundeliegenden fundamentalen Pathomechanismen molekular meist besser verstanden sind als jene Mechanismen, welche Ausheilung und Progression regulieren. Hierbei stellt das Therapieansprechen auf Ebene des individuellen Patienten eine krankheitsübergreifende Herausforderung in der Präzisionsmedizin dar.

Durch die umfassende und gezielte Bündelung erstklassiger Spitzenforschung wird der Berliner Forschungskern komplementäre Technologien der (Phospho)-Proteomik, Metabolomik, Glykomik und bildgebenden Massenspektrometrie mit fundierter klinischer Expertise in einen multimodalen Ansatz integrieren, welcher durch neueste Verfahren aus den Bereichen Computational Research und Machine Learning ergänzt wird. Während das Forschungskonzept, welches eine mechanistische und Signatur-abgeleitete Strategie umfasst, allgemein anwendbar ist, wird der medizinische Fokus in diesem Forschungskern vor allem auf Krebs und entzündlichen Erkrankungen liegen. Für dieses Vorhaben stehen dem Konsortium einzigartige sehr große Sammlungen klinischer Proben und präklinischer Modelle zur Verfügung, wobei das Kopf-Hals-Plattenepithelkarzinom als primärer Anwendungsfall fungieren wird.

SMART-CARE in Heidelberg

Der Heidelberger Forschungskern SMART-CARE *(A Systems Medicine Approach to Stratification of Cancer Recurrence)* wird koordiniert von Jeroen Krijgsveld (Deutsches Krebsforschungszentrum, DKFZ, und Universitätsklinikum Heidelberg, UKHD), Ursula Klingmüller (DKFZ) und Carsten Müller-Tidow (UKHD). Beteiligt sind darüber hinaus Wissenschaftler und Ärzte fünf verschiedener (Uni-) Kliniken in Heidelberg, Wissenschaftler der Universität Heidelberg, des EMBL und des CeMOS Forschungszentrums der Hochschule Mannheim. Ziel des Forschungskerns ist es, tiefer in die molekularen Hintergründe individueller Krebserkrankungen, insbesondere beim wiederauftretenden Krebs (Rezidiv) vorzudringen.

Der Rückfall bei Krebserkrankungen ist der Hauptgrund für krebsbedingte Todesfälle. Die Forscher wollen daher molekulare Marker in verschiedenen Gewebeproben, Blut oder Nervenwasser identifizieren, die die Wahrscheinlichkeit des Wiederauftretens von Tumoren oder das Fortschreiten einer Krebserkrankung besser bestimmbar und den Erfolg von verschiedenen Therapieansätzen vorhersagbar machen. Sie konzentrieren sich dabei auf Blutkrebs, Lungenkarzinome, Gehirntumore und Sarkome. Systematische und quantitative Analysen der Protein- und Stoffwechselzusammensetzung dieser vier Krebsarten aus verschiedenen therapeutischen Ansätzen fließen, unterstützt von künstlicher Intelligenz und maschinellem Lernen, in mathematische Modellierungen des zellulären Zusammenspiels ein und erzeugen letztlich einen vielversprechenden Datenschatz für die Entwicklung einer neuen Generation an Biomarker-Mustern.

Dadurch soll das Risiko von Rückfällen der Krebserkrankung gesenkt sowie die Möglichkeit geschaffen werden, spezifische und individuelle Therapievorschläge der Krebsmedizin zur Verfügung zu stellen.

Kontakt:

Aktuelle Sprecher der Initiative MSCoreSys:

Prof. Dr. Jeroen Krijgsveld
Deutsches Krebsforschungszentrum (DKFZ)
und
Universitätsklinikum Heidelberg (UKHD)
j.krijgsveld@dkfz-heidelberg.de

https://www.dkfz.de/en/proteomik-stammzellen-krebs/index.php

Prof. Dr. Stefan Tenzer
Institut für Immunologie
Universitätsmedizin der Johannes
Gutenberg-Universität Mainz
tenzer@uni-mainz.de

http://www.tenzerlab.de

Homepage der Initiative:
www.mscoresys.de

Fotos: Jutta Jung / DKFZ, Thomas Böhm / Universitätsmedizin Mainz

modellierung von infektionserkrankungen der atemwege

Systemmedizinische Modellierung von Therapieoptionen bei ambulant erworbener Pneumonie und COVID-19

von Peter Ahnert und Markus Scholz

Infektionen der Atemwege gehörten auch schon vor COVID-19 zu den häufigsten Infektionserkrankungen weltweit. Infekte der unteren Atemwege sind dabei unter den fünf häufigsten Todesursachen (Michaud, 2009). Einen bedeutenden Anteil daran hat die ambulant erworbene Pneumonie (CAP). Der Krankheitsverlauf der CAP ist individuell sehr unterschiedlich mit der Möglichkeit rapider teilweise schwer vorhersehbarer Verschlechterungen. Im Verbundprojekt CAPSyS untersuchen wir Mechanismen, die den Krankheitsverlauf beeinflussen, mittels systemmedizinischer Ansätze. Auch für das aktuelle COVID-19 Geschehen möchten wir mit diesem Ansatz einen Beitrag zu besseren Prognose- und Therapiemöglichkeiten leisten.

CAP – Infektionserkrankung der Lunge mit systemischen Auswirkungen

Die ambulant erworbene Pneumonie *(community acquired pneumonia, CAP)* ist eine häufige Infektionserkrankung der unteren Atemwege, die insbesondere bei älteren Menschen und Menschen mit Vorerkrankungen einen sehr schweren Verlauf nehmen kann. Der außerordentlich heterogene Krankheitsverlauf zeichnet sich durch die Möglichkeit rapider teilweise schwer vorhersehbarer Verschlechterungen aus. Zudem ist das Arsenal möglicher therapeutischer Interventionen begrenzt. Ein kritischer Krankheitsverlauf ist häufig dadurch gekennzeichnet, dass die epithelial-endotheliale Barriere, die das Infektionsgeschehen in der Lunge vom Blutkreislauf abgrenzt, in ihrer Funktionsweise derart eingeschränkt ist, dass die Infektion systemisch wird. Damit ist die Infektion nicht mehr räumlich begrenzt, sondern greift auf den gesamten Organismus über. Hierdurch entstehen septische Zustände und damit verbunden Schädigungen anderer Organsysteme, wie zum Beispiel Leber und Niere. Jedoch können Schädigungen auch bei intakter Barriere aufgrund der systemischen Immunantwort auftreten. Das Verständnis der Prozesse um die Barriere stellt deshalb einen Schwerpunkt der Arbeiten in CAPSyS dar, da hier Hoffnung auf neue therapeutische Ansätze besteht. Aus diesem Grunde werden neben der Erforschung des Geschehens beim Menschen auch gezielte Mausexperimente durchgeführt, um die Barrierefunktion unter verschiedenen Interventionen im Verlauf der Infektion zu untersuchen.

Lungenentzündungen können auch durch Infektion der Atemwege mit dem neuen Erreger SARS-CoV-2 hervorgerufen werden. In ihrem Verlauf unterscheidet sich diese Erkrankung jedoch von CAP anderer Ursache hinsichtlich Manifestation und Dynamik teilweise deutlich (Zhou *et al.,* 2020). Bei COVID-19 kann ein schwerer Lungenschaden mit einem relativen Wohlbefinden einhergehen. Im weiteren Verlauf der Erkrankung kann es dann zu einem plötzlichen Lungenversagen kommen. Unseren im Rahmen von CAPSyS entwickelten systemmedizinischen Forschungsansatz übertragen wir deshalb kurzfristig auch auf die durch SARS-CoV-2 ausgelösten Lungenentzündungen.

In der Arbeitsgruppe Genetische Statistik und Systembiologie (Abbildung 1) am Institut für Medizinische Informatik, Statistik und Epidemiologie der Universität Leipzig sind wir im Rahmen von CAPSyS für die statistische und bioinformatische Analyse der umfangreichen Daten aus klinischen Studien und Mausexperimenten zuständig. Darüber hinaus entwickeln wir dynamische biomathematische Modelle, um verschiedene Aspekte des Infektionsgeschehens beschreiben und vorhersagen zu können. Diese Arbeiten und einige Ergebnisse werden im Folgenden kurz vorgestellt.

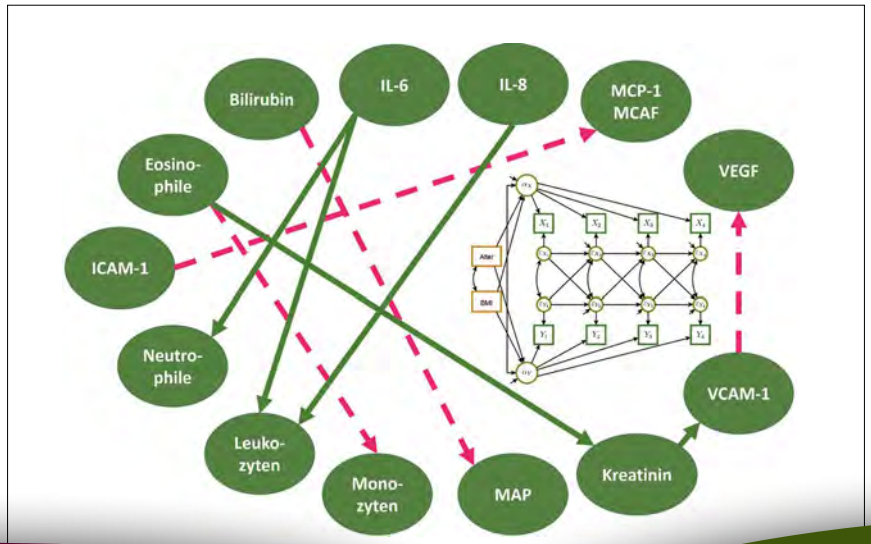
Abbildung 1: AG Genetische Statistik und Systembiologie am Institut für Medizinische Informatik, Statistik und Epidemiologie der Universität Leipzig (Professor Dr. Scholz (ganz links, zweite Reihe), Dr. Ahnert (rechts neben Professor Scholz) (Foto: Medizinische Fakultät der Universität Leipzig).

Molekulare Netzwerkanalysen zeigen kausale Zusammenhänge

Aus der PROGRESS Studie[1] zur ambulant erworbenen Pneumonie bei hospitalisierten Patienten und aus anderen Studien liegen umfangreiche molekulare Daten mehrerer Omics-Ebenen vor (Genetik, Transkriptom, Metabolom, Proteom, Zytokine), die teilweise auch im Zeitverlauf verfügbar sind. Ein wichtiger Schwerpunkt unserer Arbeiten in CAPSyS ist die Analyse dieser umfangreichen Daten. Zum einen werden verschiedene diagnostische und prognostische Biomarker-Signaturen entwickelt und validiert. Zum anderen besteht das Ziel, kausale Beziehungen zwischen verschiedenen Merkmalen einer Omics-Ebene und hin zu klinischen Parametern herzustellen. Dies erfolgt sowohl querschnittlich zum Beispiel mittels Methoden der Mendelschen Randomisierung oder durch Mediationsanalysen als auch unter Ausnutzung der erhobenen Zeitreihendaten zum Beispiel über Strukturgleichungsmodelle. Durch die sukzessive Anwendung dieser Methoden lassen sich umfangreiche kausale Netzwerke etablieren (Beispiel Abbildung 2). Ziel dieser Analysen ist die Etablierung relevanter Omics-Beziehungen, die in die mechanistische Modellierung der Krankheitsprozesse einfließen.

[1] http://capnetz.de/html/progress/project

Abbildung 2: Kausale Netzwerkanalysen anhand von Zeitreihendaten in PROGRESS



Abgeleitete kausale Beziehungen zwischen Zytokinen, Blutzellen und klinischen Parametern mit mindestens einer signifikanten Kausalbeziehung zu einem anderen Parameter. Grüne Pfeile: positive Kausalbeziehungen, Rote gestrichelte Pfeile: negative Kausalbeziehungen. Signifikante Kausalbeziehungen sind durch $p < 0{,}001$ im zugehörigen Strukturgleichungsmodell (siehe eingefügtes vereinfachtes Schema) definiert. (Quelle: AG Genetische Statistik und Systembiologie, IMISE, Universität Leipzig)

Biomathematische Modellierung der Barrierefunktion und Therapie

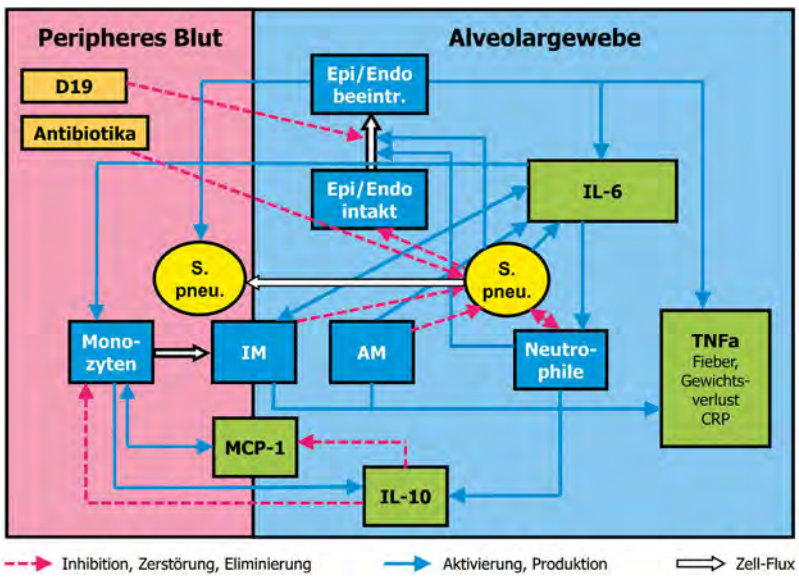
Basierend auf den identifizierten wesentlichen Beziehungen zwischen molekularen und zellulären Faktoren wurde ein biomathematisches Modell der Pneumokokken-Infektionen in Mäusen etabliert. Das Modell beschreibt mittels gewöhnlicher Differentialgleichungen die Entwicklung der bakteriellen Last in der Lunge und der Zirkulation, die Population wichtiger Immunzellen und die Konzentration wesentlicher Zytokine und Chemokine (Proteine, die das Wachstum und die Migration von Zellen regulieren) sowie entsprechende Wechselwirkungen zwischen diesen Komponenten der Immunreaktion. Zudem wurden Mechanismen der Antibiosewirkung berücksichtigt. Das Modell wurde mittels umfangreicher tierexperimenteller Zeitverlaufsdaten vor allem der frühen Phasen einer Infektion entwickelt und parametrisiert (Schirm *et al.*, 2016). Dieses Modell wurde kürzlich grundlegend erweitert, wobei vor allem die Funktion der zwischen Alveolen (Lungenbläschen) und Blutkreislauf liegenden epithelial-endothelialen Barriere genauer beschrieben sowie die Wirkung von barriereschützenden Therapiekonzepten wie zum Beispiel Complement Component 5a-Inaktivierung (C5a-Inhibitor) betrachtet wurde (Abbildung 3). So kann die Wirkung kombinierter Therapien aus Antibiose und barriereschützenden Faktoren simuliert und vorhergesagt werden. Hierbei stellte sich unter anderem heraus, dass eine möglichst frühe Antibiose kombiniert mit einer hohen Dosierung des C5a-Inhibitors zu besonders günstigen Krankheitsverläufen mit deutlich stabilisierter Barrierefunktion und entsprechend geringerer systemischer Inflammation führt.

Modell der Krankheitsschwere beim Menschen

Das in Abbildung 3 gezeigte Modell der Immunreaktion und Therapie einer Pneumokokken-Pneumonie bei Mäusen lässt sich nicht ohne weiteres auf den Menschen übertragen. Zum einen fehlen hier wesentliche für die Parametrisierung notwendige Informationen und Daten vom Ort des Infektionsgeschehens. Zum anderen sind Daten aus der Frühphase einer Infektion beim Menschen nicht erhältlich. In der Regel fehlt auch die Information, wann die Infektion stattgefunden hat, so dass die Patienten selbst bei engmaschiger Beobachtung im Krankenhaus zu unterschiedlichen Zeitpunkten nach Beginn der Infektion beobachtet werden und somit asynchrone Zeitreihen vorliegen. Um dieses Problem zu umgehen, wurde beim Menschen zunächst ein phänomenologisches Modell aufgebaut, welches Übergänge von Krankheitsstadien als zeitliche Zufallsprozesse beschreibt. Dieses sogenannte zeitkontinuierliche Markov-Modell wurde mittels Daten der PROGRESS-Studie sowie Patienten mit pneumogener (also durch Infektion der Atemwege hervorgerufener) Sepsis aus großen multizentrischen klinischen Studien der SepNet-Studiengruppe[2] entwickelt und parametrisiert. Die Krankheitsschwere wurde hierbei mittels des SOFA-Scores (Sequential Organ Failure Assessment) operationalisiert (Ahnert *et al.*, 2019, Abbildung 4). Das Modell zeigt gute Vorhersagen in einem unabhängigen Datensatz bezüglich der 28 Tage-Mortalität der Patienten und soll perspektivisch hinsichtlich der Einbeziehung individueller Risikofaktoren erweitert werden (Przybilla *et al.*, 2020).

[2] https://www.sepsis-stiftung.eu/sepnet

Abbildung 3: Therapiemodell der Pneumokokken-Pneumonie in Mäusen



Schematische Darstellung des weiterentwickelten Modells zur Beschreibung der zellulären und molekularen Immunantwort sowie der Bakterienpopulationen im Alveolargewebe und im peripheren Blut bei Infektion von Mäusen durch S. pneumoniae (S. pneu.) unter dem Einfluss von Antibiotika und Barriere-stabilisierenden Therapeutika (D19). Epi/Endo intakt: unbeeinträchtigte Epi- und Endothelzellen, Epi/Endo beeintr.: beeinträchtigte Epi- und Endothelzellen bei Barrieredefekt, IM: inflammatorische Makrophagen, AM: Alveolarmakrophagen, CRP: C-reaktives Protein. (Quelle: AG Genetische Statistik und Systembiologie, IMISE, Universität Leipzig).

COVID-19

Obwohl CAPSyS die Pneumokokken-assoziierte Pneumonie zum Forschungsschwerpunkt hat, gibt es einige Parallelen zur COVID-19 Erkrankung. In Zusammenarbeit mit Kollegen der Charité – Universitätsmedizin Berlin versuchen wir deshalb, die im Rahmen von CAPSyS entwickelten system-medizinischen Ansätze auf dieses neue Krankheitsbild zu übertragen. Wir nutzen aktuell die PROGRESS und CAPSyS Infrastrukturen, um eine geeignete Datenbasis zu hospitalisierten COVID-19 Patienten aufzubauen. Mittels dieser Datenbasis wollen wir die im vorigen Abschnitt dargestellten Markov-Modelle auf COVID-19 übertragen, um den Verlauf der Erkrankung besser verstehen zu lernen.

Des Weiteren spielt die Barrierefunktion auch bei COVID-19 eine entscheidende Rolle. Wir arbeiten deshalb aktuell an einer Übertragung und Weiterentwicklung des dargestellten, anhand von Daten aus Tierexperimenten kalibrierten Modells des Infektionsgeschehens in der Lunge und der Funktion der endothelial-epithelialen Barriere auf dieses Krankheitsbild. Um diese Arbeiten zu unterstützen, werden durch unsere Partner an der Charité umfangreiche Experimente mit SARS-CoV-2 ganz in Analogie zu den bisher in CAPSyS mit Pneumokokken durchgeführten Experimenten durchgeführt.

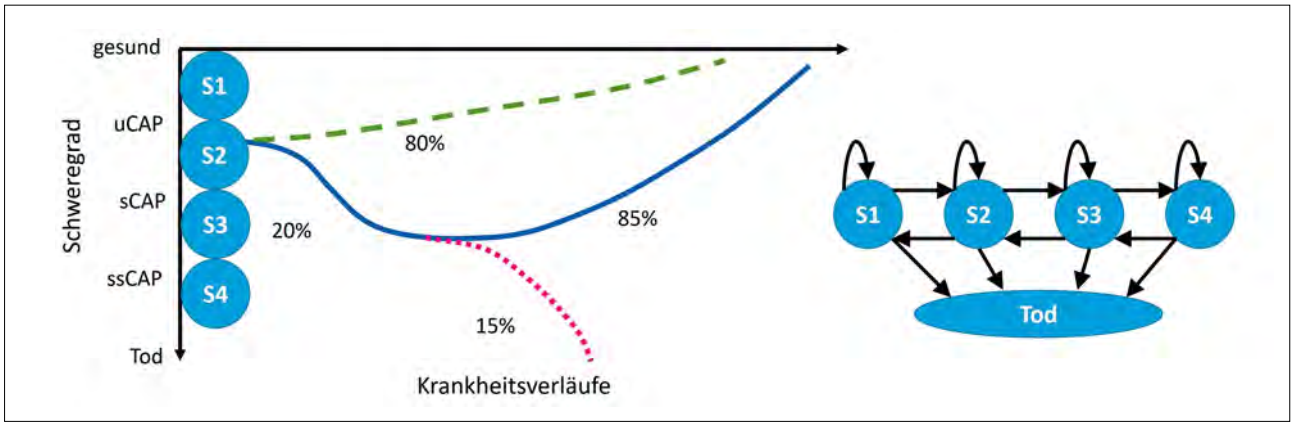
Schließlich beteiligen wir uns auch an der epidemiologischen Modellierung der COVID-19 Pandemie und stützen uns auch hier auf Modelle, die wir im Rahmen anderer Projekte bereits für die epidemiologische Beschreibung von Pneumokokken-Pneumonien aufgebaut haben. Vorhersagen zum Verlauf der Entwicklung der COVID-19 Todesfälle in Deutschland, die auf einer ersten Version dieses Modells beruhen, stellen wir bereits der Öffentlichkeit zur Verfügung. Dies geschieht über eine Plattform, auf welcher die Vorhersagen verschiedener Modellierungsgruppen gesammelt und gemeinsam sichtbar gemacht werden (https://jobrac.shinyapps.io/app_forecasts_de).

Wir hoffen, mit unseren Aktivitäten im Rahmen von CAPSyS einen systemmedizinischen Beitrag sowohl zur Entwicklung verbesserter Behandlungsmöglichkeiten als auch zum besseren Verständnis der Pathomechanismen und der Ausbreitung von COVID-19 zu leisten.

Spätfolgen

Über die Spätfolgen herkömmlicher und durch SARS-CoV-2 ausgelöster Pneumonien ist aktuell wenig bekannt. Bei herkömmlichen Pneumonien wurde auf Basis epidemiologischer Daten unter anderem eine Zunahme von Atherosklerose und damit assoziierter Ereignisse beobachtet. Die zugrundeliegenden Pathomechanismen sind weitgehend unbekannt. Um diese systemmedizinisch zu erforschen, wurde kürzlich das e:Med-Verbundprojekt „Systemmedizin der Pneumonie-aggravierten Atherosklerose“ (SYMPATH) unter Leitung der Charité (Martin Witzenrath) und der Universität Leipzig (Markus Scholz) initiiert (https://www.sys-med.de/de/verbuende/sympath). Auch hier erfolgen parallele Arbeiten in Maus und Mensch, um die zugrundeliegenden molekularen Pathomechanismen aufzuklären und zu modellieren. Eine Übertragung des Ansatzes auf COVID-19 ist geplant.

Abbildung 4: Markov-Modell der Therapieverläufe bei ambulant erworbener Pneumonie



Beispielhafte Krankheitsverläufe von einer leichten ambulant erworbenen Pneumonie (uCAP) zu einer schweren CAP (sCAP) oder einer schweren septischen CAP mit hoher Letalität (ssCAP). Für das Markov-Modell (Schema rechts) wurden vier durch den SOFA-Score definierte Krankheitszustände (S1-S4) angenommen, welche vereinfachend sequentiell durchlaufen werden können. Der Endzustand Tod hingegen kann aus jedem Krankheitszustand heraus direkt erreicht werden. Die Übergänge im Modell erfolgen zeitkontinuierlich. (Quelle: AG Genetische Statistik und Systembiologie, IMISE, Universität Leipzig).

Steckbrief Forschungsprojekt:

Der Forschungsverbund „CAPSyS – Systemmedizin der ambulant erworbenen Pneumonie“ wurde im Jahre 2014 im Rahmen des Forschungs- und Förderkonzeptes „e:Med – Maßnahmen zur Etablierung der Systemmedizin“ des Bundesministeriums für Bildung und Forschung gegründet. Ziel des Forschungsverbundes ist die systembiologische Analyse des Krankheitsverlaufes von Pneumokokken-assoziierten Pneumonien. Speziell steht die Frage im Vordergrund, wie sich der Krankheitsverlauf durch klinische, zelluläre und molekulare Faktoren beschreiben und vorhersehen lässt. Im Verbund sind sieben Einrichtungen vereint: Universität Leipzig (Institut für Medizinische Informatik, Statistik und Epidemiologie, IMISE), Charité – Universitätsmedizin Berlin (Medizinische Klinik mit Schwerpunkt Infektiologie und Pneumologie), Universitätsklinikum Erlangen (Labor für System-Tumorimmunologie), Justus-Liebig-Universität Gießen (Institut für Medizinische Mikrobiologie), Ernst-Moritz-Arndt-Universität Greifswald (Abteilung für funktionelle Genomforschung), Universitätsklinikum Jena (Institut für Klinische Chemie und Laboratoriumsdiagnostik) und die Philipps-Universität Marburg (Institut für Lungenforschung). Bisher konnten die Ergebnisse der Arbeit des Verbundes in 57 Artikeln in internationalen Fachzeitschriften veröffentlicht werden.

CAPSyS
Systems Medicine of Community Acquired Pneumonia

Referenzen:

Ahnert, P., Creutz, P., Horn, K., Schwarzenberger, F., Kiehntopf, M., Hossain, H., Bauer, M., Brunkhorst, F.M., Reinhart, K., Völker, U., *et al.* (2019). Sequential organ failure assessment

score is an excellent operationalization of disease severity of adult patients with hospitalized community acquired pneumonia - results from the prospective observational PROGRESS study. Critical care (London, England) 23, 110.
Michaud, C.M. (2009). Global Burden of Infectious Diseases. Encyclopedia of Microbiology, 444-454.
Przybilla, J., Ahnert, P., Bogatsch, H., Bloos, F., Brunkhorst, F.M., SepNet, C.C.T.G., Progress, S.G., Bauer, M., Loeffler, M., Witzenrath, M., *et al.* (2020). Markov State Modelling of Disease Courses and Mortality Risks of Patients with Community-Acquired Pneumonia. Journal of clinical medicine 9.
Schirm, S., Ahnert, P., Wienhold, S., Mueller-Redetzky, H., Nouailles-Kursar, G., Loeffler, M., Witzenrath, M., and Scholz, M. (2016). A Biomathematical Model of Pneumococcal Lung Infection and Antibiotic Treatment in Mice. PloS one 11, e0156047.
Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., *et al.* (2020). Clinical course and risk factors for mortality of adult in patients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet.

Kontakt:
Prof. Dr. Markus Scholz
AG Genetische Statistik und Systembiologie
Institut für Medizinische Informatik, Statistik und Epidemiologie
Medizinische Fakultät
Universität Leipzig
markus.scholz@imise.uni-leipzig.de

https://www.genstat.imise.uni-leipzig.de
https://www.capsys.imise.uni-leipzig.de

de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

The German Network for Bioinformatics Infrastructure offers training events for life scientists

Online Training

Basis bioinformatics training for biologists
Analysis, Vsualization and Integration of Multi-Level Omics Data

Introduction to Oxford Nanopore sequence data analysis
Nanopore Workshop - Best Practice and SARS-CoV-2 Applications

Introduction to metagenome data analysis
Metagenomics Bioinformatics

Introduction to RNA-seq data analysis
RNAseq and High Throughput Omics Analysis for Plants

Advanced methods for differential analysis of proteomics data
Advanced Analysis of Auantitative Proteomics Data Using R

Primary means to store biological data
Ontologies - Statistics, Biases, Tools, Networks and Interpretation

Analyzing scientific data using machine learning algorithms
Machine Learning using Galaxy

Introduction to the de.NBI Cloud
de.NBI Cloud User Meeting

And many more at:
www.denbi.de/training

linkedin.com/company/de-nbi
contact@denbi.de
@denbiOffice

SPONSORED BY THE
Federal Ministry of Education and Research
PTJ Projektträger Jülich Forschungszentrum Jülich

Von Daten zu Wissen – Standards für die personalisierte Medizin

EU-STANDS4PM – ein Europäisches Expertenforum zur Entwicklung von Standards für computerbasierte Modellierungen in der personalisierten Medizin

von Dr. Marc Kirschner, Projektträger Jülich

Die systematische Analyse und Interpretation großer, komplexer Datensätze (Big Data) aus verschiedenen Anwendungsfeldern hat das Potential, die Diagnose und Behandlung von Erkrankungen deutlich zu verbessern. Hierzu zählen beispielsweise Daten aus den Biowissenschaften, dem Gesundheitswesen oder der klinischen Forschung. Besonders in der personalisierten Medizin können datengetriebene Verfahren mit Hilfe von computerbasierten Modellen einen wesentlichen Beitrag zu Früherkennung, Prävention und Vorhersagen über den Therapieerfolg von Erkrankungen leisten.

Der Prozess der Erzeugung von neuem, medizinischen Wissen liegt jedoch weit hinter dem möglichen Potential zurück. Gründe hierfür sind die Heterogenität von Big Data sowie das Fehlen breit akzeptierter Standards zur Datenerhebung, -harmonisierung und -integration. Das Arbeiten mit personen- und patientenbezogenen Daten stellt zudem hohe ethisch-rechtliche Anforderungen hinsichtlich der grundlegenden Rechte von Patienten auf Information und Datenschutz.

Breit anwendbare und ethisch-rechtlich konforme Standardisierungsrichtlinien stellen zukünftig somit eine zentrale Komponente für eine personalisierte Medizin im Bereich der computerbasierten Modellierung dar. Sie sind eine Grundvoraussetzung für maßgeschneiderte Behandlungskonzepte, spezifische Früherkennung und Präventionsmaßnahmen.

Um die Weiterentwicklung der Standardisierung zu beschleunigen und einen nachhaltigen Beitrag zur Etablierung transnationaler Richtlinien für datengetriebene Modellierungsmethoden in der personalisierten Medizin zu leisten, wurde

EU-STANDS4PM
standards for in silico models for personalised medicine

das Europäisches Expertenforum, EU-STANDS4PM *(A European standardization framework for data integration and data-driven in silico models for personalized medicine)*, etabliert.

EU-STANDS4PM ist eine seit Januar 2019 unter Horizon2020 der EU-Kommission geförderte *Coordination and Support Action* mit dem Ziel, die Nutzung von Big Data für die personalisierte Medizin weiter voranzutreiben. Das EU-STANDS4PM-Konsortium mit 16 Partnern aus acht europäischen Ländern umfasst fachübergreifende Expertise aus europäischen Wissenschaftsorganisationen, Industrie, ESFRI-Infrastrukturmaßnahmen, Recht- und Ethikwissenschaften sowie Standardisierungsorganisationen.

Kernaktivitäten innerhalb der dreijährigen Projektlaufzeit von EU-STANDS4PM sind die Harmonisierung von Datenintegrationsstrategien in der medizinischen Forschung und Praxis in Europa sowie die Erarbeitung flexibler Standardisierungsrichtlinien für europäische Forschungsverbünde. Weiterhin sollen datengetriebene prädikative Modellierungsstrategien in der personalisierten Medizin gezielt gestärkt werden.

EU-STANDS4PM ist ein offenes Netzwerk für alle, die ein Interesse daran haben, datenbasierte Modellierungen in der personalisierten Medizin mittels Standards voranzubringen und zu unterstützen.

Weitere Informationen unter:
www.eu-stands4pm.eu

MTZ®-Award for Medical Systems Biology 2020

Die diesjährigen Gewinner des *MTZ®-Awards for Medical Systems Biology* stehen fest. **Herr Dr. Fabian Fröhlich** (Technische Universität München), **Frau Dr. Carolin Loos** (Technische Universität München) und **Herr Dr. Martin Scharm** (Universität Rostock) konnten das nationale Gutachtergremium und den Vorstand der MTZ®stiftung mit ihren exzellenten Doktorarbeiten überzeugen und sich gegen die Konkurrenz durchsetzen. Mit dem Preis zeichnet die MTZ®stiftung junge Wissenschaftlerinnen und Wissenschaftler aus, die in ihrer Doktorarbeit bahnbrechende und herausragende Ergebnisse auf dem Gebiet der medizinisch orientierten Systembiologie erzielt haben. Der Förderpreis soll dem vielversprechenden wissenschaftlichen Nachwuchs besondere Sichtbarkeit und öffentliche Anerkennung verschaffen. Hierzu arbeitet die MTZ®stiftung mit dem Bundesministerium für Bildung und Forschung (BMBF) sowie dem Projektträger Jülich (PtJ) zusammen.

MTZ® stiftung

– for a better future –

Die Verleihung des Preises, bestehend aus einer Urkunde und einem Preisgeld, erfolgt schon zum siebten Mal und soll in einem feierlichen Rahmen bei der 8. Internationalen Konferenz „Systems Biology of Mammalian Cells“ (SBMC2020) übergeben werden. Die Konferenz und die feierliche Preisverleihung mussten wegen der Einschränkungen aufgrund der Corona-Pandemie verschoben werden.

Weitere Informationen unter:

www.mtzstiftung.de

Quelle: Projektträger Jülich

events

COMBINE Online Forum 2020

05. Oktober – 09. Oktober 2020

von Dagmar Waltemath

Foto: Till Juncker / Universität Greifswald

Prof. Dagmar Waltemath,
Vice-Chair COMBINE

Die COVID-19-Pandemie hat viele von uns in den virtuellen Raum gezwungen. Neben Einschränken eröffnet diese unfreiwillige Digitalisierung aber auch viele neue Möglichkeiten.

So konnte das 11. COMBINE Forum erstmals als Online Forum realisiert werden und damit ein deutlich breiteres Publikum erreicht werden. Der Teilnahmerekord beim zehnjährigen Jubiläum des COMBINE Forums in 2019 (siehe auch Waltemath *et al.*, 2020)[1] konnte dieses Jahr online sogar verdreifacht werden. Die Idee, das jährliche Treffen der Standardisierungs-Community für Modellierung in der Computer-gestützten Biology (***CO**mputational **M**odeling in **BI**ology **NE**twork,* http://co.mbine.org) zu virtualisieren stand schon längere Zeit im Raum, da die Verteilung der Mitglieder*innen über alle Kontinente neben Umweltaspekten auch erhebliche Zeit- und monetäre Aufwände nach sich zog. Schon in den vergangenen Jahren wurden Vorträge aufgezeichnet und fanden Diskussionen teilweise hybrid statt (vor Ort mit Zuschaltung von Teilnehmer*innen per Videokonferenz). Durch COVID-19 wurden die Organisator*innen nun zum nächsten Schritt gedrängt - und somit begannen im Mai 2020 die Planungen für eine Online-Konferenz.

Die virtuelle COMBINE 2020 fand vom 05.-09. Oktober 2020 als fünftägige Konferenz mit täglichem 24-Stunden-Programm statt. Somit war es allen Teilnehmenden unabhängig von deren Zeitzone möglich an einem Teil der 7 Keynotes, 41 Vorträge, 10 Lightning Talks, 14 Tutorials und 22 fachspezifischen Diskussionsrunden teilzunehmen. Die ersten beiden Tage mit Keynotes und Fachvorträgen wurden - teilweise mit Live-Fragen, teilweise mit der aufgezeichneten Diskussion- zeitversetzt wiederholt, um ein breiteres Publikum anzusprechen. Der Fokus der folgenden Tage lag auf Tutorials und Breakouts. Mit 300 Teilnehmer*innen hat das diesjährige COMBINE Forum so viele Wissenschaftler*innen angezogen wie niemals zuvor. Auch die Themenvielfalt war beachtlich: sie reichte von spezifischen Erweiterungen der Standards, über Anwendungsvorträge von Biologen, Medizinern, und Modellierern, bis hin zu Industrie-Vorträgen zu Standardisierung. Chris Myers, Chair des COMBINE Koordinations-Boards fasst das diesjährige Community-Treffen mit folgenden Worten zusammen: ***While reproducibility remains a challenge in computational modeling in biology, the COMBINE community and the standards being developed within this community have the potential to make this a fully reproducible scientific endeavor.*** Das vollständige Programm ist auf der COMBINE-Webseite einsehbar (http://co.mbine.org/events/COMBINE_2020).

[1] DOI: https://doi.org/10.1515/jib-2020-0005

Steckbrief COMBINE Netzwerk:

Das COMBINE Netzwerk ist ein Zusammenschluss von Forscher*innen und Softwareingenieur*innen, die an der Entwicklung gemeinsamer Standards und Formate für die rechnergestützte Modellierung und Simulation in der System- und Synthetischen Biologie beteiligt sind. Das Netzwerk wurde 2009 mit dem Ziel gegründet verschiedene Standardisierungsbemühungen, welche ähnliche Zwecke verfolgten, zusammenzuführen. Seither finden jährlich 2 Veranstaltungen statt, die hier vorgestellte COMBINE, zur Diskussion der weiteren Entwicklungsrichtungen der Standards und HARMONY, welche sich auf die praktische Entwicklung der Standards, sowie auf Interoperabilität und Infrastruktur konzentriert (http://co.mbine.org/events).

Chair: Chris Myers, Prof. für Elektro-, Computer- und Energietechnik an der Universität von Colorado
Vice Chair: Dagmar Waltemath, Prof. für Medizinische Informatik an der medizinischen Universität Greifswald
COMBINE Koordinatoren: Gary Bader, Padraig Gleeson, Martin Golebiewski, Thomas Gorchowski, Sarah Keating, Matthias König, David Nickerson und Falk Schreiber

impressum

gesundhyte.de

Das Magazin für Digitale Gesundheit in Deutschland – Ausgabe 13, Dezember 2020

gesundhyte.de ist ein jährlich erscheinendes Magazin mit Informationen aus der deutschen Forschung im Bereich digitale Medizin.

ISSN 2702-2544

Herausgeber:

gesundhyte.de wird herausgegeben vom Berlin Institute of Health (BIH) und dem Projektträger Jülich.

Redaktion:

Chefredakteur: Prof. Dr. Roland Eils (BIH/Charité – Universitätsmedizin Berlin)
Redaktionelle Koordination: Franziska Müller (BIH/Charité – Universitätsmedizin Berlin)
Redaktion:
Dr. Silke Argo (e:Med), Dr. Stefanie Gehring (DLR-PT), Prof. Dr. Lars Kaderali (Universität Greifswald), Dr. Marco Leuer (DLR-PT), Dr. Yvonne Pfeiffenschneider (PtJ), Dr. Matthieu-P. Schapranow (Hasso-Plattner-Institut), Dr. Thomas Schmidt (Plattform Lernende Systeme), Dr. Stefanie Seltmann (BIH) und Dr. Gesa Terstiege (PtJ).

Anschrift:
Redaktion gesundhyte.de
c/o BIH/Charité – Universitätsmedizin Berlin
Zentrum Digitale Gesundheit
Kapelle-Ufer-2; D-10117 Berlin

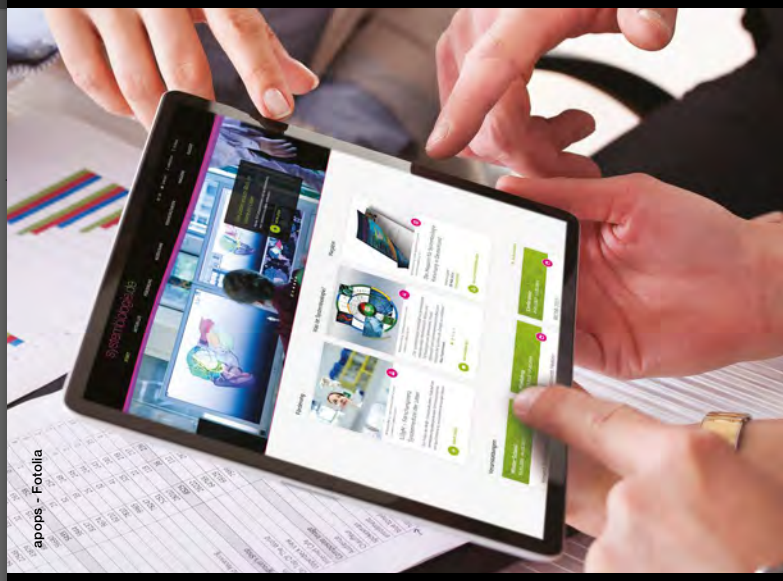
Gestalterische Konzeption und Umsetzung:
Kai Ludwig, LANGEundPFLANZ Werbeagentur GmbH, Speyer (www.LPsp.de)
Übersetzungen:
Todd Brown, Berlin, Deutschland
Druck:
Ottweiler Druckerei und Verlag GmbH, Ottweiler; Deutschland

Aboservice:
Das Magazin wird aus Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) gefördert. Diese Veröffentlichung ist Teil der Öffentlichkeitsarbeit der Herausgeber. Sie wird kostenlos abgegeben und ist nicht zum Verkauf bestimmt.

Wenn Sie das Magazin abonnieren möchten, füllen Sie bitte das Formular auf www.gesundhyte.de aus oder wenden sich an:

Redaktion gesundhyte.de
c/o BIH/Charité – Universitätsmedizin Berlin
Zentrum Digitale Gesundheit
Kapelle-Ufer-2; D-10117 Berlin
franziska.mueller@charite.de

Aus systembiologie.de wird gesundhyte.de!



Wenn Sie mehr über die Themen der Zeitschrift erfahren möchten, besuchen Sie unsere Homepage.

Das erwartet Sie:

- Spannende Geschichten aus dem Forschungsalltag – Erfahren Sie mehr über aktuelle Projekte
- Interviews und Portraits – Lernen Sie die Gesichter hinter der Forschung kennen
- Umfassende Veranstaltungsübersicht – Verpassen Sie keinen wichtigen Termin
- Informationen über aktuelle Fördermaßnahmen – Bleiben Sie stets auf dem Laufenden
- Aktive Mitgestaltung – Schlagen Sie uns Ihr Thema vor

Wir freuen uns auf Ihren Besuch auf unserer Homepage!

wir über uns

die gesundhyte.de-Redaktion stellt sich vor

gesundhyte.de möchte die Erfolge der deutschen Forschung auf anschauliche Weise einem breiten Publikum zugänglich machen. Erstellt wird das einmal jährlich auf Deutsch und Englisch erscheinende Magazin gemeinsam von einem multidisziplinären Redaktionsteam unterschiedlicher Institutionen der deutschen Forschung: Berlin Institute of Health, Charité – Universitätsmedizin Berlin, Hasso-Plattner-Institut Potsdam, Universität Greifswald, Projektträger Jülich, DLR Projektträger und Vertretern der Initiativen: Lernende Systeme – Die Plattform für Künstliche Intelligenz und e.med Systems Medicine. Finanziert wird das Magazin vom Berlin Institute of Health (BIH) und dem Bundesministerium für Bildung und Forschung (BMBF).

Die Redaktionsmitglieder von gesundhyte.de:

v.l.n.r. erste Reihe: Prof. Dr. Roland Eils (BIH/Charité – Universitätsmedizin Berlin), Franziska Müller (BIH/Charité – Universitätsmedizin Berlin), Dr. Silke Argo (e:Med), Dr. Stefanie Gehring (DLR-PT), **v.l.n.r. zweite Reihe:** Prof. Dr. Lars Kaderali (Universität Greifswald), Dr. Marco Leuer (DLR-PT), Dr. Yvonne Pfeiffenschneider (PtJ), Dr. Matthieu-P. Schapranow (Hasso-Plattner-Institut), **v.l.n.r. dritte Reihe:** Dr. Thomas Schmidt (Plattform Lernende Systeme), Dr. Stefanie Seltmann (BIH) und Dr. Gesa Terstiege (PtJ), Kai Ludwig (LANGEundPFLANZ)

Fotos: shutterstock_© RAY-BON, © David Ausserhofer, privat / Müller, privat / Argo, privat / Gehring, privat / Kaderali, privat / Leuer, privat / Pfeiffenschneider, privat / Schapranow, © Thilo Schoch, © Thomas Rafalzyk, privat / Terstiege, privat / Ludwig

kontakt

Berlin Institute of Health/Charité – Universitätsmedizin Berlin
Chefredakteur: Prof. Dr. Roland Eils
Redaktionelle Koordination:
Franziska Müller
c/o BIH/Charité – Universitätsmedizin Berlin
Zentrum für Digitale Gesundheit
Kapelle-Ufer-2; D-10117 Berlin
E-Mail: franziska.mueller@charite.de
www.hidih.org

Berlin Institute of Health
Kommunikation und Marketing, Pressestelle
Ansprechpartner:
Dr. Stefanie Seltmann
Anna-Louisa-Karsch-Str. 2; D-10178 Berlin
E-Mail: s.seltmann@bihealth.de
www.bihealth.org

Lernende Systeme – Die Plattform für Künstliche Intelligenz
Ansprechpartner:
Dr. Thomas Schmidt und Dr. Ursula Ohliger
Geschäftsstelle: c/o acatech
Pariser Platz 4a; D-10117 Berlin
E-Mail: t.schmidt@acatech.de, ohliger@acatech.de
www.plattform-lernende-systeme.de

Projektträger Jülich
Forschungszentrum Jülich GmbH
Lebenswissenschaften und Gesundheitsforschung (LGF)
Ansprechpartner:
Dr. Yvonne Pfeiffenschneider und Dr. Gesa Terstiege
D-52425 Jülich
E-Mail: y.pfeiffenschneider@fz-juelich.de, g.terstiege@fz-juelich.de
www.ptj.de

DLR Projektträger
Gesundheitsforschung (OE20)
Ansprechpartner:
Dr. Marco Leuer und Dr. Stefanie Gehring
Heinrich-Konen-Str. 1; D-53227 Bonn
E-Mail: marco.leuer@dlr.de, stefanie.gehring@dlr.de
www.dlr-pt.de

Geschäftsstelle des e:Med Projektkomitees
Ansprechpartner Leitung:
Dr. Silke Argo
c/o Deutsches Krebsforschungszentrum (DKFZ) – V025
Im Neuenheimer Feld 581; D-69120 Heidelberg
E-Mail: s.argo@dkfz.de
www.sys-med.de

systembiologie.de

DAS MAGAZIN FÜR SYSTEMBIOLOGISCHE FORSCHUNG IN DEUTSCHLAND AUSGABE 13 DEZEMBER 2020

spezial: künstliche intelligenz

ab Seite 8

potenziale und herausforderungen von KI in der medizin

Seite 15

big data und smartphones auf der intensivstation

Seite 30

interviews mit Petra Ritter Janine Felden

Seite 52 und 72

aus systembiologie.de wird gesundhyte.de

DAS MAGAZIN FÜR DIGITALE GESUNDHEIT IN DEUTSCHLAND