gesundhyte.de

DAS MAGAZIN FÜR DIGITALE GESUNDHEIT IN DEUTSCHLAND

AUSGABE 14 DEZEMBER 2021

fokus: forschungsdaten helfen heilen

ab Seite 14

zwischen individualisierten therapieoptionen und gläsernen patienten

Seite 32

SmICS – eine smarte lösung zur automatisierten ausbruchserkennung im krankenhaus

Seite 46

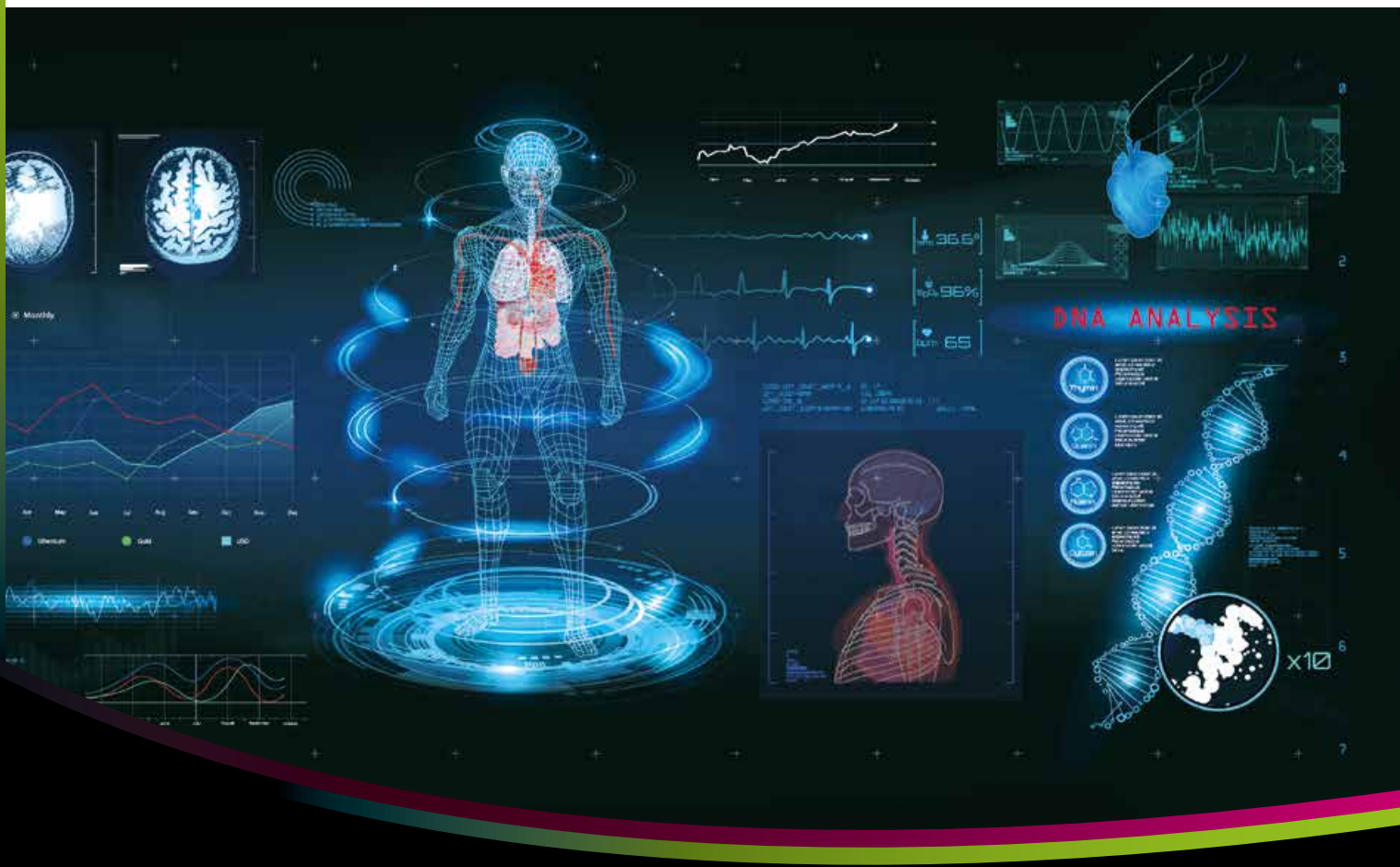
interviews mit Martin Hager Dana Westphal und Michael Seifert Simon Haas

Seite 29, 73 und 82

ISSN 2702-2552

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung



gesundhyte.de

widmet sich der Suche nach Lösungen für die wichtigsten Herausforderungen des Gesundheitswesens von heute und konzentriert sich dabei auf Digitale Gesundheit und Systemmedizin, zwei junge Disziplinen, die als die Medizin der Zukunft gelten. Der Schlüssel liegt in der Kombination von Forschungsdaten aus dem Labor mit Daten aus der Versorgung – vom Labor bis zum Krankenbett. Innovative Technologien und Methoden werden den Weg für präzisere Vorhersagen und personalisierte Therapien ebnen. Dieser Ansatz wird bereits heute in der Onkologie erfolgreich eingesetzt und soll in Zukunft auf andere Krankheiten ausgeweitet werden. Dabei spielt der Aufbau einer robusten und standardisierten IT-Infrastruktur eine große Rolle, um den sicheren Austausch von Patientendaten zwischen Forschungsteams zu ermöglichen. Eine solche Infrastruktur wird auch helfen, den dringend benötigten effizienten Datenaustausch zwischen Kliniken und ambulanter Versorgung zu verbessern. Lesen Sie im Magazin gesundhyte.de, wie diese innovativen Wissenschaftszweige an Lösungen für unsere aktuellen und zukünftigen Herausforderungen in der Medizin arbeiten.

Titelbild: iStock © Lan Zhang

Foto: AdobeStock_© SergeyBitos

grußwort

Liebe Leserinnen und Leser,

nach fast zwei Jahren COVID-19-Pandemie blicken wir zurück auf eine intensive Zeit, in der gerade die Forschung Bahnbrechendes erreicht hat. Die Bedeutung und Notwendigkeit leistungsfähiger Forschung und medizinischer Versorgung ist für die Breite der Bevölkerung ins Zentrum des Interesses gerückt und sichtbar geworden. Entscheidend für die Erfolge war dabei auch der rasche Zugang zu zuverlässigen Daten und ihr Austausch zwischen Forschung und Versorgung. Umfassende Datensätze rund um das neue Virus mussten in kürzester Zeit erhoben, miteinander geteilt und analysiert werden.

Für diese Aufgabe konnte die wissenschaftliche Gemeinschaft auf Infrastrukturen zurückgreifen, deren Aufbau das Bundesministerium für Bildung und Forschung (BMBF) in den letzten Jahren vorangetrieben hat. So wurden die bioinformatischen Werkzeuge und die Rechen-Cloud des deutschen Netzwerks für Bioinformatik-Infrastruktur (de.NBI) bundesweit von Forscherinnen und Forschern genutzt, um das Virus besser zu verstehen und seine Wirkmechanismen aufzuklären.

Auch für die Verknüpfung medizinischer Daten hat das BMBF Strukturen geschaffen. Kernelement der seit 2018 vom BMBF geförderten Medizininformatik-Initiative sind die Datenintegrationszentren, die übergreifende Analysen verschiedener Universitätskliniken ermöglichen. Zur Bewältigung von COVID-19 hat das BMBF die Vernetzung der Universitätskliniken mit der Förderung des Nationalen Netzwerks der Universitätsmedizin (NUM) noch einmal intensiviert. Zentraler Baustein ist dabei die Forschungsdatenplattform CODEX zur Bereitstellung von Forschungsdaten zu COVID-19 für alle Universitätskliniken. Die umfassenden Entwicklungen der Medizininformatik-Initiative, insbesondere die der Datenintegrationszentren, sind dabei eine entscheidende Grundlage.

In der aktuellen Ausgabe von gesundhyte.de erhalten Sie vielfältige Einblicke in Projekte aus diesen und anderen Maßnahmen. Die Ergebnisse zeigen das große Potenzial von Forschungsdaten und ihrer Analyse in den Lebenswissenschaften und der Medizin. Ich wünsche Ihnen allen eine anregende Lektüre.

Prof. Dr. Veronika von Messling
Abteilungsleiterin Lebenswissenschaften
Bundesministerium für Bildung und Forschung

Foto: Prof. Dr. Veronika von Messling

grußwort

Liebe Leserinnen und Leser,

ein weiteres Jahr der Pandemie hat uns gezeigt, wie essentiell die Verknüpfung von Versorgungsdaten mit der Forschung ist. Mit der Integration des Berliner Instituts für Gesundheitsforschung, dem BIH, in die Charité sind wir unserem Ziel, die translationale Forschung zu stärken, einen großen Schritt nähergekommen. Unter dem Motto „Wir gehören zusammen!“ wurde das BIH zum 1. Januar 2021 als Translationsforschungsbereich eingegliedert und bildet nun neben Klinikum und Medizinischer Fakultät die dritte Säule der Charité.

Bereits während der Pandemie konnte sich die Zusammenarbeit von Charité und BIH beweisen: Dank der Nähe des BIH zu unserer Klinik konnten wir schnell reagieren und viele Daten aus der Krankenversorgung direkt in die Forschung fließen lassen. In der letzten Zeit wurde auf Hochtouren an wissenschaftlichen und medizinischen Lösungen geforscht, um die Pandemie zu bekämpfen, auch an der Charité, und wir leisten in vielen Kooperationen unseren Beitrag zur Bekämpfung von Covid-19. Die Pandemiesituation ist ein Paradebeispiel dafür, wie wichtig die enge Verzahnung des Klinik- und Forschungsalltags ist. Aus der Krise hat sich unter der Koordination der Charité ein sehr starkes und eng zusammenarbeitendes Netzwerk, das Netzwerk Universitätsmedizin (NUM), gebildet. Wir sind zuversichtlich, dass dieser Zusammenschluss aller Universitätskliniken des Landes in Zukunft weiterhin Erfolge hervorbringen wird.

Dieser neue Translationsforschungsbereich kommt auf längere Sicht vor allem unseren Patientinnen und Patienten zugute. Deshalb freue ich mich darauf, gemeinsam mit dem BIH die Umsetzung von Forschungsergebnissen in die klinische Anwendung für unsere Patientinnen und Patienten weiter zu forcieren. Dabei müssen die Synergien, die sich nun zwischen Charité und BIH ergeben, konstruktiv und zielgerichtet genutzt werden. Daher ist die Integration des BIH in die Charité von so großer Bedeutung.

In dieser Ausgabe der gesundhyte.de erwartet Sie eine Vielzahl von Beispielen für vielversprechende translationale Forschung. Ich wünsche Ihnen viel Vergnügen beim Studium der einzelnen Artikel.

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

Foto: Wiebke Peitz / Charité

vorwort

Liebe Leserinnen und Leser,

„Der Ball ist rund und ein Spiel dauert 90 Minuten." Dieser berühmte Ausspruch von Sepp Herberger beschreibt ein Faktum. Fakten basieren auf Daten und beschreiben eine unbestreitbare Tatsache. Fakten können aber auch einen allgemein anerkannten Sachverhalt beschreiben. Was wahr oder falsch sein kann, wird zu Tatsachenbehauptungen.

Der berüchtigte Ausspruch von Trumps Beraterin Kellyanne Conway, die Trumps damaligen Sprecher gegen den Vorwurf der Lüge in Schutz nahm, indem sie äußerte, dass er nicht gelogen, sondern alternative Fakten angeboten hätte, war die traurige Geburtsstunde einer Ära, in der Tatsachenbehauptungen gerne auch datenfrei in den Raum gestellt werden. Es ist wohlbekannt, dass solche Tatsachenbehauptungen sehr viel schneller in den sozialen Medien viral gehen als bisweilen als trocken oder langweilig empfundene, datenbasierte Schlussfolgerungen. *„Wenn 50 Millionen Menschen etwas Dummes sagen, bleibt es trotzdem eine Dummheit"*, sagte der berühmte französische Schriftsteller Anatole France schon vor einem Jahrhundert, bevor dieser Ausspruch eine ganz neue Bedeutung in Zeiten sozialer Medien erlang.

Das Ermitteln von Tatsachen mittels Erhebung von Daten gehört zum Kerngeschäft der Wissenschaften. Daten liefern die Grundlage für Prognosen, Einschätzungen und Diagnosen auch im Gesundheitsbereich. Gerade in Zeiten, wo datenfreie Tatsachenbehauptungen gerne in die Welt gestellt werden und millionenfach repliziert werden, ist es umso wichtiger uns alle immer wieder daran zu erinnern, dass Wissen und Fakten datenbasiert sein müssen. Verstörend wird dabei aber in der Öffentlichkeit wahrgenommen, dass Tatsachenbehauptungen auch in der Wissenschaft Änderungen unterworfen sind. Etwas, was gestern noch als „wahr" empfunden wurde, erscheint morgen mitunter als „falsch". Die Wissenschaft war schon immer mit Irrtümern und späteren Revisionen konfrontiert. Nicht nur in der kurzen Historie der jetzigen Pandemie gibt es zahlreiche Beispiele, wie eine wechselnde Datenlage neue Erkenntnisse erzeugt hat.

Daten spiegeln eine Momentaufnahme der Beschreibung der Wirklichkeit wider, die Datenlandschaft entwickelt sich stetig und dynamisch. Grundlage einer datenzentrierten Wissenschaft sind Plattformen begleitet von Rahmenbedingungen, die der wissenschaftlichen Gemeinschaft einen breiten Zugang zu Daten und deren bestmöglichen Nutzung erlauben. Hierzu gibt es faszinierende Einblicke in der Ihnen vorliegenden Ausgabe. Dass Wissenschaftler*innen auch einen emotionalen Zugang zu Daten haben können, zeigt das Porträt von Claudia Langenberg. *„Wir lieben Daten"*, eine treffsichere Charakterisierung ihres Arbeitsgebietes in nur drei Worten und zugleich auch hervorragende Zusammenfassung dieser Ausgabe von gesundhyte.de.

Ich wünsche Ihnen viel Vergnügen und zahlreiche Erkenntnisgewinne bei der Lektüre der verbleibenden 36.594 Worte.

Bleiben Sie uns und den Daten gewogen!
Ihr

Roland Eils
Chefredakteur gesundhyte.de

Foto: David Ausserhofer / BIH

inhalt

grußwort 3
Prof. Dr. Veronika von Messling, Abteilungsleiterin Lebenswissenschaften, Bundesministerium 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

„das leben ist unvorhersehbar!“ 8
Porträt: Barbara Di Ventura
von Katharina Kalhoff und Gesa Terstiege

wie habe ich die pandemie erlebt? 11
1,5 Jahre Corona aus der Sicht eines Nachwuchswissenschaftlers
von Sören Lukassen

fokus: forschungsdaten helfen heilen

besserer zugriff auf klinisch relevante daten 14
BMBF-Fördermaßnahme „i:DSem“ zieht ein positives Resümee
von Björn Dreesen-Daun, Markus Löffler, Stephan Kiefer, Juliane Fluck, Mischa Ubachs und Maksims Fiosins

nutzung von klinischen routinedaten für forschung und versorgung 21
Voraussetzung: Der digitale modulare Broad Consent
von Clemens Spitzenpfeil, Alexandra Stein, Martin Bialke, Torsten Leddig und Wolfgang Hoffmann

lost in translation? semantische standardisierung! 25
Terminologien wie LOINC und SNOMED CT reduzieren Babylonische Sprachverwirrung
von Cora Drenkhahn und Josef Ingenerf

eine forschende versorgung wäre längst möglich 29
Interview und Firmenporträt:
Martin Hager, Roche Pharma AG

zwischen individualisierten therapieoptionen und gläsernen patienten 32
Wie Patientenvertretungen Chancen und Herausforderungen von KI in der Medizin bewerten
von Klemens Budde, Matthieu Schapranow, Elsa Kirchner und Thomas Zahn

innovative software-werkzeuge für die lebenswissenschaften 37
BMBF-Initiative CompLS unterstützt die Entwicklung rechnergestützter Methoden und Analysewerkzeugen mit 34 Millionen Euro
von Daniel Heinrichs, René Wolf-Eulenfeld, Bernhard Renard, Jakub Bartoszewicz, Melania Nowicka, Birte Kehr und Sebastian Niehus

zukunftsorientierte analyse von forschungsdaten 42
Mitglieder des de.NBI-Netzwerkes nutzen KI-basierte Technologien
von Vera Ortseifen, Peter Belmann und Alfred Pühler

SmICS 46
Eine smarte Lösung zur automatisierten Ausbruchserkennung im Krankenhaus
von Antje Wulff, Claas Baier, Michael Marschollek und Simone Scheithauer für alle Mitglieder und Standorte des Use Case Infection Control des HiGHmed Projektes und die Infection Control Study Group

nationale forschungsdateninfrastruktur NFDI 50
Der gemeinsame Aufbau eines FAIRen Forschungsdatenmanagements in Deutschland
von Nathalie Hartl

GAIA-X als enabler für einen gesundheitsdatenraum 54
Das 2019 gestartete deutsch/französische Projekt GAIA-X geht ins dritte Jahr. Eine Momentaufnahme.
von Harald Wagener

FAIRe datennutzung: erfahrungen aus verbundprojekten 57
Anforderungen und schrittweise Umsetzung in datengetriebenen Forschungsverbünden
von von Matthias Ganzinger und Matthieu-P. Schapranow

meldungen aus dem BMBF 62

news aus dem BIH 66

ein digitales gesundheitsportemonnaie für menschen ohne krankenversicherung 70
Firmenporträt: mTOMADY
von Stefanie Seltmann

„komplexe fragestellungen kann man nur durch interdisziplinarität beantworten“ 73
Interview: Dana Westphal und Michael Seifert
von Marco Leuer.

INCOME – integrative, kollaborative modellierung in der systemmedizin 77
Per Hackathon in eine Kultur des *Data and Model Sharing*
von Nina Fischer, Wolfgang Müller, Dagmar Waltemath, Olaf Wolkenhauer und Jan Hasenauer

e:Med juniorverbund LeukoSyStem 82
Interview: Simon Haas
von Silke Argo

„we love data!“ 87
Porträt: Claudia Langenberg
von Stefanie Seltmann

news 90

events 94

impressum 101

wir über uns 102

kontakt 103

„das leben ist unvorhersehbar!“

Portrait über die Systembiologin Barbara Di Ventura

von Katharina Kalhoff und Gesa Terstiege

Zellen möglichst exakt zu beeinflussen, ist das Ziel der Optogentik. Diese Kombination von optischen und genetischen Methoden hat seit den ersten Ansätzen vor gut zwanzig Jahren große Fortschritte gemacht. Professor Barbara Di Ventura möchte diese Methoden weitervorantreiben und vor allem ihr therapeutisches Potenzial ausloten. Dabei setzt sie neben ihren Erfahrungen aus den Ingenieurswissenschaften und der Molekularbiologie vor allem auf ihre Leidenschaft und Motivation für das Thema. Sie ist überzeugt, dass diese Methoden Krebstherapien entscheidend verbessern können.

Ihr Vater wünschte sich für sie eine Karriere als Rechtsanwältin, weil sie als Kind nicht aufhörte zu reden, bis ihr alle zustimmten. *„Meine Mutter hatte Sorge, dass ich dann mit zu vielen korrupten Menschen oder Mafiosi zu tun habe. Mich hat allerdings hauptsächlich der Blick ins bürgerliche Gesetzbuch abgeschreckt – das war mir zu langweilig“*, erklärt die Wissenschaftlerin Barbara Di Ventura von der Universität Freiburg und lacht. Statt auf Jura fiel die Studienwahl in ihrer Heimat Rom dann auf Ingenieurwissenschaften mit Schwerpunkt Informatik. Heute ist sie Professorin für biologische Signalforschung an der Universität Freiburg und möchte mit ihrer Forschung in Zukunft bessere Krebstherapien ermöglichen.

„Schnell hatte ich keine Lust mehr, immer auf die Daten der Experimentierenden zu warten und habe mich selbst ins Labor gestellt – ohne jegliche Vorkenntnisse.“

Ihr Wechsel zur Biologie begann mit einer Masterarbeit zur Modellierung von Gennetzwerken. Ursprünglich wollte die Ingenieurin gar nicht promovieren, aber sie reizte ein Auslandsaufenthalt. Das Europäische Labor für Molekularbiologie (EMBL) suchte damals Doktoranden an mehreren Standorten in Europa. *„Deutschland hatte ich für meine Zukunft gar nicht auf dem Schirm, aber ich habe mich sehr schnell in Heidelberg verliebt. Das Leben ist unvorhersehbar“*, kommentiert sie ihren Werdegang. Geplant war, dass sie Computermodelle entwickelt, die auf biologischen Experimenten beruhen. *„Schnell hatte ich keine Lust mehr, immer auf die Daten der Experimentierenden zu warten und habe mich selbst ins Labor gestellt – ohne jegliche Vorkenntnisse. Mein damaliger Betreuer Luis Serrano hat mich zum Glück sehr unterstützt“*, erinnert sich die Wissenschaftlerin.

„Wobei ich auch immer inspirierende Mentoren und Unterstützer hatte, die mich gefördert und in meinen wissenschaftlichen Visionen unterstützt haben – Wissenschaft ist Teamarbeit.“

Gute Zusammenarbeit im Team ist entscheidend

Ein wichtiger Schritt in diese Richtung war der Aufbau einer eigenen Arbeitsgruppe. Das Bundesministerium für Bildung und Forschung (BMBF) hat sie hierbei im Rahmen einer „e:Bio-Nachwuchsgruppe“ mit knapp 1,5 Millionen Euro unterstützt. *„Die Förderung war sehr wichtig für meine weitere Karriere. Ohne die Nachwuchsgruppe hätte ich die Professur in Freiburg wahrscheinlich nicht bekommen. So konnte ich zeigen, dass ich unabhängig arbeiten und Themen vorantreiben kann“*, erklärt die Wissenschaftlerin. *„Wobei ich auch immer inspirierende Mentoren und Unterstützer*

Von Rom nach Süddeutschland: Die Systembiologin Prof. Dr. Barbara Di Ventura forscht inzwischen an der Universität Freiburg (Foto: Harald Neumann).

hatte, die mich gefördert und in meinen wissenschaftlichen Visionen unterstützt haben – Wissenschaft ist Teamarbeit“. In ihrem Forschungsfeld geht es dabei vor allem um eine Zusammenarbeit von Menschen, die sich den biologischen Fragestellungen auf verschiedene Weise nähern. Ein Teil ihres Teams entwickelt sogenannte optogenetische Werkzeuge, andere modellieren die Vorgänge und wieder ein anderer Teil nutzt diese Werkzeuge, um biologische Fragen zu beantworten.

„Auf dem Gebiet der Optogenetik wäre die Entwicklung eines einkomponentigen, kleinen, einfach zu bedienenden optogenetischen Werkzeugs, das auf rotes Licht reagiert, ein echter Durchbruch. Diese Art von Licht kann tiefer in das Gewebe eindringen und könnte eines Tages mit Erfolg bei Patienten eingesetzt werden, um Krebs zu behandeln.“

Optogenetik bezieht sich dabei, wie der Name schon verrät, auf die Kombination aus optischen Methoden und Genetik. Die Forscherinnen und Forscher nutzen Licht, um z. B. einzelne Gene an- oder auszuschalten. Licht als Trigger für molekulare Prozesse zu nutzen hat einige Vorteile. Zum einen ist die Reaktion sehr schnell und sehr einfach anzuwenden. Licht löst keine unerwünschten Wirkungen in der Zelle aus – im Gegensatz zu kleinen Molekülen, die sonst häufig für Manipulationen in der Zelle eingesetzt werden.

Ihre Begeisterung für die Arbeit schwingt in jedem Satz mit, wenn die 45-Jährige über ihre Arbeit spricht. *„Auf dem Gebiet der Optogenetik wäre die Entwicklung eines einkomponentigen, kleinen, einfach zu bedienenden optogenetischen Werkzeugs, das auf rotes Licht reagiert, ein echter Durchbruch. Diese Art von Licht kann tiefer in das Gewebe eindringen und könnte eines Tages mit Erfolg bei Patienten eingesetzt werden, um Krebs zu behandeln.“*

Als Ingenieurin, die erst spät in die Molekularbiologie eingestiegen ist, sei ihr Fachgebiet, die Synthetische Biologie, genau das Richtige für sie gewesen. *„Hier kann ich mein Wissen aus beiden Bereichen perfekt kombinieren“*, erzählt sie begeistert. Mittler-

Wissenschaft ist Teamarbeit: Barbara Di Ventura legt großen Wert darauf, dass in ihrer Arbeitsgruppe viele Disziplinen gemeinsam an einem Ziel arbeiten (Foto: Harald Neumann).

weile hat sie eine Arbeitsgruppe mit zehn Mitarbeiterinnen und Mitarbeitern aufgebaut, von denen 80 Prozent experimentell und 20 Prozent an theoretischen Modellen arbeiten.

„Auch wenn es verrückt klingt: Ich genieße jeden Arbeitstag“

Privat ist die Italienerin ebenfalls seit Jahren in Deutschland angekommen und genießt das Leben in Süddeutschland mit Mann und Sohn. Die internationale Forschung ist und bleibt aber genauso wichtig für sie. Sie erfreue sich an der Freiheit und der Kreativität, die mit der Forschung verbunden seien und genieße die Interaktion mit ihrem Team und den Studierenden: *„Auch wenn es verrückt klingt: Ich genieße jeden Arbeitstag“.*

Nur eine Sache kann bei der lebensfrohen Italienerin für kurzzeitige Verstimmung sorgen: *„Da ich selbst ein extrem engagierter und motivierter Mensch bin, frustriert es mich, wenn meine Studierenden unmotiviert sind. Mich irritiert mangelnde Hingabe.“* Wer die Wissenschaftlerin in Zeiten ohne pandemiebedingte Einschränkungen nicht an der Universität oder auf internationalen Kongressen antrifft, hat vielleicht einmal die Chance, sie über die musikalische Seite kennenzulernen. *„Ich singe leidenschaftlich gern in einer Rockband. Es ist natürlich keine Voraussetzung, aber wer wissenschaftlich hochmotiviert ist und ein Instrument spielt, kann sich gerne bei uns bewerben“*, verrät sie mit einem Augenzwinkern.

> „Da ich selbst ein extrem engagierter und motivierter Mensch bin, frustriert es mich, wenn meine Studierenden unmotiviert sind. Mich irritiert mangelnde Hingabe.“

Kontakt:

Prof. Dr. Barbara Di Ventura
Institut für Biologie II
Arbeitsgruppe: Molecular and Cellular Engineering
Albert-Ludwigs-Universität Freiburg
barbara.diventura@biologie.uni-freiburg.de

http://diventura.bioss.uni-freiburg.de

wie habe ich die pandemie erlebt?

1,5 Jahre Corona aus der Sicht eines Nachwuchswissenschaftlers

von Sören Lukassen

Als sich COVID-19 Anfang 2020 weltweit auszubreiten begann, war absehbar, dass diese Pandemie mindestens kurzfristig dramatische Auswirkungen auf unser Leben haben würde. Der Ausbruch in Wuhan hatte gezeigt, dass auch bei einem lokalen Ausbruch in einem Land mit eigentlich guten Ressourcen die Gesundheitsversorgung an ihre Grenzen stoßen konnte. Auch war zu diesem Zeitpunkt wenig über die Übertragung, Risikofaktoren, und Letalität von COVID-19 bekannt. Meine Bedenken kreisten zum damaligen Zeitpunkt um existenzielle und organisatorische Fragen: Wie kann ich eine Ansteckung verhindern? Was passiert, wenn meine Freunde, Familie, oder ich uns anstecken? Die größte Umwälzung, die die Pandemie für mich mit sich bringen sollte, tauchte zu diesem Zeitpunkt gar nicht in meinen Gedanken auf: meine Arbeit als Wissenschaftler.

Als Bioinformatiker war ich in der glücklichen Situation, auch von zu Hause aus arbeiten zu können, ein Lockdown hätte mich also nicht großartig behindert. Noch bevor der erste COVID-19 Fall in Berlin auftrat holte mich aber gewissermaßen meine Vergangenheit ein: bevor ich mich als Post-doc mit der Identifikation funktioneller Gensets in single-cell RNA-Seq Daten beschäftigte, hatte ich eine Zeit lang als Bioinformatiker in einer Core Unit gearbeitet und unter anderem RNA-Seq und single-cell RNA-Seq Datensätze ausgewertet. In meiner Arbeitsgruppe gab es mehrere Projekte die sich mit Lungenerkrankungen beschäftigten, und jetzt wurde jemand gebraucht, der schnell die verschiedenen Kontrolldatensätze zusammenführen, analysieren und als Referenzdatensatz für andere Forscher aufbereiten konnte. Währenddessen sollten Proben von so vielen COVID-19 Patienten wie möglich genommen und sequenziert werden, um die Erkrankung auf Einzelzellebene charakterisieren zu können.

Was als kurzes und spontanes Projekt begann, entwickelte sich schnell zu einer Herkulesaufgabe mit einer Vielzahl von Beteiligten und resultierte in mehreren Publikationen und Datensätzen, die wiederum von Dutzenden anderen Gruppen für ihre eigene Forschung verwendet wurden. Die Mitarbeit an diesem Projekt war für Nachwuchswissenschaftler wie mich sicherlich anstrengend, aber auch lehrreich. Durch das hohe Tempo treten einerseits Probleme deutlich stärker hervor, andererseits kann man aber auch Prozesse die sonst Jahre dauern würden und für Nachwuchswissenschaftler schwer zu verfolgen sind im Zeitraffer betrachten. Im Folgenden möchte ich ein paar meiner Beobachtungen aus den letzten anderthalb Jahren festhalten.

Die Dringlichkeit einer Pandemie erzwingt Kooperation

Viele Forschungsprojekte werden primär durch eine einzelne oder wenige Personen vorangetrieben. Ein großes Ziel von Doktorarbeits- oder Post-docprojekten ist die Erstautorenschaft auf dem „eigenen“ Paper. Zwar gibt es die Möglichkeit der geteilten Erstautorenschaft, aber der Verweis auf Paper mit „Mustermann *et al.*“ zeigt, dass der ersten Position auf der Autorenliste in der Wahrnehmung eine besondere Bedeutung zukommt. In der Regel ist dies ein Post-doc oder Doktorand, der das Projekt hauptsächlich vorangetrieben und sich über Jahre die nötige Expertise erarbeitet hat.

„Was als kurzes und spontanes Projekt begann, entwickelte sich schnell zu einer Herkulesaufgabe mit einer Vielzahl von Beteiligten.“

Sören Lukassen, Nachwuchsgruppenleiter Medical Omics (Foto: Sören Lukassen).

Diese Arbeitsweise funktioniert in einer Situation wie der COVID-19 Pandemie nicht. Sie wäre einerseits zu langsam, andererseits bedeutet das Auftreten einer neuen Erkrankung, dass keine Arbeitsgruppe eine ausreichende Expertise hat, um komplexe Datensätze adäquat zu analysieren und interpretieren. Hierfür müssen sich im Grunde Konsortien im Kleinen bilden, bei denen Forscher mit unterschiedlichster Expertise eng zusammenarbeiten. In unserem Fall gab es nahezu täglichen Austausch zwischen Bioinformatikern, Immunologen, Virologen, und Klinikern, was zeitgleich einen viel intensiveren Blick über den Tellerrand ermöglichte als es sonst der Fall gewesen wäre. Gleichzeitig lässt diese intensive Kommunikation aber auch Konflikte deutlich hervortreten, was mich zum nächsten Punkt bringt:

Forschung als Intensivkurs in Gruppendynamik

In Managementkursen wird gerne das Phasenmodell der Teambildung gelehrt, bei dem es nach einem ersten zögerlichen Beschnuppern zu Konflikten kommt, die schließlich gelöst werden, was produktives Arbeiten ermöglicht. Für altgediente Arbeitsgruppenleiter mag das nichts Neues sein, aber als Nachwuchswissenschaftler kommt man in der Regel in eine bestehende Gruppe und erlebt selten die Dynamiken einer neu gegründeten Arbeitsgruppe. Bei Kooperationsprojekten wiederum gibt es oft genug Abstand zwischen den Parteien, um das Konfliktpotenzial minimal zu halten. Für eine schnelle Analyse der Daten zu COVID-19 war aber eine deutlich engere Zusammenarbeit nötig: Generierung der Daten, Analyse und Interpretation fanden parallel durch Gruppen mit unterschiedlichster Exper-

Über den Autor:

Sören Lukassen ist Nachwuchsgruppenleiter am BIH-Zentrum für Digitale Gesundheit der Charité. Nach dem Studium der Molekularen Medizin und der Promotion in Humangenetik arbeitete er als Bioinformatiker an der Erlanger Genomics Core Unit. In seiner Zeit als Post-doc in Heidelberg und Berlin beschäftigte er sich mit Machine Learning für single-cell RNA-Seq, bevor er Anfang 2020 die Arbeit an verschiedenen Studien zu den Mechanismen schwerer Erkrankung bei COVID-19 begann.

tise statt, die sich in unserem Fall täglich austauschten. Da es keine vorbestehenden Strukturen gab, mussten die Zuständigkeiten dynamisch geklärt werden. Die Beteiligten mussten hierbei einen Plan für das Gesamtprojekt aushandeln, der die Interessen und Ressourcen aller Gruppen vertrat, was anfangs durchaus zu lebhaften Diskussionen führte. Auch wenn die enge Kooperation zunächst zu Konflikten führte, zahlte sie sich mittelfristig jedoch aus: statt Meinungsverschiedenheiten aus dem Weg zu gehen mussten diese ausdiskutiert werden, was nach einer Eingewöhnungszeit zu starkem gegenseitigem Vertrauen und einer hocheffizienten Arbeitsatmosphäre führte, wie sie bei einer lockereren Kooperation wohl nicht entstanden wäre.

Kurzfristige Projekte erfordern trotzdem langfristige Planung

Aus dieser engen Zusammenarbeit entstanden schnell mehrere gemeinsame Publikationen, die Datenmengen wuchsen und es kamen immer neue Ideen für Projekte zusammen. Hatten wir anfangs noch single-cell RNA-Seq Datensätze mit ca. 20 Proben und einigen zehntausend Zellen, wurden daraus schnell weit über hundert Proben und hunderttausende von Zellen. Was uns einzigartige Einblicke in die Biologie von COVID-19 ermöglichte, wurde aber auch mehr und mehr zu einem Problem: unsere Analysepipelines hielten nicht mit der steigenden Datenmenge schritt. Bei der Wahl der Tools hatten wir Anfang 2020 vor allem auf für uns bewährte und vor allem zu anderen Projekten rückwärtskompatible Pipelines geachtet. Dies ermöglichte die schnelle Übernahme von gesunden Kontrollproben aus anderen Projekten und reduzierte den Aufwand bei der Analyse erheblich. Auch wussten wir, dass unsere Analysescripts bis 150,000 oder 200,000 Zellen funktionieren würden, was angesichts der geplanten vielleicht hunderttausend Zellen ausreichend erschien. Als auf unsere ersten Projekte weitere folgten, begann die Pipeline aber zu einem Problem zu werden. Zuerst hielten wir trotzdem an ihr fest, da wir ja schon mit der Herangehensweise publiziert hatten und Konsistenz wahren wollten. In der Finanzwelt würde man hier wohl von einer „sunk cost fallacy“ sprechen – wir hatten so viel Zeit in die Analysescripts investiert und so viele Daten damit generiert, dass wir uns nicht davon trennen wollten. Letztendlich kamen wir natürlich doch an den Punkt, an dem wir unsere Pipeline von Grund auf neu aufstellen mussten. Die Zeit, die wir hierfür investieren mussten, wurde problemlos durch die um ein Vielfaches schnellere Laufzeit aufgewogen, und der Satz „könntet ihr das nochmal mit zusätzlichen Proben laufen lassen?“ verlor seinen Schrecken.

Diese Erfahrung hat bei mir zu mehreren Erkenntnissen geführt:

1. **Analysepipelines und Infrastruktur sollten so ausgelegt sein, dass sie auch aktuell unrealistisch erscheinende Datenmengen bewältigen können**
2. **Ähnlich wichtig wie ein Blick auf die absolute Performance eines Algorithmus ist der Blick auf seine Skalierbarkeit. Wo diese Angaben fehlen, sollte man sich die Zeit nehmen und diese selbst bestimmen**
3. **Sobald absehbar wird, dass die Datenmenge für die bestehenden Verfahren zu groß werden könnte, sollte man rechtzeitig einen Umstieg planen. In die alte Pipeline investierte Zeit ist kein Argument, einen Wechsel hinauszuzögern.**

In der Pandemie haben sich einige Erkenntnisse durchgesetzt, die an sich nicht neu waren. Im Wissenschaftsbetrieb wären hier beispielsweise die Vor- und Nachteile von Preprints und damit verbunden der Nutzen von post-publication peer review zu nennen. Neben diesen Lerneffekten im Großen gibt es aber auch die im Kleinen, sei es bezüglich Zeitmanagement, Homeoffice, oder eben Zusammenarbeit und Projektplanung. Wenn wir diese für uns nutzen können, hätten die letzten anderthalb Jahre zumindest in dieser Hinsicht etwas Gutes gehabt.

Kontakt:

Dr. Sören Lukassen
Leiter Nachwuchsgruppe Medical Omics
Digital Health Center
Berlin Institute of Health at Charité
soeren.lukassen@charite.de

www.hidih.org/research/medicalomics

besserer zugriff auf klinisch relevante daten

BMBF-Fördermaßnahme „i:DSem“ zieht ein positives Resümee

von Björn Dreesen-Daun

Klinische oder experimentelle Primärdaten aber auch Sekundärdaten, wie Fachpublikationen wachsen explosionsartig. Erschwerend kommt hinzu, dass diese Daten häufig sehr heterogen, d. h. unstrukturiert in verschiedenen Formaten und unterschiedlicher Qualität vorliegen. Der Forschungsansatz der „Integrativen Datensemantik“ adressiert dieses Dilemma. Ziel ist es, Instrumente für die Homogenisierung dieser heterogenen Daten zu entwickeln und eine überwiegend computergestützte Verarbeitung möglich zu machen. Dies ist eine notwendige Voraussetzung, um in Zukunft sämtliche klinisch relevante Daten für die Entwicklung verbesserter Therapien nutzen zu können.

Die Fördermaßnahme „i:DSem – Integrative Datensemantik in der Systemmedizin“ zählt zu den Initiativen des Bundesministeriums für Bildung und Forschung (BMBF), die Innovationen im Bereich der computergestützten Datenverarbeitung für die Gesundheitsforschung voranbringen wollen. Anlass zum Start der Initiative im Jahr 2016 war, dass immer mehr Wissen produziert wird, jedoch erhebliche Hürden bestehen, dieses für die Problemlösung der Ärztinnen und Ärzte zur Verfügung zu stellen. Diese Herausforderung bedarf der Entwicklung neuartiger Instrumente, sogenannter semantischer Technologien, die im Zentrum der Maßnahme „i:DSem“ stehen.

Ende des Jahres 2021 läuft die Förderung nach etwas mehr als fünf Jahren und zahlreichen Erfolgen aus. Im Förderzeitraum hat das BMBF über 20 Millionen Euro für insgesamt acht interdisziplinäre Forschungsprojekte zur Verfügung gestellt.

Die Projekte der Maßnahme gliederten sich in zwei Phasen. Zunächst gab es eine dreijährige Entwicklungsphase, nach einer Zwischenevaluierung im Jahr 2019 traten dann alle Projekte erfolgreich in die zweijährige Translationsphase ein, in der die erfolgten Entwicklungen für die konkrete Anwendung validiert werden sollen. Mit dem Abschluss dieser Phase enden im Laufe des Jahres 2021 nun die Projekte.

Das BMBF blickt insgesamt auf eine erfolgreiche Maßnahme zurück, wie auch die digitale Abschlussveranstaltung im Sommer 2021 bestätigte. Seit dem Start von „i:DSem“ hat das BMBF weitere Fördermaßnahmen zur Stärkung der computergestützten Gesundheitsforschung initiiert: So soll mit der Maßnahme „Computational Life Sciences“ insbesondere das erhebliche Potential der künstlichen Intelligenz für die lebenswissenschaftliche Forschung erschlossen werden. Zudem baut die BMBF-geförderte Medizininformatik-Initiative wichtige IT-Infrastrukturen an den Universitätskliniken auf, um u.a. klinische Routinedaten so aufzubereiten, dass sie institutionenübergreifend für die medizinische Forschung nutzbar sind.

Im Folgenden werden beispielhaft an drei ausgewählten Projekten die Erfolge der „i:DSem-Maßnahme“ illustriert.

Kontakt:

Dr. Björn Dreesen-Daun
Projektträger Jülich
Forschungszentrum Jülich GmbH
b.dreesen@fz-juelich.de

www.ptj.de/idsem

Foto: FZJ, R. U. Limbach

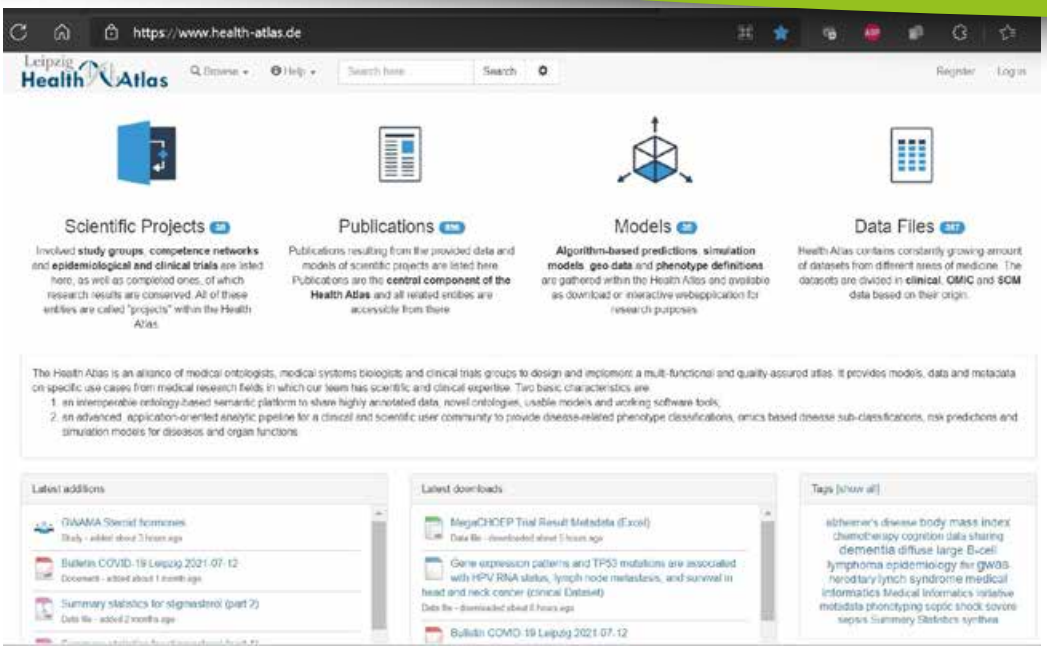


Abbildung 1: Screenshot der Startseite des Health Atlas, erreichbar unter dem Link: https://www.health-atlas.de

health atlas als plattform für gesundheitsforschung

Der Health Atlas stellt strukturierte Forschungs(meta)daten, Applikationen und Modelle nach FAIR-Kriterien zur Verfügung

von Markus Löffler

Der Health Atlas (https://www.health-atlas.de) ist eine interoperable Plattform der Gesundheitsforschung. Publikationen, strukturierte Daten, Modelle und Applikationen klinischer und epidemiologischer Forschung stehen darin FAIR zur Verfügung.

Publikationen zeigen den wissenschaftlichen Erkenntnisgewinn, nicht nur in der Medizin. Sie fußen auf Daten und durchgeführten Analysen (Analysepläne und -Skripte), die für eine reproduzierbare und nachvollziehbare Forschung zunehmend von Journalen eingefordert werden. Der Zugang zu diesen Ressourcen treibt die Weiterentwicklung von Ideen und das Aufkommen weiterer Forschungsfragen an.

Der Health Atlas bietet Zugang zu diesen Ressourcen mit einer Plattform, die neben den Publikationen ebenso qualitätsgesicherte Daten und Metadaten interoperabel zur Verfügung stellt. Anwenderinnen und Anwender können ihre Publikationen samt zugrundeliegenden Forschungsdaten, Supplements und Metadaten nach einem Registrierungsprozess selbstständig importieren. Publikationsmetadaten (Autoren, Keywords, eine Projektbeschreibung und Zuweisung zu der entsprechenden Erkrankung in der Human Disease Ontology) optimieren deren Findability und Nachnutzbarkeit. Während die Publikationen frei zugänglich sind, kann der Zugang zu Daten begrenzt werden („Zugang mit Passwort“, „Zugang auf Anfrage“, oder „frei zugänglich“).

Was ist besonders?

Neben Publikationen und Daten bietet der Health Atlas den Zugang zu verschiedenen Modellen, z. B. zur Prädiktion und Risikobeurteilung von Erkrankungen, die von neuartigen Modellen zur Typisierung von Körperformen und zur COVID-19-Pandemie erweitert werden. Alle Modelle sind mit modernen Web-Applikationen umgesetzt, die zum Ausprobieren und zur Nutzung in Forschungsprojekten dienen sollen. So können Wissenschaftlerinnen und Wissenschaftler z. B. zwischen diversen publizierten Prädiktionsmodellen für familiären Brust-

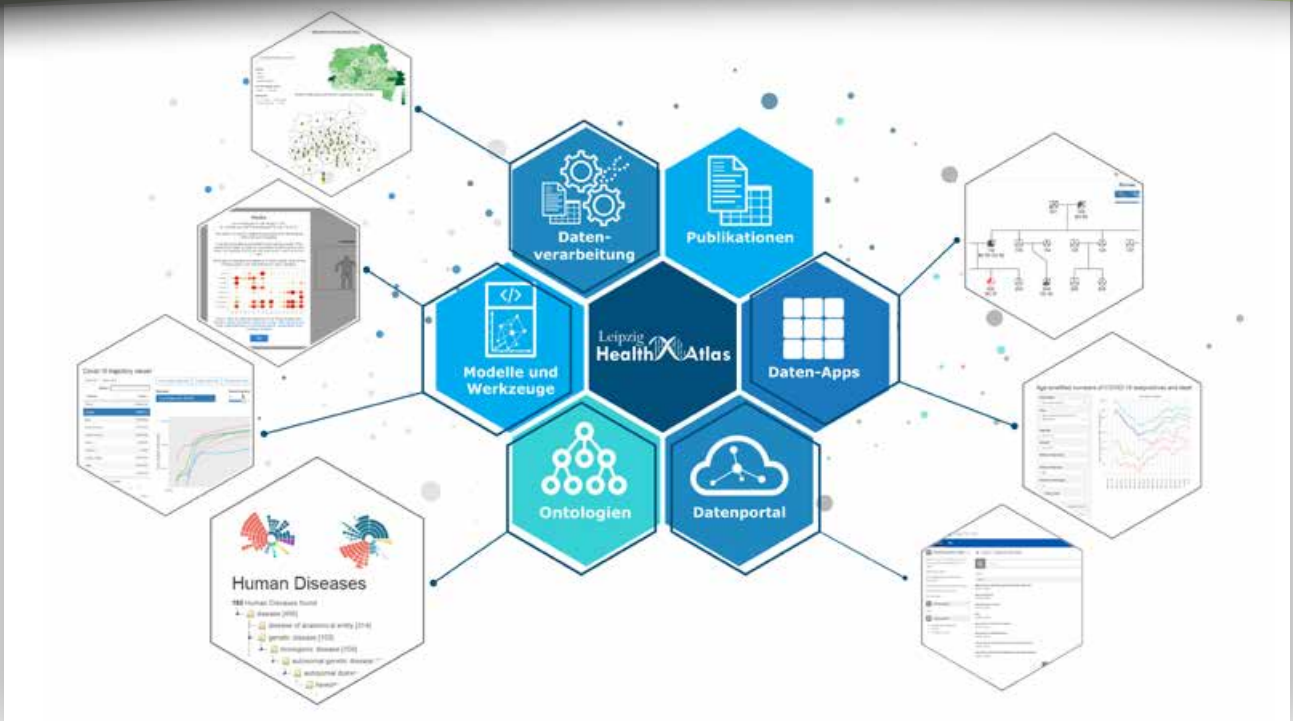


Abbildung 2: Die Diversität des Health Atlas als interoperable Plattform der Gesundheitsforschung. Die Bausteine sind das Datenportal, Modelle und Applikationen zur online Datenverarbeitung und Ontologien zur besseren Suche (Quelle: René Hänsel, IMISE).

und Eierstockkrebs wählen und die erhaltenen Ergebnisse vergleichen, ohne die Modelle selbst implementieren zu müssen. Andere Modelle fokussieren auf normative Daten. Oftmals sind solche Daten in Tabellenform in Publikationen verfügbar, die bei konkreten Messungen zur Einordnung der Messwerte dienen. Deren Nutzung und Einbeziehung in weitergehende Analysen und Visualisierungen obliegt jedoch dann den Wissenschaftlern. Andere Modelle visualisieren die COVID-19 Fall-Statistiken oder geben eine Einordnung zur aktuellen Pandemiesituation.

Was macht uns FAIR?

Die Intention des Health Atlas liegt auf dem Teilen von Daten. Daher war es ein Grundanliegen, den FAIR-Kriterien Rechnung zu tragen. Eindeutige Identifier („Health Atlas ID“) werden verwendet, um Inhalte referenzieren zu können. Durch Nutzung der SEEK-Plattform sind die Inhalte neben der Benutzerschnittstelle über aktuelle Schnittstellen erreichbar. Der Datenzugang kann abgestuft vorgenommen werden.

Auf dem eWorkshop „FAIR Data Infrastructures for Biomedical Communities“ am 15.10.2020 verwies Professorin Carole Goble (University of Manchester und u. a. Koordinatorin von FAIR-DOM) auf den Health Atlas als Beispiel für gelungenes Data Sharing.

Der Health Atlas (ursprünglich Leipzig Health Atlas, LHA) wurde im Rahmen der i:DSem-Initiative zu integrativer Datensemantik in der Systemmedizin des BMBF vom Institut für Medizinische Informatik Statistik und Epidemiologie (IMISE), dem Interdisziplinären Zentrum für Bioinformatik (IZBI) und dem LIFE-Research Center for Civilization Diseases entwickelt. Er ist integraler Bestandteil weiterer Konsortien der Medizininformatik-Initiative (SMITH) und der Nationalen Forschungsdateninfrastruktur (NFDI4Health).

Haben wir Ihr Interesse geweckt? Werden Sie Teil des Health Atlas und registrieren sich kostenlos unter:
https://www.health-atlas.de/signup

Kontakt:

Prof. Dr. Markus Löffler
Leiter des Instituts für
Medizinische Informatik, Statistik und
Epidemiologie – IMISE
Universität Leipzig
info@health-atlas.de

www.imise.uni-leipzig.de

Foto: M. Löffler

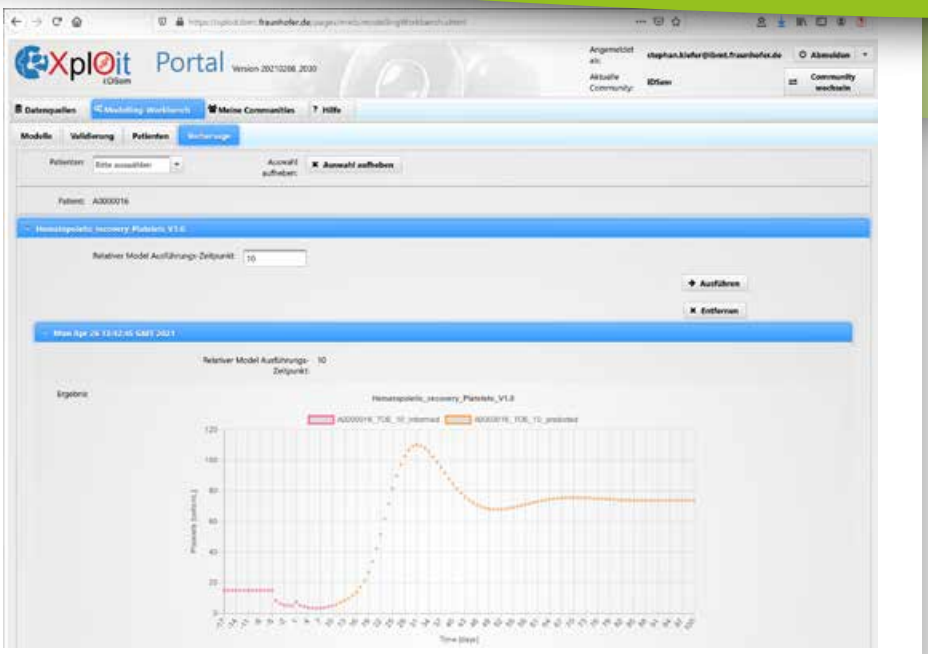


Abbildung 3: Darstellung der Ergebnisse einer beispielhaften Ausführung des Modells der Blutbildung auf der XplOit-Plattform. Das Modell wurde auf Basis der roten Datenpunkte bis Tag 10 nach Stammzelltransplantation informiert. Die gelben Datenpunkte stellen die vom Modell vorhergesagten Blutplättchen bis Tag 100 nach Stammzelltransplantation dar. (Abbildung: Fraunhofer IBMT und Universität des Saarlandes).

XplOit: effizient gesundheitsvorhersagemodelle entwickeln

IT-Unterstützung zur Entwicklung und Validierung von Vorhersagemodellen bei Stammzelltransplantation

von Stephan Kiefer

Prädiktive Computermodelle können Ärzte dabei unterstützen, maßgeschneiderte Therapien anzubieten, indem sie das individuelle Risiko von Komplikationen und das Ansprechen auf die Behandlung vorhersagen. In der Stammzelltransplantation werden solche Modelle dringend gebraucht. Jedoch ist ihre Erstellung wegen der großen Vielfalt benötigter Daten bei gleichzeitig kleinen Fallzahlen aufwendig und die Umsetzung im klinischen Alltag schwierig.

Um dieses Hindernis zu überwinden, stand im Zentrum des kürzlich abgeschlossenen Projektes XplOit die Entwicklung einer neuen speziellen IT-Plattform, die sog. XplOit-Plattform. Mit ihrer Hilfe sollen validierte Computermodelle für die Stammzelltransplantationsmedizin bereitgestellt werden.

Die XplOit-Plattform

Die gemeinsam von IT-Experten des Fraunhofer-Instituts für Biomedizinische Technik und der Universität des Saarlandes entwickelte webbasierte XplOit-Plattform (Weiler *et al.*, 2018) ermöglicht es, Ärzten, vielfältige klinische Daten zur Erstellung von Vorhersagemodellen unter Einhaltung von Datenschutzbestimmungen zusammenzuführen und mit Modellentwicklern zu teilen, sowie die entwickelten Modelle zu erproben. Kliniker können mit lokalen Tools die benötigten Daten pseudonymisieren, dabei Datumsangaben in Bezug zum Transplantationsdatum relativieren und die Daten per Upload über spezielle Datenpipelines in einem Data Warehouse zur Verfügung stellen. Für semistrukturierte Befundtexte hat der industrielle Projektpartner Averbis entsprechende Instrumente zur Informationsextraktion geschaffen. So konnten die beteiligten Transplantationszentren des Universitätsklinikums

Essen und des Universitätsklinikums des Saarlandes sowie ihre virologischen Institute Daten von über 2500 Patienten nach allogener Stammzelltransplantation mit mehr als 15 Millionen Datenpunkten für die Modellentwicklung zusammentragen.

Um unterschiedliche Formate und Benennungen in den Daten zu harmonisieren, verfolgte das Projekt einen Datenintegrationsansatz, bei dem eine sogenannte Ontologie als Wissensbasis die Fachtermini in der Transplantationsmedizin und ihre Beziehungen untereinander beschreibt. Ein spezielles Werkzeug erlaubt es, benötigte Konzepte für diese projektspezifische Ontologie aus etablierten Ontologien zu entnehmen und zu ergänzen. Modellentwickler können so in den bereitgestellten Datensätzen die für ihr geplantes Modell relevanten Daten annotieren und zusammenführen, um anschließend mit Analyse- und Visualisierungswerkzeugen Korrelationen zu finden und Hypothesen zu generieren. Die für das Trainieren von Modellen benötigten Daten können in entsprechende Umgebungen übernommen werden. Die entwickelten Modelle werden anschließend in ein Modell-Repository hochgeladen, und nach Annotation der benötigten Eingabeparameter mit Daten aus dem Data Warehouse der Plattform gespeist und zur Ausführung gebracht. Dabei können die Modelle zu unterschiedlichen Zeitpunkten eines retrospektiven Behandlungsverlaufs ausgeführt werden und mit der tatsächlichen Gesundheitsentwicklung oder auch mit Einschätzungen der behandelnden Kliniker zu diesem Zeitpunkt verglichen und so auf ihren Nutzen überprüft werden.

Vorhersagemodelle nach hämatopoetischer Stammzelltransplantation

Projektpartner der Eberhard Karls Universität Tübingen und der Universität des Saarlandes entwickelten erste prädiktive Modelle für Patientinnen und Patienten nach Zelltransplantation, die mögliche lebensbedrohliche Komplikationen wie die gefürchtete Graft-versus-Host-Disease oder eine Reaktivierung des Zytomegalievirus als auch die Blutneubildung und das Überleben individuell vorhersagen und die Kliniker dabei unterstützen, lebensrettende Maßnahmen früher als bisher anzuwenden. Es wurden sowohl Machine-Learning-Modelle für die Vorhersage von Ereignissen, als auch systembiologische Modelle für die Vorhersage eines zeitlichen Verlaufs entwickelt. Abbildung 3 zeigt exemplarisch die Vorhersage der Blutplättchenneubildung (dargestellt als gelbe Datenpunkte) eines hypothetischen Patienten nach Stammzelltransplantation. Eine erste Bewertung der Leistungsfähigkeit der in XplOit entwickelten Modelle im Vergleich zur Einschätzung der Kliniker und dem tatsächlichen Ergebnis wird auch durch die XplOit-Plattform unterstützt und steht kurz vor dem Abschluss.

Referenzen:

Weiler, G., Schwarz, U., Rauch, J., Rohm, K., Lehr, T., Theobald, S., Kiefer, S., Götz, K., Och, K., Pfeifer, N., Handl, L., Smola, S., Ihle, M., Turki, AT., Beelen, DW., Rissland, J., Bittenbring, J. und Graf, N. (2018). XplOit: An Ontology-Based Data Integration Platform Supporting the Development of Predictive Models for Personalized Medicine. Stud Health Technol Inform.247:21-25.

Kontakt:

Stephan Kiefer
Stellv. Leiter der AG
Biomedizinische Daten & Bioethik
Fraunhofer-Institut für
Biomedizinische Technik
stephan.kiefer@ibmt.fraunhofer.de

www.ibmt.fraunhofer.de

Foto: Stephan Kiefer

Abbildung 4: Übersicht der IDSN Architektur (A) und Ergebnis-Beispiele interaktiver Analysen von Daten aus SCA (B) und Genexpressionsdaten (C). A gibt einen Überblick der Architektur, die im Kapitel IDSN Konzept und in Madan, S. *et al.*, 2018 erläutert ist. B zeigt visuelle Auswertungen des Vergleichs von SCA Krankheitsbeginn in Abhängigkeit zu den Genvarianten SCA 1, 3 und 6 und der entsprechenden Allel-Länge (links). C visualisiert die Expression des Gens MAPT in Abhängigkeit verschiedener Krankheiten als Tabelle (links) oder graphisch (rechts) (Quelle: IDSN).

relevanz FAIRer daten für die neurodegenerative forschung

Das BMBF Projekt Integrative Datensemantik für die Neurodegenerationsforschung ermöglicht die Analyse komplexer Daten

von Juliane Fluck, Mischa Uebachs und Maksims Fiosins

Die Verfügbarkeit großer Mengen personenbezogener Forschungsdaten über Neurodegenerative Erkrankungen wie Alzheimer und Parkinson ermöglichen neue präventive oder therapeutische Interventionen. Die semantischen Technologien und FAIR Prinzipien, die das Projekt IDSN anwendet, öffnen neue Horizonte für die Therapie-Entwicklung oder Anwendung der Künstlichen Intelligenz (KI) in diesem Bereich.

Um das Leiden der Betroffenen und ihrer Familien sowie die wirtschaftliche Belastung einer wachsenden, alternden Bevölkerung zu reduzieren, ist im Bereich der neurodegenerativen Erkrankungen und der Demenz eine schnelle Translation von Forschung zu neuen Therapien notwendig. Das Hauptziel des BMBF-geförderten Projektes Integrative Datensemantik für die Neurodegenerationsforschung (IDSN,www.idsn.info) ist, Daten aus verschiedenen Forschungsbereichen zu integrieren und diese mit bestehenden Krankheitsinformationen zu kombinieren. Dazu haben sich Forschende des Deutschen Zentrums für Neurodegenerative Erkrankungen, des Fraunhofer-Institutes SCAI, der Universitätskliniken Bonn und Hamburg-Eppendorf und des ZB MED Informationszentrums Lebenswissenschaften in IDSN zusammengeschlossen, um (1) Standards, Methoden und Software zur Integration von Daten aus Hochdurchsatz-Screening, klinischen Kohorten und/oder klinischen Routinedaten zu entwickeln und (2) diese Daten in neuen Analysen zu nutzen. Im Folgenden wird das IDSN Konzept erläutert und 2 Analysebeispiele gegeben.

IDSN Konzept

Primärdaten werden lokal analysiert und verbleiben dort. Die Analyseergebnisse werden mit semantischen Annotationen unter Verwendung von Standard-Ontologien und Terminologien versehen, um Daten interoperabel zu machen. Diese Standards werden in der separaten semantischen Nachschlageplattform

SemLookP (https://semanticlookup.zbmed.de/) gespeichert, um geeignete Konzepte zur Annotation von Daten abzurufen, Informationen über die Konzepte bereitzustellen und Mappings zwischen verschiedenen Standards bereitzustellen (Übersicht, Abbildung 4A). Die semantisch angereicherten Daten inklusive Provenienz zu den Originaldaten werden an eine zentrale Datenmanagement-Plattform übergeben (Madan, S. *et al.,* 2018). Semantische Indizierungen ermöglichen schnelle Datenabfragen und -abrufe, interaktive grafische Benutzeroberflächen eine dedizierte Datenvisualisierung und eine schnelle und flexible weitere Analyse der Daten (Abbildung 4B, C).

Mechanismen der frontotemporalen Demenz

Frontotemporale Demenzen (FTD) sind eine heterogene Gruppe der Demenzerkrankung, die mit Funktionsverlust im Stirn- und Schläfenlappen einhergeht, oft frühzeitig (vor dem 65. Lebensjahr) beginnt und häufig genetisch oder durch Stoffwechselfehler bedingt ist. Ein Fokus von IDSN ist die kombinierte KI-Analyse genetischer, epigenetischer Gen- und Proteinexpressionsdaten sowie klinischer Daten, um neue Einblicke in FTD zu bekommen. Die semantische Annotation in IDSN erlaubt es, diese verschiedenen Datentypen in Beziehung zu setzen und so eine integrale Analyse von Expressionsdaten und klinischen Patientendaten zu ermöglichen. Ein weiterer Vorteil der Semantik in IDSN ist die einfache Verbindung und Nutzung von Sekundärdaten (wie z. B. Assoziationen zwischen Genen und Krankheiten). Die Kombination von semantisch annotierten Metadaten, großen Datensätzen und neuartigen KI-Modellen erlaubte es uns, neue potentielle Krankheitsmechanismen von FTD zu entdecken, wie z. B. eine stark erhöhte Vaskularisierung des erkrankten Gehirns (Menden, K. *et al.*: 2020 [1] und [2]).

Analyse von Kohorten-Daten spino-zerebelläre Ataxien

Die spino-zerebellären Ataxien (SCA) sind seltene erbliche Erkrankungen des Kleinhirns, die zu einer fortschreitenden Gleichgewichts- und Koordinationsstörung mit körperlicher Behinderung führen. Die Erhebung ausreichender Mengen klinischer Daten ist sehr aufwändig. Im IDSN-Projekt wurden zur Zusammenführung von vier weltweit verfügbaren klinischen Studien ein semantischer Standard und Importsoftware entwickelt, um eine integrierte Analyse der Daten zu ermöglichen (Uebachs, M. *et al.,* 2020). Die Entwicklung und Zuordnung zu diesem Standard ist nicht trivial: zum einen erfordert sie klinisches Fachwissen, zum anderen werden ähnliche Informationen mit unterschiedlichen Methoden und Skalierungen erfasst. Der Gewinn ist die kombinierte Analyse in einer intuitiven interaktiven Benutzeroberfläche, die gemeinsam mit Klinikern in IDSN entwickelt wurden (Abbildung 4).

Referenzen:

Madan, S. *et al.* (2018). A Semantic Data Integration Methodology for Translational Neurodegenerative Disease Research11th Semantic Web Applications and Tools for Healthcare and Life Sciences (SWAT4(HC)LS. Antwerp, Belgium, December 3-6, 2018. CEUR Workshop Proceedings, Vol-2275, C.J.O. Baker, A. Waagmeester, A. Splendiani, O. Beyan, M Scott-Marshall, eds. Paper 9.

Menden, K. *et al.* [1] (2020). Deep learning–based cell composition analysis from tissue expression profiles. Science advances 6(30).

Menden, K. *et al.* [2] (2020). Integrated multi-omics analysis reveals common and distinct dysregulated pathways for genetic subtypes of Frontotemporal Dementia. https://www.biorxiv.org/content/10.1101/2020.12.01.405894v3

Uebachs, M. *et al.* (2020). A new tool for prompt and comprehensive visualization and comparison of cohort characteristics in spinocerebellar ataxias. SCA &ARCA Global Conference 19.-21.10.2020.

Kontakt:

Prof. Dr. Juliane Fluck
Leitung Wissensmanagement
Fraunhofer Institut SCAI und ZB MED
Informationszentrum Lebenswissenschaften
und Universität Bonn
fluck@zbmed.de

www.scai.fraunhofer.de/de/geschaeftsfelder/bioinformatik.html

Dr. Mischa Uebachs
Facharzt Neurologie
Klinik und Poliklinik für Neurologie,
Universitätsklinikum Bonn
m.uebachs@gmail.com

https://neurologie.uni-bonn.de

Dr. Maksims Fiosins
Wissenschaftlicher Mitarbeiter
Universitätsklinikum Hamburg-Eppendorf
maksims.fiosins@gmail.com

https://www.uke.de/kliniken-institute/institute/medizinische-systembiologie/index.html

Foto: ZB MED, Uebachs, Fiosins

nutzung von klinischen routinedaten für forschung und versorgung

Voraussetzung: Der digitale modulare Broad Consent

von Clemens Spitzenpfeil, Alexandra Stein, Martin Bialke, Torsten Leddig und Wolfgang Hoffmann

Informationen über den Krankheitsverlauf und die Wirkung von Therapien sind für Ärzte wie auch Forschende von großer Bedeutung. Als Basis werden hierfür bislang klinische und epidemiologische Studien herangezogen, wohingegen die Daten aus dem Behandlungsalltag zumeist ungenutzt bleiben. Dabei hätten diese Daten den Vorteil, dass sie durch ihre vielfältigen Ausprägungen einen realitätsnahen Einblick in den klinischen Alltag erlauben. Genau hier setzt die MII mit dem Broad Consent an: diese Daten sollen der Forschung zugänglich gemacht werden. Für den Erfolg dieses Vorhabens ist ein Verständnis auf Seiten der Patienten und eine effiziente Umsetzung des Einwilligungsprozesses unumgänglich.

Worin liegt das Potential von Versorgungsdaten?

Das Verständnis des Einflusses von Risiko- und Einflussfaktoren auf die Entstehung und den Verlauf von Krankheiten und die Interaktionen zwischen verschiedenen Krankheitsbildern ist essentiell für die Verbesserung von Diagnostik und Therapie. Es ist gegenwärtig gängige Praxis, verschiedene klinische Studien miteinander zu vergleichen, um eine ausreichende Datenmenge als Referenzmaterial zu erhalten. Mit fortschreitendem Erkenntnisstand wird die Komplexität medizinischer Krankheitsbilder jedoch immer höher und damit steigen auch die Anforderungen an klinische Studien. Die Datenerhebung wird weiterhin durch noch nicht bekannte Komorbiditäten oder Spätfolgen verkompliziert, die im Vorhinein im Studiendesign nicht berücksichtigt werden können. Die Anforderung an Studien, eine möglichst genaue Zweckbindung anzugeben, kann daher nicht erfüllt werden, wenn ein solcher Zusammenhang mit der Erkrankung noch nicht bekannt ist. Für viele Patient*innen bedeutet dies aktuell den Einschluss in eine Vielzahl von Studien und das wiederholte Durchlaufen umfangreicher Interviews und Untersuchungen.

Versorgungsdaten haben das Potential, hier eine deutliche Reduktion des Aufwandes sowohl auf Seiten der Forschenden als auch der Patient*innen zu ermöglichen, da sie das gesamte Spektrum der in der medizinischen Versorgung auftretenden Erkrankungen abbilden. Somit erlauben sie auch die Untersuchung von Krankheitsinteraktionen und bergen das Potential zur Entdeckung bisher unbekannter Erkrankungen. Sie ergänzen daher klinische Studien, um neue Zusammenhänge zu entdecken und gezielte Hypothesen zu generieren.

Wesentliche Voraussetzung für die Ausschöpfung dieses Potentials ist dabei die adäquate Abbildung von Krankheitsverläufen und Behandlungspfaden durch die langfristige und sektorenübergreifende Erhebung von Versorgungsdaten. Eine solche Datensammlung kann effektiv nur auf Basis von pseudonymen Datensätzen gelingen. Die Sekundärnutzung von solchen personenbezogenen Daten ist gemäß Art. 6 der Datenschutzgrundverordnung unter bestimmten Voraussetzungen gestattet, wobei die informierte Einwilligung der Patient*innen meist das Mittel der Wahl ist. Die Herausforderung beim Broad Consent liegt darin, eine Einwilligung in die Datennutzung für noch nicht näher definierte Forschungsprojekte und einen langen Zeitraum zu erhalten im Vergleich zur zeitlich und inhaltlich begrenzten Einwilligung für eine konkrete klinische Studie. Außerdem können die Patientendaten nach Einwilligung bis zu 30 Jahre

gespeichert und für Forschungszwecke verwendet werden, sofern kein Widerruf der Einwilligung erfolgt. Eine unabhängige Ethikkommission und ein Fachgremium an jeder teilnehmenden Klinik entscheiden, ob die Daten für ein Forschungsvorhaben genutzt werden dürfen.

Die Sicht der Patienten

Wie Richter *et al.*, 2017 zeigen konnten, stehen Patienten einer breiten Einwilligung ohne Einschränkung des Forschungsgebietes grundsätzlich positiv gegenüber. Die Gründe für eine Teilnahme sind dabei primär prosozial ausgerichtet, d. h. resultieren u.a. aus Altruismus, Solidarität oder auch Verbundenheit mit dem behandelnden Arzt. Auch einer Forschungsnutzung von Versorgungsdaten ohne Einwilligung, nach Art. 9 II (j) DSGVO, stehen etwa 75% der Patienten positiv gegenüber (Richter *et al.*, 2019). Somit wäre aus Patient*innensicht theoretisch ein Verzicht auf eine informierte Einwilligung denkbar. Aus rechtlicher und ethischer Sicht sowie um eine möglichst repräsentative Abdeckung des Patient*innenkollektivs zu erreichen, ist es aber ratsam, eine informierte Entscheidungsmöglichkeit vorzusehen. Diese kann entweder als Opt-out-Verfahren oder als Broad Consent gestaltet werden – wie er im Bereich der Biobanken bereits etabliert ist, und im Rahmen der Medizininformatik-Initiative (MII) mit stärkerem Fokus auf Versorgungsdaten derzeit eingeführt wird. Dieser Ansatz scheint insbesondere deshalb ratsam, da die MII auch eine Rekontaktierung von Patient*innen für die Mitteilung von Zufallsbefunden, die Einholung zusätzlicher Informationen oder auch der Rekrutierung von Patienten für themenspezifische klinische und epidemiologische Studien vorsieht. Um das Wohlwollen der Patient*innen nicht zu verspielen, sollte eine solche erneute Kontaktaufnahme nur nach vorheriger Zustimmung erfolgen.

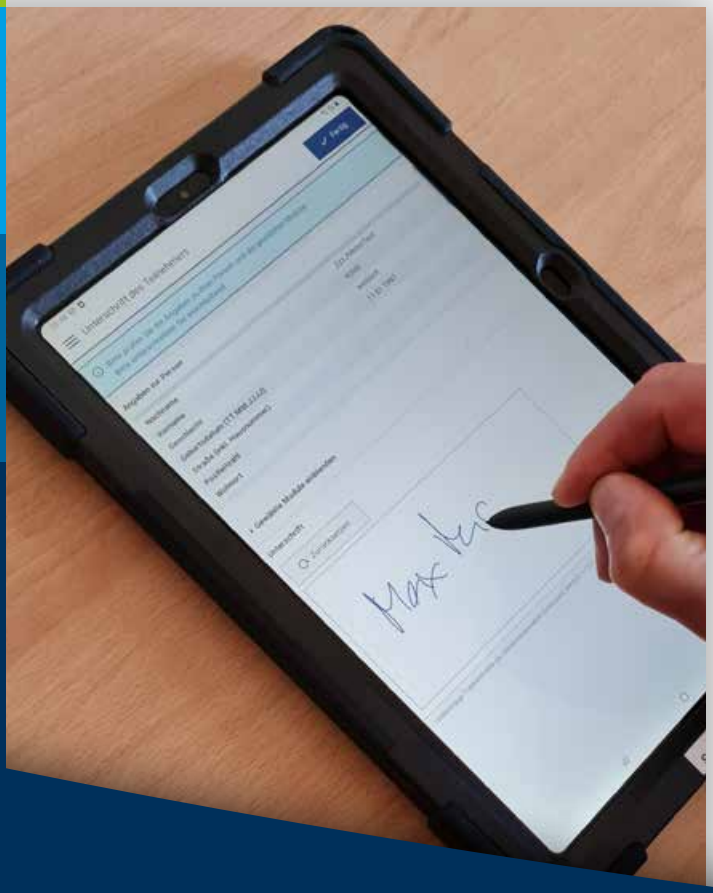


Abbildung 1: Vollständig digitale Erhebung des Broad Consent mittels des gICS an der Universitätsmedizin Greifswald (Foto: A. Stein).

Digitalisierung des Consent-Prozesses

Wesentlich für die Nutzung von Versorgungsdaten im Forschungskontext ist eine einheitliche und informierte Einwilligung der Patient*innen. Diese sollte dabei so einfach wie möglich erfolgen. Gleiches gilt natürlich auch für den Widerruf derselbigen.

Sowohl die fehlerfreie Abbildung des Patientenwillens als auch die zügige Umsetzung von Widerrufen wird durch eine vollständig elektronische Erhebung und Verarbeitung der Consent-Informationen begünstigt. Die DSGVO – und einige darauf Bezug nehmende Normen im Landesrecht MV und einigen anderen Bundesländern – sowie deren Umsetzung in deutsches Recht sehen derzeit kein explizites Schriftformerfordernis im Sinne von § 126 BGB vor. Dies schafft die Möglichkeit einer elektronischen Einwilligung (Dierks *et al.*, 2021) und gestattet die vollständig digitale Erhebung des Consents (z. B. mittels Tablet-PC, vgl. Abbildung 1).

Für die technische Implementierung des Consent-Prozesses stehen dabei eine Vielzahl von technischen Standards und eine kleine Auswahl freier, etablierter Werkzeuge, wie z. B. dem generic Informed Consent Service (Rau *et al.*, 2020), zur Verfügung. Wesentlich für standortübergreifende Projekte wie die MII ist dabei die standardisierte Abbildung der Einwilligung. Hierzu wurde im Rahmen der MII eine generische Abbildung eines modularen Consents vorgeschlagen (Bild *et al.*, 2020). Dieser Vorschlag wurde durch die aktuellen Arbeiten der MII Task Force „Consent-Umsetzung“ erweitert und wird in aktuellen Implementierungen[1] bereits berücksichtigt.

[1] https://www.ths-greifswald.de/einwilligungs-management-gics-in-neuer-version-2-13-0-verfuegbar/

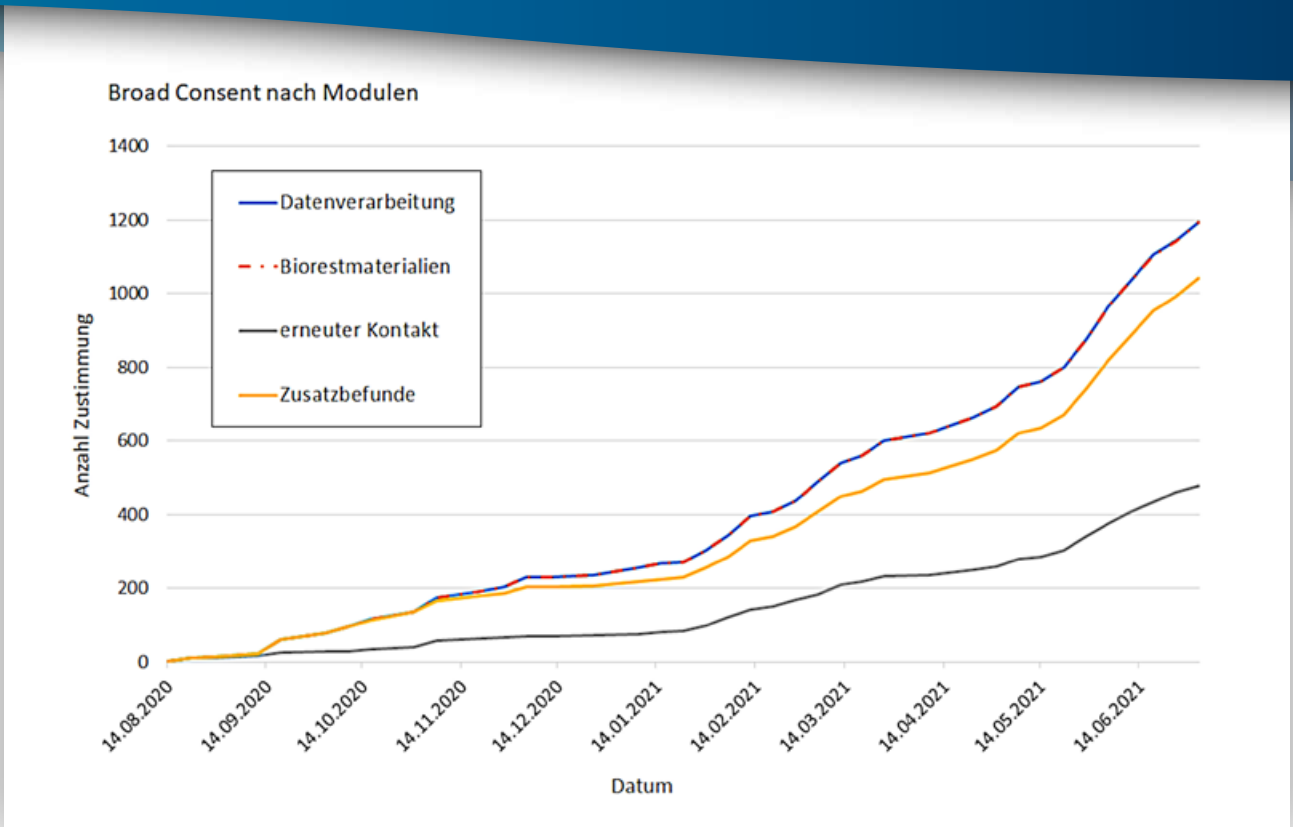


Abbildung 2: Zahl der Consente an der UMG im Zeitraum vom 14.08.2020 bis 2.07.2021, aufgeschlüsselt nach Consent-Modulen (Quelle: C. Spitzenpfeil).

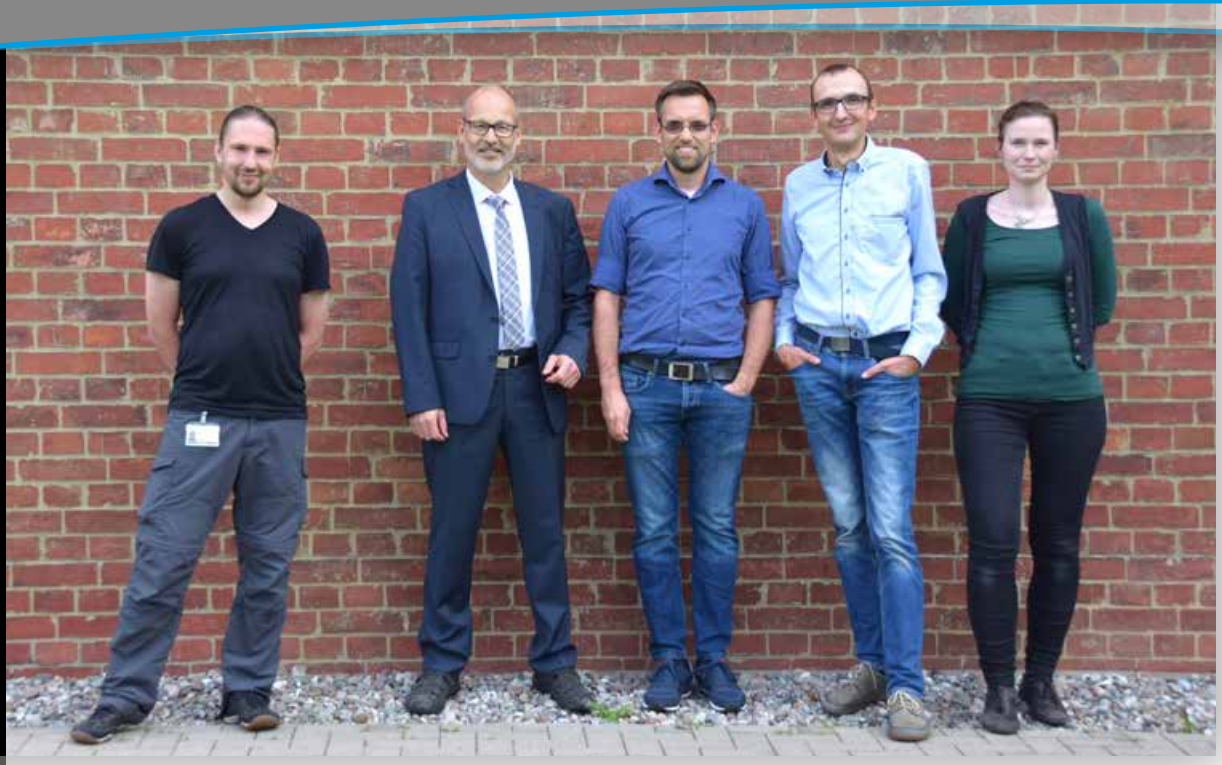
Erste Erfahrungen mit dem eBroad Consent

Im Rahmen der MII beginnen bereits mehrere deutsche Universitätskliniken damit, die Patient*innen bei der Aufnahme ins Krankenhaus um diese elektronische Einwilligung (eBroad Consent) zu bitten. An der Universitätsmedizin Greifswald (UMG) haben wir bereits im August 2020 damit begonnen. Diesem wichtigen Meilenstein vorausgegangen war eine Untersuchung, in welchem Setting für Patient*innen eine freie und unbeeinflusste Entscheidung zum Broad Consent möglich ist. Konkret wurde untersucht, inwieweit eine räumliche und personelle Trennung zum Aufnahmeprozess für die freie Willensbildung auf Seiten der Patient*innen erforderlich wäre. Die Untersuchung hat gezeigt, dass aus Sicht der Patient*innen eine Erhebung des Broad Consent im direkten Anschluss an die stationäre Aufnahme eine uneingeschränkte freie Entscheidung ermöglicht. Wesentlich hierbei ist die klare Trennung von Aufnahme, als zwingende Voraussetzung für die Behandlung, und Aufklärung zum Broad Consent, als freiwilliger Zusatz. Hierfür ist eine entsprechende Gesprächsführung ausreichend.

Um den Aufklärungsprozess zeitlich möglichst schlank zu halten und den Aufwand für die Mitarbeiter*innen der zentralen Patientenaufnahme als auch der Patient*innen zu reduzieren, wird der Aufklärungsfilm der MII zum Broad Consent im Wartebereich der zentralen Patientenaufnahme gezeigt. Die Erfahrung zeigt, dass so bereits eine gute Vorbefassung auf Seiten der Patient*innen mit dem Consent erreicht werden kann.

In dieser Weise konnte die UMG bereits über 1000 Broad Consents mit einem überschaubaren Aufwand auf allen Seiten erheben. Als Grundlage für zukünftige Forschungsprojekte ist es interessant, dass eine hohe Zustimmung zur Verwendung von Versorgungsdaten und Restbiomaterialien besteht, wohingegen eine Rekontaktierung für die Erhebung von zusätzlichen Daten bzw. den Einschluss in weiterführende Studien aktuell nur von ca. einem Drittel der einwilligenden Patienten gewünscht wird (vgl. Abbildung 2).

Die Erfahrungen an der UMG zeigen, dass die Erhebung des Broad Consents sich gut in den Prozess der stationären Patientenaufnahme integrieren lässt. Für den Erfolg des Broad Consents wird es aber wesentlich sein, nicht nur die stationär behandelten Patienten einzuschließen, sondern möglichst das gesamte Patientenaufkommen abzudecken. Entsprechende Feldversuche in ausgewählten Ambulanzen der UMG sind derzeit in Vorbereitung.



Die Autoren v.l.n.r.: Clemens Spitzenpfeil, Prof. Dr. Wolfgang Hoffmann, Dr. Martin Bialke, Dr. Torsten Leddig, Dipl.-Jur. Alexandra Stein (Foto: Martin Bialke).

Referenzen:

Richter G. *et al.* (2017): Broad consent for health care-embedded biobanking: understanding and reasons to donate in a large patient sample. Genetics in Medicine, Volume 20, Number 1

Richter G. *et al.* (2019): Patient views on research use of clinical data without consent: Legal, but also acceptable?. European Journal of Human Genetics (2019) 27:841-847

Rau H. *et al.* (2020): The generic Informed Consent Service gICS ®: implementation and benefits of a modular consent software tool to master the challenge of electronic consent management in research. Journal of Translational Medicine (2020) 18(1):287

Dierks C. *et al.* (2021): Data Privacy in European Medical Research: A Contemporary Legal Opinion. (Berlin: Schriftenreihe der TMF, Band 18)

Bild R. *et al.* (2020): Towards a comprehensive and interoperable representation of consent-based data usage permissions in the German medical informatics initiative. BMC Medical Infromatics and Decision Making (2020) 20:103

Kontakt:

Prof. Dr. Wolfgang Hoffmann, MPH
Abteilungsleiter
Institut für Community Medicine,
Abt. Versorgungsepidemiologie und Community Health
Universitätsmedizin Greifswald
wolfgang.hoffmann@uni-greifswald.de

www2.medizin.uni-greifswald.de/icm/index.php?id=18

Über die Abteilung Versorgungsepidemiologie und Community Health

Die Abteilung Versorgungsepidemiologie und Community Health (ICMVC) am Institut für Community Medicine der Universitätsmedizin Greifswald befasst sich u. a. mit Themen des Datenschutzes und dessen Umsetzung in der medizinischen Forschung. Hierzu hat das ICMVC bereits 2014 die unabhängige Treuhandstelle an der Universitätsmedizin Greifswald etabliert, und die notwendigen Werkzeuge für ID-, Pseudonym- und Consent-Management erstellt. Begleitend zur technischen Unterstützung des Consent-Prozesses hat das ICMVC bereits frühzeitig an der Etablierung eines Broad Consent am Standort gearbeitet. Als Konsortialpartner im MIRACUM Konsortium fungiert das ICMVC als Consent Competence Center.

lost in translation? semantische standardisierung!

Terminologien wie LOINC und SNOMED CT reduzieren Babylonische Sprachverwirrung

von Cora Drenkhahn und Josef Ingenerf

Immer mehr Daten aus der Krankenversorgung werden digital abgelegt und bergen damit einen Schatz an verborgenen Erkenntnissen für Qualitätssicherung und Forschung. Jedoch unterbinden fehlende semantische Standardisierung und Mängel derzeit genutzter Codiersysteme deren zielführende Auswertbarkeit. Soll dieser Wissensschatz zukünftig gehoben werden, ist eine bedeutungserhaltende Codierung mit den im Folgenden thematisierten Terminologien wie LOINC oder SNOMED CT unumgänglich. In Kombination mit dem international aufstrebenden FHIR-Standard entstehen aktuell vielversprechende interoperable Softwarelösungen. Wir möchten aufzeigen, warum sich die Investition in die Verwendung von Codes aus anspruchsvolleren Terminologien lohnt.

In der Routineversorgung des Gesundheitswesens fallen kontinuierlich beachtliche Mengen medizinischer Daten zu hunderttausenden Einzelfällen an. Diese werden mit der fortschreitenden Digitalisierung auch zunehmend elektronisch dokumentiert und kommuniziert. Hierzu gehören wesentliche Kerninformationen wie Diagnosen, Prozeduren, Laborwerte und Medikamente, die im Rahmen der Krankenversorgung dokumentiert und abgerechnet werden. Idealerweise werden diese und weitere Informationen für andere Zwecke wiederverwendet. Hierdurch ist einerseits eine Qualitätssteigerung der Versorgung für den Patienten selbst möglich, indem unnötige Mehrfachuntersuchungen und Behandlungsfehler durch Unkenntnis bestehender Erkrankungen oder Medikationen vermieden werden – sofern die benötigten Informationen zur richtigen Zeit am richtigen Ort in geeigneter Form bereitstehen. Andererseits ergibt sich langfristig eine Verbesserung der Behandlungsoptionen durch eine übergreifende Nutzung der Falldaten zu Forschungszwecken – das Einverständnis des Patienten vorausgesetzt. Durch die Zusammenführung großer Datenmengen und ihre gemeinsame Auswertung wird hier das Erkennen übergreifender Entwicklungen und die Ableitung krankheitsrelevanter Zusammenhänge ermöglicht, was unweigerlich zu neuen wissenschaftlichen Erkenntnissen führt, die letztlich wieder dem einzelnen Patienten zugutekommen (Safran *et al.*, 2007).

Variabilität natürlicher Sprache behindert Weiterverwendbarkeit erhobener Daten

Dieses immense Potential prinzipiell verfügbarer Gesundheitsdaten bleibt jedoch bislang weitgehend ungenutzt. Ein wesentlicher Grund liegt in der Unschärfe unserer Sprache im Allgemeinen, aber auch der medizinischen Fachsprache im Besonderen. Insbesondere Abkürzungen wie HWI tragen zu einer ***babylonischen Sprachverwirrung*** bei. Die mehrdeutige Abkürzung (Homonymie) lässt sich je nach Kontext als Harnwegsinfekt oder als Hinterwandinfarkt interpretieren. Umgekehrt lassen sich Sachverhalte wie ein Hautabszess mehrfach umschreiben (Synonymie), etwa mit „Apostem“, „Abszess an der Haut“ oder „Abscess of skin“. Mehrsprachige Synonyme sind insbesondere für die internationale Forschung relevant. Viele weitere Phänomene wie Quasi-Synonymie (z. B. „Abszess“ versus „Furunkel“: Ist das dasselbe?) oder Negation (z. B. „kein ausgedehnter Abszess“: Liegt kein Abszess vor oder einer, der nicht ausgedehnt ist?) erschweren eine zuverlässige Interpretation sprachlicher Daten. Das gilt für Menschen. Das gilt aber vor allem für Rechner, um deren Einsatz es hier geht.

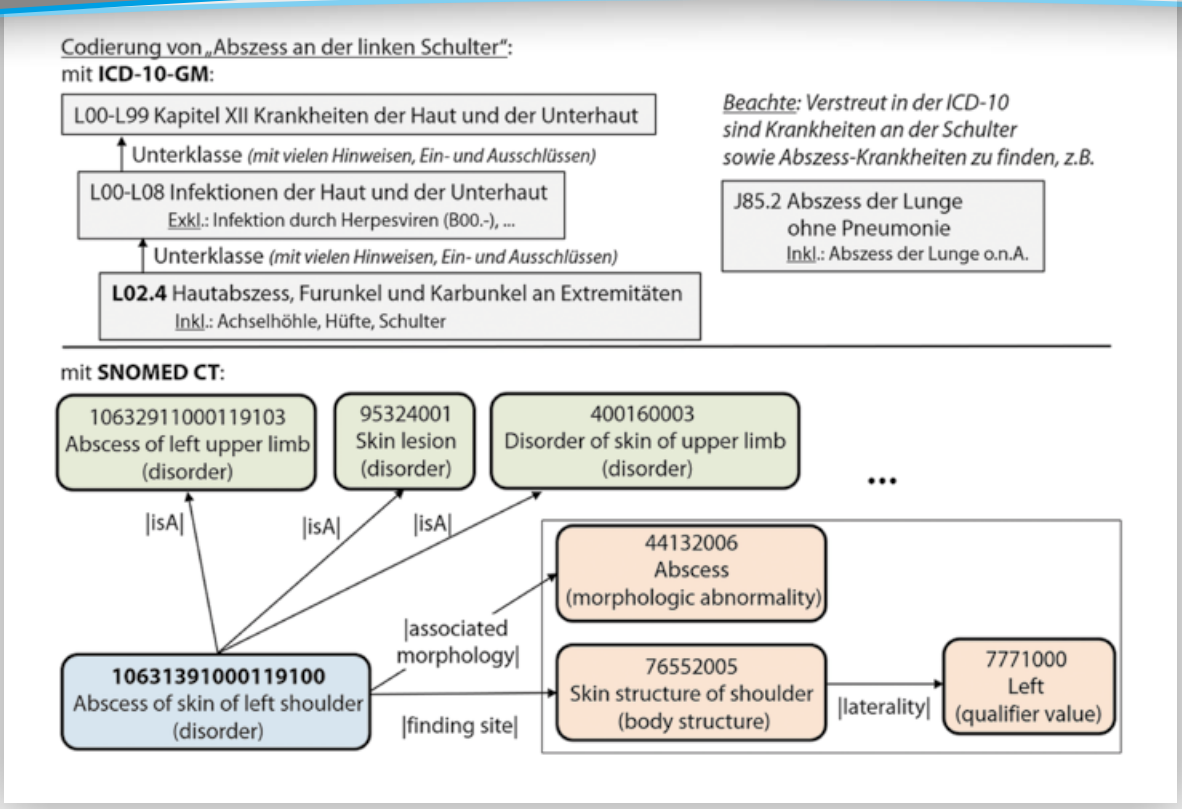


Abbildung 1: ICD-10-Klassifikation im Vergleich zur ausdrucksstärkeren SNOMED CT-Terminologie (Quelle: Erstellt von Josef Ingenerf).

Insofern erschwert die verbreitete Verwendung von Freitext in der ärztlichen Dokumentation eine erwünschte Weiterverwendung von Versorgungsdaten etwa für die Forschung. Entsprechend ist ein zentrales Ziel der Medizininformatik-Initiative (MII), die Nutzung von Versorgungsdaten für die Forschung zu erleichtern (Semler *et al.*, 2018). Auch wenn fortschrittliche Verfahren des Natural Language Processing (NLP), Text Minings und Maschinellen Lernens inzwischen durchaus Erfolge erzielen, können die wesentlichen Informationen aus einer freitextlichen Dokumentation momentan nicht zuverlässig korrekt und vollständig extrahiert werden (Névéol *et al.*, 2018). Angesichts der genannten Eigenheiten natürlicher Sprache überrascht dieser Befund jedoch nicht: Synonymie und Homonymie, implizit enthaltener Kontext, Abkürzungen und Negationen – die gerade in der Medizin so beliebte Flexibilität, einen Sachverhalt in individueller Ausdrucksweise zu beschreiben, wird hier zum Verhängnis.

Warum die Codierung mit ICD-10 keine Lösung ist

Momentan werden in Deutschland Klassifikationen wie die **ICD-10-GM** (International Classification of Diseases, 10th Revision, German Modification) gesetzlich vorgeschrieben. Hiermit codierte Diagnosen werden primär für Abrechnungszwecke oder statistische Auswertungen verwendet. Um tatsächlich eine vielfältige Weiterverwendbarkeit der Daten zu erreichen, ist diese Praxis jedoch ungenügend. Die vordefinierten ca. 13.000 Klassen decken den gesamten Anwendungsbereich vollständig und untereinander überschneidungsfrei ab, sodass eine eindeutige Zuordnung einer jeden Diagnose in genau eine ICD-10-Klasse erreicht wird. Obwohl nützlich für die Statistik – Mehrfachzählungen desselben Sachverhalts sind ausgeschlossen – bedeutet die Klassenbildung einen Informationsverlust, der andere Arten der Auswertung erschwert oder verhindert. Außerdem ist die inhaltliche Bedeutung der Klassen nur sehr begrenzt maschinell auswertbar, da insbesondere ihre Ein-/Ausschlusskriterien nur durch Menschen interpretierbar sind. In Abbildung 1 erkennt man in der oberen Hälfte die Klasse „L02.4“, die eine recht unscharfe „Schublade“ repräsentiert, wo auch verwandte Krankheitsbilder (z. B. „Furunkel“) eingeschlossen sind, aber eine Verursachung durch Herpesviren ausgeschlossen ist. Die exakte Semantik der Klasse ist für Rechner „unsichtbar“, da ausschließlich im Text notiert. Ein Zugriff auf alle Abszess-Krankheiten ist genauso wenig möglich wie eine Präzisierung der Anatomie einschließlich der Seitigkeit.

Nur semantische Standardisierung mit Terminologien sichert Interoperabilität und Weiterverwendbarkeit erhobener Patientendaten

Somit kann durch eine Codierung mit einfachen Klassifikationen auch keine semantische Interoperabilität erreicht werden. Für eine flexible, *zweckungebundene* Standardisierung ist vielmehr eine möglichst präzise Darstellung der medizinischen Inhalte nötig, welche optimalerweise auch weitergehend maschinenverwertbar ist. Diesen Bedarf erfüllen verschiedene **Terminologien**, von denen LOINC und SNOMED CT zwei der

LOINC	LangBezeichner	Komponente	Eigenschaft	Timing	System	Skala	Methode	Klasse
595-9	Bacteria identified in Abscess by Aerobe culture	Bakterien identifiziert	Prid	Pkt	Abs	Nom	Aerobe Kultur	MIKRO

Abbildung 2: LOINC-Code mit Bezeichnung, sechs Achsen mit Teilbedeutungen sowie eine grobe Klasseneinteilung (hier: Mikrobiologie) (Quelle: https://loinc.org/search/).

bekanntesten und verbreitetsten Vertreter darstellen. Durch die weltweite Verbreitung dieser Standards können die Daten sogar international kommuniziert und zu computergestützten Analysen herangezogen werden. Bei der medizinischen Dokumentation ist trotzdem nicht der Umgang mit obskuren Zahlenreihen zu befürchten; innerhalb der Standards sind die identifizierenden Codes mit natürlichsprachlichen Bezeichnungen für den menschlichen Anwender verknüpft. Durch eine gute Integration in (bestehende) Softwaresysteme kann ebenso die Dokumentation unterstützt und Fehleingaben vorgebeugt werden.

LOINC (Logical Observation Identifiers Names and Codes) ist ein internationaler, frei verfügbarer Standard des US-amerikanischen Regenstrief Institute und beschränkt sich primär auf die standardisierte Repräsentation von Labortests und anderen klinischen Messungen. Ein LOINC-Code steht hierbei für die eindeutige Kombination von mindestens fünf (in Abbildung 2 hervorgehobenen) Aspekten, die die Eigenschaften des Tests explizit darlegen.

Von besonderer Bedeutung und Aktualität ist außerdem SNOMED CT (SNOMED Clinical Terms) (Drenkhahn and Ingenerf, 2020). Anfang dieses Jahres wurde Deutschland Mitglied der herausgebenden Organisation SNOMED International, sodass die Terminologie nun national lizenziert verwendbar ist. Im Gegensatz zu anderen thematisch begrenzten semantischen Standards gilt SNOMED CT als Referenzterminologie, die den medizinischen Bereich möglichst umfassend abbilden soll. Außerdem verfügt SNOMED CT über einen komplexen Aufbau sowie besondere Eigenschaften, wodurch gemeinsam fortschrittliche Funktionalitäten ermöglicht werden.

Das Kernstück von SNOMED CT bilden die sogenannten Concepts, die jeweils einen für die Medizin relevanten Sachverhalt als abstrahiertes Konstrukt anhand eines eindeutigen Identifiers repräsentieren. Für jedes dieser ca. 350.000 Konzepte sind aber auch mehrere natürlichsprachliche Bezeichnungen (Descriptions) enthalten, die als Synonyme für Suchanfragen etwa bei der Dateneingabe sehr hilfreich sind. Das wahre „Schwergewicht" unter den Terminologien wird SNOMED CT aber durch eine dritte Komponente: Mithilfe von Relationships gelingt eine maschinenverarbeitbare Definition der inhaltlichen Bedeutung der Konzepte. Durch diese logische Formalisierung entstehen komplexe Verknüpfungen der Konzepte untereinander (siehe unteren Abschnitt in Abbildung 1), welche sich anschließend auch für computergestützte Ableitungen von Schlussfolgerungen (mithilfe eines Beschreibungslogik-Reasoners) nutzen lassen. Entsprechend bietet SNOMED CT beste Voraussetzungen auch für komplexe Datenauswertungen und eine vielfältige Wiederverwendung einmal erhobener Informationen.

Aktuelle Entwicklungen: FHIR bringt Terminologien voran

Wie sich nun bereits abzeichnet, ist die fehlende Integration semantischer Standards in die klinischen Primärsysteme an sich nicht der einzige Grund für ihren schleppenden Einsatz. Insbesondere die Komplexität und Heterogenität der verschiedenen Terminologien erschwert bisher eine flächendeckende Umsetzung (Rector, 1999). Aktuell zeichnen sich hier jedoch vielversprechende Entwicklungen ab: Auf der Ebene des Informationsmodells sorgt der Strukturstandard **HL7 FHIR** (Health Level 7 – Fast Healthcare Interoperability Resources) seit einigen Jahren für neuen Schwung. Neben der zentralen Forderung, semantische Standards einzubinden und der Bereitstellung einfacher Mechanismen hierfür, bietet FHIR auch ein spezielles Modul für Terminology Services. Mit den hier definierten Datenstrukturen und Diensten lassen sich die Terminologie-Inhalte sowohl einheitlich repräsentieren als auch anfragen. Somit wird eine Auslagerung der Semantikkomponenten auf einen dedizierten Terminologieserver ermöglicht, mit welchem die betreffenden Anwendungen über diese standardisierten Dienste niederschwellig kommunizieren können, wie in Abbildung 3 dargestellt.

Mehrere aktuelle Projekte und Initiativen demonstrieren bereits, dass aus der Kombination von Semantik- und Strukturstandards sowie Terminologieserver synergistische Vorteile zu erzielen sind. So verwendet etwa das Konsortium HiGHmed der MII in allen drei Anwendungsfällen „Cardiology", „Oncology" und „Infection Control" den FHIR-basierten „Ontoserver" der australischen Firma CSIRO als zentralen Knotenpunkt für die Einbindung semantischer Inhalte (Haarbrandt *et al.*, 2018). Gleiches gilt für das Projekt „CODEX" des Netzwerks der Universitätsmedizin (NUM), in dem eine nationale Forschungsdatenplattform zu Covid-19 etabliert werden soll. Auch auf Ebene

Abbildung 3: Anwendungen nutzen semantische Standards über FHIR-konforme Terminologieserver
(Erstellt von Cora Drenkhahn Quellen: https://www.informatik.tu-darmstadt.de/gris/forschung_1/visual_search_and_analysis/_visual_search_and_analysis_projects_1/project_highmed_/project_highmed___interactive_visualization_for_infection_control_1.en.jsp, https://www.hl7.org/fhir/overview.html, https://de.cleanpng.com/png-j1o44m/, https://loinc.org/, https://commons.wikimedia.org/wiki/File:Snomed-logo.png).

der Kassenärztlichen Bundesvereinigung wird mit den MIOs (Medizinische Informationsobjekte) zunehmend auf eine Standardisierung mittels FHIR einerseits und Terminologien – mit Fokus auf SNOMED CT – andererseits gesetzt.

Zusammenfassend lässt sich sagen, dass in jüngster Zeit dringend erforderliche Fortschritte hinsichtlich semantisch interoperabler Anwendungssysteme zu verzeichnen sind. Deutschland ist zu Beginn dieses Jahres als 40ste Nation der non-profit-Organisation „SNOMED International" beigetreten. Unter Verwendung struktureller Standards wie HL7 FHIR werden Datenstrukturen (inkl. Codes aus Semantikstandards wie LOINC, SNOMED CT, usw.) für den Datenaustausch definiert; sowohl für gesetzlich geforderte amtliche Aufgaben (z. B. Impfpass) als auch für die oben genannten Forschungsprojekte wie MII und NUM. Somit sollten Daten beider „Welten" zunehmend austauschbar und weiterverwendbar werden, zum Wohle der Patientenversorgung als auch der Forschung.

Referenzen:

Drenkhahn, C., and Ingenerf, J. (2020). Referenzterminologie SNOMED CT. Forum der Medizin_Dokumentation und Medizin_Informatik 22, 81–84.

Haarbrandt, B., Schreiweis, B., Rey, S., Sax, U., Scheithauer, S., Rienhoff, O., Knaup-Gregori, P., Bavendiek, U., Dieterich, C., Brors, B., *et al.* (2018). HiGHmed – An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. Methods Inf Med 57, e66–e81.

Névéol, A., Dalianis, H., Velupillai, S., Savova, G., and Zweigenbaum, P. (2018). Clinical Natural Language Processing in languages other than English: opportunities and challenges. J Biomed Semantics 9, 12.

Rector, A.L. (1999). Clinical terminology: why is it so hard? Methods Inf Med 38, 239–252.

Safran, C., Bloomrosen, M., Hammond, W.E., Labkoff, S., Markel-Fox, S., Tang, P.C., and Detmer, D.E. (2007). Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. Journal of the American Medical Informatics Association 14, 1–9.

Semler, S.C., Wissing, F., and Heyder, R. (2018). German Medical Informatics Initiative: A national approach to integrating health data from patient care and medical research. Methods of Information in Medicine 57, e50.

Kontakt:

Cora Drenkhahn
Wissenschaftliche Mitarbeiterin
Institut für Medizinische Informatik |
IT Center for Clinical Research
Universität zu Lübeck
c.drenkhahn@uni-luebeck.de

www.imi.uni-luebeck.de

Das Forschungsprojekt in Kurzform

2019 hat sich der Standort Lübeck des Universitätsklinikums Schleswig-Holstein (UKSH) dem HiGHmed-Konsortium der Medizininformatik-Initiative (MII) angeschlossen. Im Anwendungsfall „Infection Control" arbeiten sieben Partnerstandorte gemeinsam an einem Früherkennungssystem für nosokomiale Infektionsausbrüche. Das Lübecker Team unterstützt hierbei vor allem beim Terminologiemanagement und der Anbindung eines Terminologieservers. Mit einem ähnlichen Schwerpunkt engagiert sich der Standort Lübeck seit 2020 im Teilprojekt CODEX (COVID-19 Data Exchange Platform) des Netzwerks der Universitätsmedizin (NUM).

„eine forschende versorgung wäre längst möglich“

Interview mit Dr. Martin Hager

Medizinischer Leiter im Bereich Personalisierte Gesundheitsversorgung (PHC) bei der Roche Pharma AG

Eine personalisierte Gesundheitsversorgung, die mit jeder Behandlung heute neues Wissen für die Zukunft generiert, wäre längst möglich – davon ist Dr. Martin Hager fest überzeugt. Im Interview spricht er über aktuelle Chancen und Herausforderungen und betont, warum Vernetzung und Austausch dabei eine zentrale Rolle einnehmen.

gesundhyte.de: *Herr Hager, Roche steht für personalisierte Medizin – wie fest verankert ist diese bereits in unserer medizinischen Versorgung?*

Dr. Martin Hager: Zunächst einmal steht außer Frage, dass im Bereich der personalisierten Medizin in den letzten Jahren enorme Fortschritte erzielt wurden. Schauen wir beispielsweise in die Onkologie, die hier zweifellos eine Vorreiterrolle einnimmt: Hier steht uns heute bereits eine Vielzahl an Biomarker-basierten Therapien zur Verfügung. Und es kommen jedes Jahr neue Therapien hinzu, die immer spezifischer auf immer kleinere Patientengruppen zugeschnitten sind. Wir sehen hier fantastische Fortschritte, die sich in Behandlungserfolge übersetzen, die vor wenigen Jahren noch illusorisch erschienen. Zu unserer Realität gehört aber auch, dass diese Fortschritte bislang nicht bei allen Patientinnen und Patienten ankommen und personalisierte Medizin in der breiten Versorgung nicht in dem Maße verankert ist, wie es heute möglich wäre.

„Wir müssen heute gemeinsam Lösungen entwickeln, die einzig auf den Nutzen für Patientinnen und Patienten zielen.“

Was sind die Gründe?

Die Gründe dafür sind vielfältig, liegen zu oft aber auch in Versorgungsstrukturen, die mit der heutigen Innovationsgeschwindigkeit und dem technologischen Fortschritt einfach nicht mehr Schritt halten. Schauen wir zum Beispiel auf die molekulare Diagnostik: Ein umfassendes Tumorprofiling mittels Next Generation Sequencing ist in der Routineversorgung nach wie vor die Ausnahme. Und hier müssen wir uns schon fragen: Was bringt es uns, wenn wir immer mehr personalisierte Therapien haben, die sich gegen seltene Alterationen richten, wenn wir in der breiten Versorgung nicht die Patien-

Foto: Roche

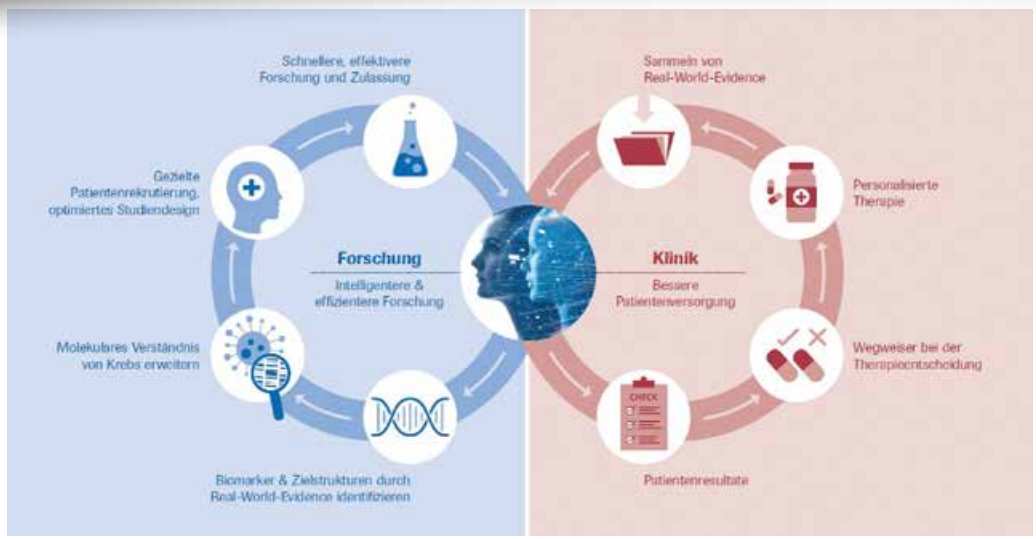


Abbildung 1: Wissen generierende Versorgung (Quelle: Roche).

tinnen und Patienten identifizieren, die von diesen Therapien profitieren können? Hier braucht es mehr Mut zu Veränderungen - wir müssen heute gemeinsam Lösungen entwickeln, die einzig auf den Nutzen für Patientinnen und Patienten zielen.

Wie kann dies gelingen?

Es gibt bereits sehr gute Initiativen, die Lösungen aufzeigen - beispielsweise, um Spitzenmedizin aus den Zentren heraus in die breite Versorgung zu tragen. Ich denke hier unter anderem an das Netzwerk Hauptstadt Urologie, das genau dieses Ziel verfolgt - und gleichzeitig auch zeigt, wie wichtig in diesem Kontext die Zusammenarbeit der verschiedenen Player im Gesundheitssystem ist. So arbeiten wir aktuell - am Beispiel des Prostatakarzinoms - zusammen mit dem Netzwerk an Lösungen für eine niederschwellige und vom Wohnort unabhängige Versorgung im Sinne der Präzisionsonkologie. Das gemeinsame Projekt reicht von der digitalen Unterstützung qualitätsgesicherter Versorgungsprozesse im Kontext der Präzisionsdiagnostik bis hin zu Fragen von Datenschutz und Datensicherheit.

Stichwort Digitalisierung: Welche Rollen spielt diese im Kontext der personalisierten Medizin?

„Die Möglichkeiten, eine datenbasierte, personalisierte Gesundheitsversorgung zu realisieren, waren nie größer als heute."

Die digitale Transformation unseres Gesundheitswesens und die konsequente Nutzung von Gesundheitsdaten spielen nicht nur eine entscheidende Rolle, sie sind vielmehr die Voraussetzung, um zu einer personalisierten Medizin im eigentlichen Sinne zu kommen: Nämlich einer personalisierten Gesundheitsversorgung, die tatsächlich die Gesundheit des Einzelnen in den Mittelpunkt rückt.

Das klingt ziemlich visionär ...

Nein, absolut nicht. Die Technologien sind dafür heute längst vorhanden. Denken wir an moderne Informationstechnologien, die Millionen von Datensätzen strukturieren und vernetzen, an digitale Gesundheitsanwendungen, die heute über Wearables oder das Smartphone individuelle Gesundheitsdaten in Echtzeit erfassen, oder an künstliche Intelligenz, die uns bei der Analyse all dieser Informationen unterstützt. Die Möglichkeiten, eine datenbasierte, personalisierte Gesundheitsversorgung zu realisieren, waren nie größer als heute.

Ein zentraler Aspekt ist dabei auch der Austausch von Gesundheitsdaten aus der Versorgung.

Absolut, Daten aus der Versorgung spielen dabei eine ganz zentrale Rolle. Wir gehen heute beispielsweise davon aus, dass etwa 95% der potentiell vorhandenen Daten in der Onkologie im Rahmen der Routineversorgung generiert werden. Wenn es uns gelingt, diese Daten gemäß gemeinsamer Qualitätsstandards strukturiert zu erfassen, miteinander zu vernetzen und auszutauschen, können wir eine forschende Versorgung etablieren, die mit jeder Behandlung heute neues Wissen für die Zukunft generiert.

Roche gestaltet die Personalisierte Medizin (Quelle: Roche).

Was braucht es dafür?

Natürlich braucht es dafür eine entsprechende Infrastruktur jenseits von Insellösungen sowie Rahmenbedingungen, die den sicheren Austausch anonymisierter und pseudonymisierter Gesundheitsdaten überhaupt erst ermöglichen. Hier ist es in unseren Augen auch wichtig, dass die private Forschung als gleichberechtigter Partner im Innovationsprozess anerkannt wird und gleichermaßen auf Daten zu Forschungszwecken zugreifen kann. Vor allem braucht es aber auch den gemeinsamen Willen aller Beteiligten im Gesundheitssystem, über Partikularinteressen hinweg aufeinander zuzugehen. Dass die Zusammenarbeit und der Wissenstransfer über Sektorengrenzen hinweg möglich und vor allem auch wertvoll ist, zeigt beispielsweise auch eine Kooperation, die wir Ende letzten Jahres mit der LMU eingegangen sind: Hier wollen wir gemeinsam zeigen, dass man aus anonymisierten klinisch-genomischen Daten aus der Versorgung für die Behandlung zukünftiger Patientinnen und Patienten lernen kann. Ganz konkret entwickeln wir zusammen einen Prädiktionsalgorithmus, mit dem wir anhand von Daten vorhersagen können, welche Patientinnen und Patienten wann von einer molekularen Testung profitieren. Ich bin überzeugt davon, dass diese Formen der Zusammenarbeit die Zukunft sind, denn nur so können wir gemeinsam die Zukunft der Gesundheitsversorgung im Sinne von Patientinnen und Patienten gestalten.

„Hier ist es in unseren Augen auch wichtig, dass die private Forschung als gleichberechtigter Partner im Innovationsprozess anerkannt wird und gleichermaßen auf Daten zu Forschungszwecken zugreifen kann."

„Die digitale Transformation unseres Gesundheitswesens und die konsequente Nutzung von Gesundheitsdaten spielen nicht nur eine entscheidende Rolle, sie sind vielmehr die Voraussetzung, um zu einer personalisierten Medizin im eigentlichen Sinne zu kommen."

Das Interview führte: Franziska Müller

Kontakt:

Dr. Martin H. Hager, MBA

Medical Therapeutic Area Lead Personalized Healthcare (PHC)

Roche Pharma AG

martin.hager@roche.com

www.roche.de/ueber-roche/was-uns-antreibt/personalisierte-medizin

zwischen individualisierten therapieoptionen und gläsernen patienten

Wie Patientenvertretungen Chancen und Herausforderungen von KI in der Medizin bewerten

von Klemens Budde, Matthieu Schapranow, Elsa Kirchner und Thomas Zahn – Plattform Lernende Systeme

Lernende Systeme, die Algorithmen der Künstlichen Intelligenz (KI) verwenden, können PatientInnen, ÄrztInnen sowie Pflegefachkräfte bei der Prävention, frühzeitigen Diagnose oder patientenindividuellen Therapien wirksam unterstützen. Eine qualitative Befragung zeigt, dass InteressensvertreterInnen von PatientInnen KI im Gesundheitswesen Potenziale für eine stärker personalisierte Behandlung sowie eine umfassendere und schnellere Diagnosefindung attestieren. Herausforderungen stellen Datenmissbrauch, Cyberkriminalität, durch KI-Systeme getroffene fehlerhafte und diskriminierende Entscheidungen sowie mangelhafte Datensätze und schlecht optimierte Algorithmen dar. Hierzu wünschen sich die ExpertInnen eine stärkere Einbindung von Betroffenen in Forschung, Entwicklung und Anwendung von KI-Technologien.

Die Debatte um einen digitalen Impfnachweis zeigte, dass Deutschland bei der Digitalisierung des Gesundheitswesens noch vor großen Herausforderungen steht. Die zunehmende Digitalisierung ist nicht nur für die Bewältigung der Corona-Pandemie wichtig, sondern auch die zentrale Voraussetzung für den Einsatz von Künstlicher Intelligenz in der Gesundheitsfürsorge. Die Nutzung von patientenindividuellen medizinischen Daten und KI-Assistenzsystemen kann künftig bei der Prävention, frühzeitigen Diagnose sowie bei individualisierten Therapien zu besseren Behandlungsergebnissen führen, die Entdeckung neuer medizinischer Zusammenhänge und innovativer Präventionsansätze ermöglichen und somit unsere Gesundheitsversorgung verbessern. Der Einsatz Lernender Systeme, die abstrakt beschriebene Aufgaben auf Basis von Daten durch Maschinelles Lernen (ML) selbstständig lernen, kann medizinisches und pflegerisches Personal entlasten und PatientInnen im Alltag nachhaltig unterstützen. Gleichzeitig ergeben sich daraus hohe Anforderungen an die IT-Sicherheit der Systeme, um bei Betroffenen sowie medizinischem und pflegerischem Personal Vertrauen in die neuen Technologien zu schaffen. Dafür gilt es, Hemmschwellen im Umgang mit

Publikation „KI in der Medizin und Pflege aus der Perspektive Betroffener“
Der Beitrag basiert auf dem Tagungsbericht der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme. Der Tagungsbericht eines Runden Tischs mit VertreterInnen verschiedener Betroffenengruppen fasst verschiedene Wahrnehmungen von Chancen und Herausforderungen für den Einsatz von KI im Gesundheitswesen zusammen.

Klemens Budde *et al.* (Hrsg.): KI in der Medizin und Pflege aus der Perspektive Betroffener. Tagungsbericht zum Runden Tisch mit Patientenvertretungen aus der Plattform Lernende Systeme, München 2020. Online verfügbar unter: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG6_Whitepaper_Medizin_Pflege_Tagungsbericht.pdf

Die zentralen Studienergebnisse wurden auch im Deutschen Ärzteblatt vorgestellt. Online unter: https://www.aerzteblatt.de/archiv/216998/Kuenstliche-Intelligenz-Patienten-im-Fokus

Die Plattform-Mitglieder Elsa Kirchner (DFKI; im Hintergrund) und Klemens Budde (Charité – Universitätsmedizin Berlin; rechts) diskutieren mit den PatientenvertreterInnen (Quelle: Plattform Lernende Systeme).

KI-basierten Assistenzsystemen zu erkennen und Bedürfnisse Betroffener bereits während der Entwicklung zu berücksichtigen. Die Perspektive Betroffener ist bisher bei der Analyse von Potenzialen und Herausforderungen der Digitalisierung des Gesundheitswesens und dem Einsatz von KI-Systemen nur teilweise berücksichtigt worden – meist liegt der Fokus auf technischen Potenzialen. Bei allen technologischen Fortschritten müssen aber die Interessen von PatientInnen und Pflegebedürftigen im Mittelpunkt stehen – auch im Hinblick auf die Qualitätssicherung und -verbesserung sowie die Zufriedenheit mit der medizinischen Betreuung. Daher müssen PatientInnen in die unterschiedlichen Prozesse einbezogen werden, um Behandlungsabläufe optimal auf die Bedürfnisse zuzuschneiden.

Die Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme hat dazu gemeinsam mit dem Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) und der Charité Berlin eine qualitative ExpertInnenbefragung von Patientenvertretungen in Deutschland durchgeführt. Zielsetzung war, Chancen und Herausforderungen beim Einsatz von KI-Systemen im Gesundheitswesen aus Sicht von PatientenvertreterInnen zu identifizieren. Dazu wurde die Wahrnehmungen der PatientenvertreterInnen – in Anlehnung an die Methode der Delphi-Befragung – mit Hilfe eines mehrstufigen, qualitativen Befragungsverfahren erhoben. Zunächst wurden 12 PatientenvertreterInnen mit Hilfe eines Fragebogens einzeln zu bestimmten Themen der Digitalisierung und KI-basierten Assistenzsystemen im Gesundheitswesen befragt. Im Rahmen einer darauf aufbauenden zweiten Teilstudie fand im Herbst 2019 – organisiert durch die Plattform Lernende Systeme und dem DFKI – ein Runder Tisch statt, bei dem zehn der 12 PatientenvertreterInnen aus der Vorabumfrage sowie sechs weitere auf Basis der Ergebnisse der qualitativen Vorabumfrage Standpunkte interaktiv diskutieren konnten. Die Auswahl der PatientenvertreterInnen erfolgte unter Berücksichtigung möglichst unterschiedlicher Krankheitsbilder.

Behandlungsqualität mit KI verbessern und Patientensouveränität stärken

Die Befragungsergebnisse lassen insgesamt auf eine eher positive, in Teilen ambivalente, Bewertung von KI-Technologien im Gesundheitswesen durch PatientenvertreterInnen schließen – so werden dem Einsatz von KI große Potenziale aber auch einige Risiken attestiert. Gleichzeitig variiert der Kenntnisstand zu KI-Technologien zwischen den PatientenvertreterInnen und deren Nutzung steht noch am Anfang. Die Mehrheit der befragten PatientenvertreterInnen konnten die Frage nach der Sicherheit KI-basierter Assistenzsysteme in der Medizin nicht beantworten, was auf weiteren Aufklärungsbedarf schließen lässt. Darüber hinaus wünschte sich die Mehrheit mehr methodische Transparenz bei KI-basierten Technologien und Anwendungen. Daher sollte der Kompetenzaufbau zu KI Technologien stärker vorangetrieben werden.

Zu Beginn der interaktiven Diskussion aller 16 PatientenvertreterInnen im Rahmen der zweiten Teilstudie wurden die Themengebiete zu **Chancen und Wünschen** von KI im Gesundheitswesen – gegenüber **möglichen Risiken** und

Tabelle 1: Potenziale von KI in der Medizin aus Sicht von PatientenvertreterInnen

VERBESSERUNG DER BEHANDLUNGSQUALITÄT

- Personalisierte Medizin
- Sektorenübergreifende Prozesse und vernetzte Behandlung
- Diagnostik
- Prozess- und Ablaufoptimierung
- Wissensdiffusion und Erkenntnisgewinn
- Unterversorgung entgegenwirken
- Verringerung von Nebenwirkungen der Behandlung

STÄRKUNG DER PATIENTEN-SOUVERÄNITÄT

- Stärkung der Patientenrechte
- Soziale Teilhabe und Inklusion
- Unterstützung im Alltag

Datenschutzbedenken – als am wichtigsten eingestuft. Dieses Stimmungsbild zeigt, dass für die ExpertInnen die potenziellen Chancen von KI im Gesundheitswesen gegenüber der Risikoeinschätzung überwiegen und sie den möglichen Potenzialen neuer Technologien aufgeschlossen und grundsätzlich positiv gegenüberstehen.

Die Potenziale von KI in der Medizin können den PatientenvertreterInnen zufolge vor allem in zwei Kernbereiche untergliedert werden: So können KI-Assistenzsysteme die **Behandlungsqualität verbessern** und die **Souveränität von Betroffenen** stärken. Als die vielversprechendsten Vorteile von KI-Assistenzsystemen in der Medizin sehen die PatientenvertreterInnen die personalisierte Behandlung und Diagnosefindung, um einer Einheitsbehandlung und den assoziierten Nebenwirkungen und Therapieversagen entgegenzuwirken. Die Dimensionen, die hier als wichtig eingeschätzt wurden, sind personalisierte Behandlungsziele, eine optimierte Therapiewahl und Timing, die Früherkennung von Krankheitsverschlechterungen und neue Optionen für eine ganzheitlichere Therapie. Auch die neuen Möglichkeiten im Rahmen der Diagnostik, Krankheiten frühzeitiger und umfassender zu erkennen, gehörten zu den häufigsten priorisierten Unterthemen. Ebenso häufig wurde der Wunsch nach einer vernetzten Behandlung mit sektorenübergreifenden Prozessen als wichtig eingestuft. Die Prozessoptimierung im Gesundheitswesen durch den Einsatz von Lernenden Systemen, um transparentere und effizientere Abläufe zu ermöglichen, sind weitere wichtige Anliegen. So könnte etwa der bürokratische Aufwand für das Personal durch eine automatisierte Pflegedokumentation, z. B. mittels Spracherkennung, verringert werden. Außerdem könnten Lernende Systeme eine bessere Umsetzung der Leitlinien in der Regelversorgung fördern, etwa durch Entscheidungsunterstützungssysteme, die stets aktuellste Leitlinien berücksichtigen. Ebenso wichtig ist den PatientenvertreterInnen ein optimierter Zugang zu Wissen (Wissensdiffusion) und die Gewinnung neuer wissenschaftlicher Erkenntnisse, etwa durch die KI-assistierte Auswertung von Studiendaten und gesundheitsrelevanten Daten, z. B. Genominformationen. Etwas weniger wichtig wurde die Möglichkeit eingestuft, der Unterversorgung mit neuen Behandlungszugängen, z. B. Telemedizin, KI-basierte Applikationen, entgegenzuwirken.

Der Einsatz KI-basierter Medizinprodukte kann parallel auch zur Stärkung der Patientenrechte beitragen. Dabei ist den Patientenvertretungen zufolge der einfache Zugriff von Betroffenen auf ihre elektronische Patientenakte von besonders hoher Bedeutung. Als wichtig werden in diesem Zusammenhang auch

Tabelle 2: Priorisierungen der Patientenvertretungen

RISIKEN UND HERAUSFORDERUNGEN VON KI IN DER MEDIZIN

- ! Fehlerhafte und diskriminierende Entscheidungen von KI-Systemen
- ! Risiken für die Versorgungsqualität
- ! Abhängigkeiten und Ökonomisierung des Gesundheitswesens
- ! Ethische Fragestellungen

neue Möglichkeiten der Inklusion und sozialen Teilhabe sowie die dadurch gestärkte Autonomie durch KI-Technologien eingestuft, als Beispiel wurden Exoskelette genannt, die die Wiedererlangung motorischer Fähigkeiten unterstützen. Große Bedeutung kommt zudem Unterstützung von Betroffenen im Alltag zu, etwa in Form von neuen Möglichkeiten der Gesunderhaltung, z. B. KI-basierte Applikationen.

Sorgen vor unzureichend geschützten Daten und Datenmissbrauch

Die Angst vor fehlerhaften oder diskriminierenden Therapieentscheidung des medizinischen oder pflegerischen Personals, die auf Basis von KI-Technologien getroffen werden, nannten die Patientenvertretungen als eine der größten Sorgen (vgl. Tabelle 2). Dabei wird vor allem befürchtet, dass falsche oder für die Betroffenen nachteilige Entscheidungen getroffen werden, entweder durch fehlerhafte Interpretationen der KI-Ergebnisse oder durch die mangelhafte Qualität der Datensätze und Algorithmen – etwa wegen unzureichender Nachvollziehbarkeit der Algorithmen-Entscheidung oder dem Training mit irrelevanten Datenpunkten. Dazu ist die kritische Analyse der Datensätze hinsichtlich Zweck, Qualität und Größe notwendig, allerdings mangelt es hier an allgemeingültigen Metriken und anerkannten Standards.

Besonders kritisch sehen die Patientenvertretungen auch die möglichen Risiken für die Versorgungsqualität durch den Einsatz von KI-Technologien. Dabei besteht die größte Sorge darin, dass die menschliche Interaktion durch die standardisierte Behandlung eingeschränkt wird. Ebenso werden die Gefahren der fehlenden Patientenbeteiligung bei der Entwicklung der KI-Systeme sowie die eingeschränkte Mitentscheidung Betroffener befürchtet. Als relevant werden auch ethische Fragestellungen zu den gesellschaftlichen Folgen ungesunder Selbstoptimierung bewertet, wozu die Festlegung eines ethischen Rahmens gefordert wird.

Die Entstehung neuer Abhängigkeiten durch KI-Systeme, etwa durch Zugang und Zugriff zu Daten und Software, zählt zu den weiteren Befürchtungen der Patientenvertretungen. In dem Zusammenhang wurde auch vor der Gefahr gewarnt, dass zu strenge Regulationen innovationshemmend wirken könnten und so etwa die Entstehung von Start-ups verhindern.

Die größte Gefahr beim Einsatz von KI im Gesundheitswesen sehen die VertreterInnen der verschiedenen Verbände in der fehlenden Datensicherheit und in den Folgeproblemen von Datenmissbrauch und Cyberkriminalität (vgl. Tabelle 3). Dabei wünschen sie sich eine stabile IT-Sicherheit, um dies zu verhindern, und Sanktionsmöglichkeiten, um gegen Datenmissbrauch vorgehen zu können. Die Gesundheitsdaten sind elementar, um KI-gestützte Assistenzsysteme zu entwickeln und nutzbar zu machen. Ein umsichtiger Datenschutz ist daher eine notwendige Voraussetzung, um die KI-gestützte Versorgung zu stärken und die DatenspenderInnen zu schützen. Wichtig ist in diesem Zusammenhang die informationelle Selbstbestimmung sowie die Klärung der Frage, wer die Daten besitzen und nutzen darf. Hier wünschen sich die befragten ExpertInnen Vorgaben zur Datenspeicherung (z. B. Vertraulichkeit, Verfügbarkeit, Rechtsverbindlichkeit, Zugriffsrechte). Ansonsten wird etwa eine Datensammlung ohne Einwilligung befürchtet oder, dass das Recht auf Vergessen nicht mehr gewährleistet ist. Zudem sollte eine Anlaufstelle für Rechtsbeistand geschaffen werden, damit Betroffene selbstbestimmt ihre Rechte einfordern können.

Tabelle 3: Priorisierungen der Patientenvertretungen – Wichtige Datenschutzbestimmungen für KI in der Medizin

DATENMISSBRAUCH UND CYBERKRIMINALITÄT

- ! Sanktionierung von Datenmissbrauch
- ! Sichere und stabile IT-Systeme

INFORMATIONELLE SELBSTBESTIMMUNG

- ! Besitz und Nutzung der Daten
- ! Fehlende Vorgaben zur Datenspeicherung
- ! Rechtsbeistand für Betroffene
- ! Datensammlung ohne Einwilligung
- ! Recht auf Vergessen

Die Ergebnisse der Befragung von VertreterInnen verschiedener Betroffenengruppen legen Handlungsoptionen und politisch-regulatorische Gestaltungserfordernisse aus Sicht von Betroffenen für die weitere Implementierung von KI im Gesundheitswesen offen. Der Austausch mit den Patientenvertretungen zeigt zudem, dass noch weiterer Aufklärungsbedarf über die Funktionsweise und Anwendungsmöglichkeiten von KI-basierten Assistenzsystemen besteht. So müssen Potenziale und Grenzen von KI-Systemen in der Medizin öffentlich kommuniziert und politisch diskutiert werden, um die Wahrnehmung der gesellschaftlichen Relevanz zu stärken und die Akzeptanz weiter zu erhöhen. Dazu braucht es einen breit angelegten gesellschaftlichen Dialog, der KI-Entwickler, Anwender und Betroffene einbezieht.

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. Matthieu-P. Schapranow
ist Mitglied der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme, Leiter der Arbeitsgruppe „In-Memory Computing for Digital Health“ und wissenschaftlicher Leiter Digital Health Innovations am Hasso-Plattner-Institut (HPI).
schapranow@hpi.de

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

Prof. Dr.- Thomas Zahn
ist Mitglied der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme und Geschäftsführer des DSI- Data Science Institute @ bbw sowie Professor für Data Science an der bbw Hochschule Berlin.
thomas.zahn@bbw-hochschule.de

www.bbw-hochschule.de

Prof. Dr. Elsa Andrea Kirchner
ist Mitglied der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme und Professorin für Systeme der Medizintechnik an der Universität Duisburg-Essen sowie Teamleiterin des Teams „Intelligent Healthcare Systems“ am DFKI Robotics Innovation Center, Bremen.
elsa.kirchner@dfki.de

https://www.dfki.de/web/forschung/forschungsbereiche/robotics-innovation-center/

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.

innovative software-werkzeuge für die lebenswissenschaften

BMBF-Initiative CompLS unterstützt die Entwicklung rechnergestützter Methoden und Analysewerkzeugen mit 34 Millionen Euro

von Daniel Heinrichs und René Wolf-Eulenfeld

Um die Entwicklung von innovativen rechnergestützten Methoden und Analysewerkzeugen in den Lebenswissenschaften voranzutreiben, hat das Bundesministerium für Bildung und Forschung (BMBF) die Fördermaßnahme „Computational Life Sciences- CompLS“ ins Leben gerufen und unterstützt diese seit dem Jahr 2018 in insgesamt vier Runden mit bisher 34 Millionen Euro in 53 Projekten.

In den Lebenswissenschaften ist der Fortschritt im Bereich experimenteller Methoden und moderner (Hochdurchsatz-) Technologien ein bedeutender Treiber für Innovationen. Es werden aktuell vermehrt quantitative, zeitaufgelöste Messungen von zellulären Systemen durchgeführt. Gleichzeitig nimmt auch die Menge der digitalisierten Daten in der Patientenversorgung und der klinischen Forschung rasant zu. Dadurch steigt der Bedarf an neuen bioinformatischen Werkzeugen und intelligenten Algorithmen für die effiziente Verarbeitung und Analyse der gemessenen Daten sowie neuen Methoden zur mathematischen Modellierung und Simulation komplexer biologischer Systeme.

Mit der Fördermaßnahme „Computational Life Sciences“ möchte das BMBF die Entwicklung innovativer Methoden und Softwarewerkzeuge zur bioinformatischen Verarbeitung, Modellierung und Simulation in den Lebenswissenschaften weiter voranbringen. Dadurch sollen der lebenswissenschaftlichen Forschung in Deutschland effiziente und zuverlässige Hilfsmittel, insbesondere aus dem Bereich der künstlichen Intelligenz (KI), zur Verfügung gestellt werden, um die durch neueste experimentelle Methoden gewonnenen Daten sowie Datensätze aus verschiedenen Datenquellen geeignet zu modellieren und zu analysieren.

Während die erste Auswahlrunde noch themenoffen war, wurden die weiteren Runden auf verschiedene Schwerpunktthemen fokussiert: „Deep Learning in der Biomedizin“, „Maschinelles Lernen für die Krebsforschung“ und „KI-Methoden für die Infektionsforschung“.

Die bisherigen Erfolge der Maßnahmen werden im Folgenden exemplarisch anhand von zwei Beispielen veranschaulicht.

Kontakt:

Dr. René Wolf-Eulenfeld
Projektträger Jülich
Forschungszentrum Jülich GmbH
r.wolf-eulenfeld@fz-juelich.de

Dr. Daniel Heinrichs
Projektträger Jülich
Forschungszentrum Jülich GmbH
d.heinrichs@fz-juelich.de

https://www.ptj.de/computational-life-sciences

Fotos: FZJ, R. U. Limbach und Daniel Heinrichs

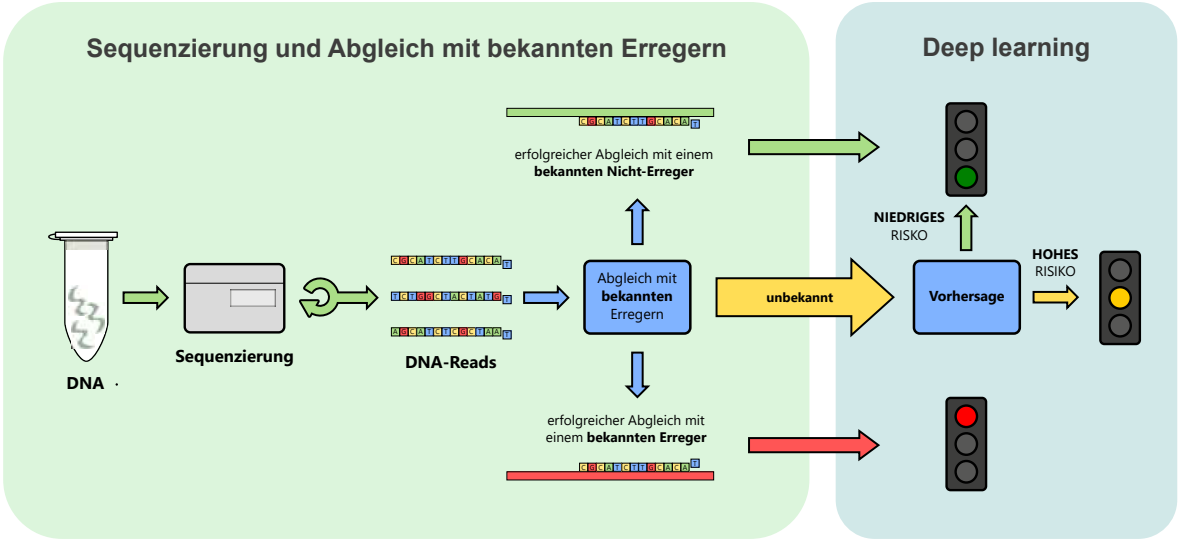


Abbildung 1: Arbeitsablauf der Infektiositätsvorhersage.
Eine biologische Probe wird sequenziert und die resultierenden Reads werden mit Genomen von bekannten Erregern und Nicht-Erregern abgeglichen. Von unbekannten Viren stammende Reads werden dann an ein neuronales Netz weitergeleitet, um vorherzusagen, ob das neue Virus den Menschen infiziert (Bild: Jakub Bartoszewicz und Melania Nowicka).

vorhersage des infektionspotenzials neuartiger viren

Interpretierbare neuronale Netze erkennen schnell neue virale Erreger auf Basis ihres Genoms

von Jakub Bartoszewicz, Melania Nowicka und Bernhard Renard

Neue Krankheiten tauchen immer wieder auf, aber die Standardmethoden der molekularen Diagnostik können nur bereits bekannte Erreger erkennen. Sie sind darauf angewiesen, sehr spezifische Übereinstimmungen mit bereits untersuchten Organismen und Viren zu finden. Neue Viren, die sich deutlich von den alten unterscheiden, werden daher leicht übersehen. Im DeepPath-Projekt entwickeln Forscher am Hasso-Plattner-Institut neue Deep-Learning-Ansätze, die vorhersagen können, ob ein neuartiges Virus den Menschen infiziert. Dies ermöglicht sowohl eine extrem schnelle Erregererkennung aus wenigen Daten als auch die Analyse ganzer Genome.

Im letzten Jahr konnten wir rasche Fortschritte bei der Entwicklung von Tests zum Nachweis des SARS-CoV-2-Virus beobachten. PCR, Antikörpertests und schließlich Schnell- oder sogar selbst durchgeführte Antigentests wurden in vielen Ländern zur täglichen Routine. Zu Beginn der Pandemie waren jedoch noch keine gezielten Tests entwickelt worden. Die am besten geeignete Methode, eine völlig neuartige Bedrohung zu erkennen und zu analysieren, ist die Genomsequenzierung. Diese Technologie ist besonders wichtig, da mit dem Auftreten weiterer neuer Erreger zu rechnen ist – sie entwickeln sich extrem schnell und viele bleiben noch unentdeckt.

Genomsequenzierung und neuronale Netze

Anstatt nach sehr spezifischen Signaturen des Erregers zu suchen, zielt die Genomsequenzierung darauf ab, seine gesamte genomische Information zu entziffern. Die Sequenzierungsgeräte lesen Milliarden von Nukleotiden pro Probe, aber die resultierenden Reads sind nur Fragmente des ursprünglichen Genoms - um Größenordnungen kürzer und aus dem Zusammenhang gerissen. Es ist immer noch schwierig, die Reads wieder in ihre ursprüngliche Reihenfolge zu bringen. Aus diesem

Grund verlassen sich standardmäßige bioinformatische Methoden zur Erkennung von Pathogenen auf den Abgleich dieser Reads mit Datenbanken bekannter Genome. Daher können sie nur bereits bekannte Krankheitserreger erkennen. DeepPath, ein vom BMBF gefördertes Projekt, das am Fachgebiet Data Analytics and Computational Statistics des Hasso-Plattner-Instituts (HPI) in Potsdam durchgeführt wird, will das ändern.

Die im Rahmen des Projekts entwickelten tiefen neuronalen Netze sagen voraus, ob die Reads von Viren stammen, die den Menschen infizieren, oder von solchen, die das nicht tun. Diese maschinellen Lernverfahren erkennen Muster, die sie mit Krankheitserregern assoziieren. Daher können sie für jede Sequenz ein „infektiöses Potenzial" vorhersagen, nicht nur für solche, die bekannten Genomen ähneln. Dies führt zu einer wesentlich genaueren Erkennung. Es ist sogar möglich, Teilergebnisse zu analysieren, während das Sequenzierungsgerät noch läuft, was Ergebnisse nahezu in Echtzeit liefert. Während jedoch der direkte Abgleich ähnlicher Sequenzen für einen Experten relativ einfach zu interpretieren ist, sagen die neuronalen Netze nur voraus, ob eine Sequenz ein Anzeichen für einen neuen Erreger sein könnte, ohne die Gründe dafür zu erklären.

Interpretierbarkeit für Netze und neue Genome

Diese Herausforderung kann auf zwei verschiedene Arten angegangen werden. Erstens können die standardmäßigen Methoden zusammen mit den neuronalen Netzen verwendet werden. Bei dieser Herangehensweise kann die Pathogenität für alle Reads in einer Probe genau vorhergesagt werden. Dazu werden die Reads mit den Genomen eng verwandter Viren abgeglichen. Der zweite Ansatz verwendet Berechnungsmethoden, die die wichtigsten vom Netzwerk gelernten Muster visualisieren. Sogar Regionen mit besonders hohem infektiösem Potenzial können in bisher ungesehenen Genomen erkannt werden. Dies erlaubt sowohl zu verstehen, wie das neuronale Netz zu seinen Vorhersagen kommt, als auch zu sehen, welche Gene der analysierten Erreger Marker für ihre Infektionsfähigkeit beim Menschen sind.

In Zukunft könnte eine ähnliche Technologie eingesetzt werden, um klinische Proben zu analysieren, die Mischungen mehrerer verschiedener Erregerklassen (z. B. Viren, Bakterien und Pilze) auf einmal enthalten, oder sogar, um Sicherheitsfragen in der synthetischen Biologie zu adressieren – synthetische DNA-Sequenzen könnten noch vor ihrer Synthese auf potenzielle Gefahren überprüft werden. Während die Erkennung und Erforschung neuartiger Krankheitserreger immer menschliche Experten erfordern wird, bieten die im DeepPath-Projekt entwickelten interpretierbaren Deep-Learning-Ansätze neue Möglichkeiten für extrem schnelle Vorhersagen aus extrem wenigen Daten. Dies könnte kostbare Zeit sparen und es uns ermöglichen, zukünftige biologische Bedrohungen zu vereiteln, noch bevor es zu einer Ausbreitung kommt.

Steckbrief Forschungsprojekt:

Das **DeepPath-Projekt** wird im Rahmen der „Computational Life Science" Fördermaßnahme des BMBF gefördert und am Hasso-Plattner-Institut mit Unterstützung des Robert Koch-Instituts durchgeführt. Es wird koordiniert von Prof. Dr. rer. nat. Bernhard Renard, Leiter des Fachgebiets „Data Analytics and Computational Statistics" am HPI.

Kontakt:

Prof. Dr. rer. nat. Bernhard Renard
Hasso-Plattner-Institut,
Digital Engineering Fakultät,
Universität Potsdam
Bernhard.Renard@hpi.de

https://hpi.de/forschung/fachgebiete/data-analytics-and-computational-statistics.html

Jakub Bartoszewicz
Hasso-Plattner-Institut,
Digital Engineering Fakultät,
Universität Potsdam
Jakub.Bartoszewicz@hpi.de

Melania Nowicka
Hasso-Plattner-Institut,
Digital Engineering Fakultät,
Universität Potsdam
Melania.Nowicka@hpi.de

Fotos: Kay Herschelmann, HPI; Kay Herschelmann, HPI und Melania Nowicka/privat

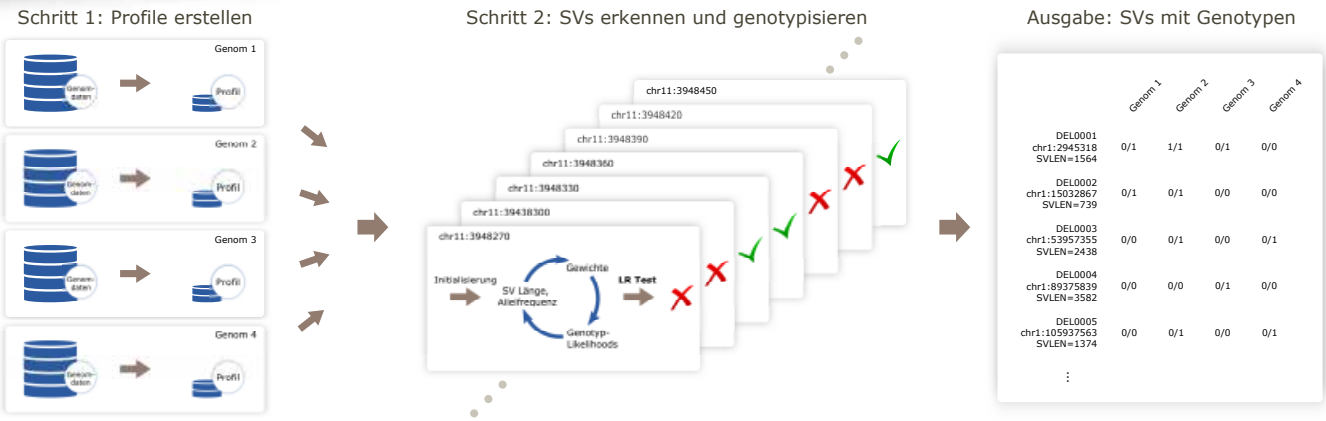


Abbildung 2: Überblick über die Berechnungsschritte im Programm PopDel.
Im ersten Schritt reduziert PopDel die Eingabedaten jedes Genoms auf kleine Profile. Im zweiten Schritt bestimmt es für alle Genome gemeinsam mit Hilfe von Likelihood-Quotienten-Tests (LR-Test), ob Strukturvarianten in jedem möglichen genomischen Abschnitt vorliegen. Die Ausgabe von PopDel enthält Informationen zur Position und Größe der Varianten sowie die entsprechenden Genotypen aller Genome (Bild: Birte Kehr).

auf der suche nach variationen in der sequenzstruktur unserer genome

… wie die Software PopDel das Potential großer Datenmengen nutzt

von Sebastian Niehus und Birte Kehr

Jeder Mensch trägt eine Vielzahl genetischer Varianten in seiner DNA-Sequenz, welche unser Aussehen und unsere Gesundheit maßgeblich beeinflussen. Um diese Unterschiede zu finden sowie ihre Auswirkungen besser zu verstehen, müssen die Genome zehntausender Personen durchsucht und analysiert werden – ein schwieriges und zeitintensives Unterfangen, welches bisherige Rechenverfahren an ihre Grenzen stoßen lässt. Mit neuen Ansätzen zur Analyse riesiger Datenmengen kann diese Herausforderung jedoch bewältigt werden, um wertvolle Forschungserkenntnisse zu gewinnen und neue Therapiemöglichkeiten für schwerwiegende Erkrankungen zu entwickeln.

Neben den Millionen von Variationen einzelner DNA-Stellen gibt es Tausende Variationen der Sequenzstruktur in jedem Genom. Solche Strukturvarianten (SVs) können verschiedenste Formen annehmen, bei denen DNA-Abschnitte umgedreht (Inversionen), verdoppelt (Duplikationen), neu eingefügt (Insertionen) oder verloren gegangen (Deletionen) sind. Bestimmte SVs stehen im Zusammenhang mit genetisch bedingten Erkrankungen wie etwa Alzheimer, Muskeldystrophien oder Autismus. Daher sind SVs von großem medizinischen Interesse. Um SVs und ihre Auswirkungen besser zu verstehen, werden heute Tausende von menschlichen Genomen sequenziert und auf SVs untersucht. Spezielle Computerprogramme analysieren die digitalisierte DNA-Sequenz und geben Auskunft über das Vorhandensein von SVs. Eine große Chance aber ebenso große technische Herausforderung stellen dabei die rasch wachsenden Mengen an verfügbaren Sequenzdaten dar. Immer mehr Genome werden sequenziert, sodass es spezieller Ansätze bedarf, um diese riesigen Datenmengen zu bewältigen.

Informationen aus zehntausenden Genomen

In dem Projekt Cohort-SV stellen wir uns dieser Herausforderung. Das von uns entwickelte Programm PopDel (Niehus *et al.*, 2021) setzt auf effiziente Algorithmen, geschicktes Datenmanagement und ein statistisches Modell, um die Genome von zehntausenden Menschen gemeinsam zu analysieren. So können wir herausfinden, welche SVs in jedem einzelnen Genom vorhanden sind und gleichzeitig Informationen über die Vererbung und Verbreitung von SVs in der Bevölkerung gewinnen.

Zur Bewältigung der großen Datenmengen arbeitet PopDel in zwei Schritten. Im ersten Schritt extrahiert es die relevanten Informationen aus jedem sequenzierten Genom und speichert diese jeweils in einem kleinen Profil. Damit wird die später zu analysierende Datenmenge auf ein bis zwei Prozent der ursprünglichen Größe reduziert, so dass PopDel anschließend

die Profile sehr vieler Genome gemeinsam analysieren kann. Im zweiten Schritt liest es die genetischen Informationen aus allen verfügbaren Profilen in kleinen genomischen Abschnitten ein und verrechnet sie in einem statistischen Modell miteinander, um zuverlässig Auskunft über das Vorkommen von SVs zu geben. Eine besondere Herausforderung hierbei ist, dass bei der Analyse sehr vieler Genome unterschiedliche SVs an der gleichen Stelle auftreten können, die als unterschiedliche Varianten erkannt werden sollen. Außerdem muss gewährleistet werden, dass SVs, die nur in einem einzigen Genom vorkommen, nicht von den Tausenden anderer Genomen verdeckt werden. Unser Anspruch ist es, dass die individuellen Ergebnisse beim gemeinsamen Analysieren von zehntausenden Genomen qualitativ mindestens genauso gut sind, wie bei der Analyse eines einzelnen Genoms.

Entdeckung einer neuen Gen-Variante

Um die Qualität der Ergebnisse und Tauglichkeit für die Praxis zu belegen, musste sich PopDel in verschiedenen Testszenarien bewähren. Auf simulierten wie auch realen Daten von einzelnen bis knapp 50.000 Genomen gelang es uns mit PopDel hochwertige Ergebnisse zu erzielen. Der bisher bedeutendste Fund von PopDel war eine bis dato unbekannte Deletion im sogenannten LDL-Rezeptor-Gen durch unseren Partner deCODE genetics (Björnsson *et al.*, 2020). Mittlerweile konnten unsere Partner experimentell bestätigen, dass die Deletion einen positiven Einfluss auf die Funktionsweise des LDL-Rezeptors hat. Da dieser eine zentrale Rolle im Cholesterinstoffwechsel unseres Körpers spielt, kann die Entdeckung der Deletion als Ansatzpunkt für zukünftige Forschung zur Behandlung von Herz-Kreislauf-Erkrankungen dienen.

Ausblick

Derzeit ist PopDel in der Lage, Deletionen zuverlässig zu bestimmen. In der laufenden Projektphase erweitern wir das Programm auf weitere Arten von SVs. Dank der schnellen Laufzeit und einem Ansatz, der explizit für große Datenmengen ausgelegt ist, kann PopDel zukünftig in den immer größeren DNA-Sequenzierungsstudien zur Anwendung kommen und so dabei helfen, SVs und die mit ihnen in Zusammenhang stehenden Krankheiten besser zu verstehen.

Steckbrief Forschungsprojekt:

Name des Projektes:
Cohort-SV – Auffinden und Genotypisieren von struktureller Variation in Sequenzierungsdaten ganzer Genome für große Kohorten

Unterstützende Institutionen:
Regensburger Centrum für Interventionelle Immunologie (RCI); Berliner Institut für Gesundheitsforschung in der Charité – Universitätsmedizin Berlin

Beteiligte Partner:
Prof. Dr. Birte Kehr

Referenzen:

Niehus, S., Jónsson, H., Schönberger, J., Björnsson, E., Beyter, D., Eggertsson, H. P., Sulem, P., Stefánsson, K., Halldórsson, B. V., und Kehr, B. (2021). PopDel identifies medium-size deletions simultaneously in tens of thousands of genomes. Nature communications, 12(1), 1-10.

Björnsson, E., Gunnarsdóttir, K., Halldórsson, G. H., Sigurðsson, Á., Árnadóttir, G. A., Jónsson, H., Olafsdottir, E. F., Niehus, S., Kehr, B., Sveinbjörnsson, G., Gudmundsdottir, S., Helgadottir, A., Andersen, K., Thorleifsson, G., Eyjolfsson, G. I., Olafsson, I., Sigurdardottir, O., Saemundsdottir, J., Magnusson, O. Th., Masson, G., Stafansson, H., Gudbjartsson, D. F., Thorgeirsson, G., Holm, H., Halldorsson, B. V., Melsted, P., Norddahl, G. L., Sulem, P., Thorsteinsdottir, U., und Stefánsson, K. (2020). Lifelong Reduction in LDL Cholesterol Due to a Gain-of-Function Mutation in LDLR. Circulation: Genomic and Precision Medicine.

Kontakt:

Prof. Dr. Birte Kehr
Leitung AG Algorithmische Bioinformatik
Regensburger Centrum für Interventionelle Immunologie (RCI)
Birte.Kehr@ukr.de

Sebastian Niehus
Doktorand AG Algorithmische Bioinformatik
Regensburger Centrum für Interventionelle Immunologie (RCI)
Sebastian.Niehus@ukr.de

https://www.rcii.de

Fotos: Universität Regensburg, Birte Kehr

zukunftsorientierte analyse von forschungsdaten

Mitglieder des de.NBI-Netzwerkes nutzen KI-basierte Technologien

von Vera Ortseifen, Peter Belmann und Alfred Pühler

Forschungsdaten sind im digitalen Zeitalter die elementare Währung für datenbasierte Innovationen. Dieses neue Gold hat einen erheblichen Einfluss auf die rasante Weiterentwicklung von Künstlicher Intelligenz (KI). Dies gilt jedoch auch umgekehrt – die Analyse von Daten kann immens von KI-basierten Methoden profitieren. Das Potential wird im Umfeld des deutschen Netzwerkes für Bioinformatik Infrastruktur (de.NBI) mit Blick auf KI-basierte Projekte ersichtlich. Diese Projekte profitieren darüber hinaus von der hervorragenden Ausstattung der de.NBI-Cloud, welche allen Forschenden aus den Lebenswissenschaften gebührenfrei zur Verfügung steht.

Die Bedeutung der Forschungsdaten für KI-basierte Entwicklungen

Forschungsdaten langfristig und nachhaltig nutzbar zu machen, ist das Gebot der Stunde. Die sogenannten *„FAIR Data Principles - Findable, Accessible, Interoperable, and Re-usable“* definieren dabei die elementaren Grundsätze, die auch im deutschen Netzwerk für Bioinformatik Infrastruktur (de.NBI) und der europäischen Life Science Infrastruktur für biologische Informationen (ELIXIR) gelebt werden. Dabei spielen Speicherung und Management der Daten nach "FAIR Data Principles" für die aktuelle Entwicklung der Künstlichen Intelligenz (KI) eine essenzielle Rolle. Es ist davon auszugehen, dass die Weiterentwicklung der KI weltweit den Fortschritt auf vielen Forschungsgebieten beeinflussen wird. Dabei lebt Maschinelles Lernen (Machine Learning, ML) und Tiefes Lernen (Deep Learning, DL) von dem Vorhandensein von Daten. Je größer die Datenmengen beziehungsweise die Trainings- und Testdatensätze sind, desto besser können KI-Programme trainiert werden. Dieser einfache Zusammenhang führt seit Jahren durch das stetige Wachstum von Datenmengen zur Bedeutungssteigerung von KI-Ansätzen. Der Einsatz von Künstlicher Intelligenz in den Lebenswissenschaften ist daher nicht nur unverzichtbar, sondern birgt auch enormes Potential, um Themen wie personalisierte Medizin, Medikamentenentwicklung oder biologische Grundlagenforschung voranzutreiben.

Dabei wächst de.NBI (siehe Steckbrief Deutsches Netzwerk für Bioinformatik-Infrastruktur) mit aktuellen Entwicklungen und hat die Chancen von KI-basierten Methoden für die Analyse von Forschungsdaten früh erkannt. Mitglieder des Netzwerkes setzen aus diesem Grund verstärkt auch auf KI-basierte Technologien, um Forschungsdaten zu analysieren. Der Umfang an KI-Projekten ergibt sich aus einer Abfrage innerhalb des Netzwerkes, wonach insgesamt 25 Projekte aus den Bereichen Modellierung, Phenotypisierung, *Omics* sowie der Meta-Analyse, Datenbanken und dem Support zurückgemeldet wurden (Abbildung 1). Einige dieser KI-basierten Projekte der de.NBI-Mitglieder sollen im Folgenden näher erläutert werden.

Identifizierung neuer antimikrobieller Resistenzziele durch Deep Learning im Hochdurchsatz

An den Universitäten Gießen und Marburg in den Arbeitsgruppen von Alexander Goesmann, Dominik Heider und Trinad Chakraborty arbeiten die Wissenschaftler an der Identifizierung neuer antimikrobieller Resistenzziele. Antibiotikaresistente Bakterien sind weltweit zu einer ernsten Bedrohung für die öffentliche Gesundheit geworden. Ihre Überwachung und Eindämmung sowie die rechtzeitige Identifizierung neuer genetischer Determinanten von Antibiotikaresistenzen sind wichtige Werkzeuge, um sowohl aufkommende als auch etablierte Krankheitserreger zu bekämpfen. Um diesen Herausforderungen zu begegnen, haben Forschende der Universitäten Gießen und Marburg ein BMBF-gefördertes Projekt namens *„Deep-iAMR – Identifizierung von neuen antimikrobiellen Resistenz-*

Rubrik	Projekte	Schlagwörter	Forschungseinrichtungen
Modellierung & Phänotypisierung	7	Krankheitsmerkmale, Patientenstratifikation, Risikomodellierung, antimikrobielle Substanzen, Wirkmechanismen, Mikroskopie	• Leibniz Institut IPK, Gatersleben • Berlin Institut für Gesundheitsforschung in der Charité (BIH) • Universität Heidelberg • Max-Delbrück-Centrum für Molekulare Medizin, Berlin • Justus-Liebig-Universität Gießen • Deutsches Krebsforschungszentrum (DKFZ) Heidelberg
Omics- & meta-Analysen	13	Genomics, Postgenomics, ncRNA, Proteomics, single-cell, Meta-Analysen, krankheitsübergreifende Vergleiche	• Ruhr Universität Bochum • Universität Rostock • Universität Leipzig • Forschungszentrum Jülich • European Molecular Biology Laboratory, Heidelberg • Eberhard Karls Universität Tübingen • Max Planck Institut Magdeburg
Datenbanken & Support	5	Kuratieren, Datenhoheit, Tools-Empfehlungen, Bildverarbeitung, Forschungsförderung	• Heidelberger Institut für Theoretische Studien • Albert-Ludwigs-Universität Freiburg • Charité, Berlin • Universitätsklinikum Tübingen • Justus-Liebig-Universität Gießen

Abbildung 1: Die 25 KI-basierten Projekte der Mitglieder des de.NBI-Netzwerkes in verschiedenen Bereichen der Lebenswissenschaften (Quelle: de.NBI Geschäftsstelle; Hintergrund Ipopba – stock.adobe.com).

Targets durch Deep-Learning-Verfahren im Hochdurchsatz“ ins Leben gerufen. Dieses Projekt kombiniert die Expertise aus den Gebieten der medizinischen Mikrobiologie, dem Maschinellem Lernen und von Cloud-Computing-Techniken (Abbildung 2) zur Entwicklung automatischer, skalierbarer Bioinformatik-Workflows. Neue Deep Learning-Ansätze werden eingesetzt, um antimikrobielle Resistenzen schnell und zuverlässig vorherzusagen und um bisher verborgene Signalwege und neuartige Ziele für die Entwicklung neuer Antibiotika aufzudecken. Dabei werden die Daten auf komplexe Muster zwischen genomischen und phänotypischen Daten untersucht, um die genetischen Determinanten von schwer zu entdeckenden oder bislang unbekannten antimikrobiellen Resistenzen zu finden. Die Anwendung moderner Methoden der Künstlichen Intelligenz für die Auffindung dieser Muster, erfordert jedoch massive Rechenressourcen sowie spezialisierte Hardware, wie Grafikprozessoren (GPUs) und Tensoren. Forschende innerhalb des Deep-iAMR-Projekts nutzen die Vorteile der Cloud-Computing-Infrastruktur des de.NBI-Netzwerks, die sowohl die benötigten Rechenressourcen als auch spezialisierte Hardware bei Bedarf zur Verfügung stellt.

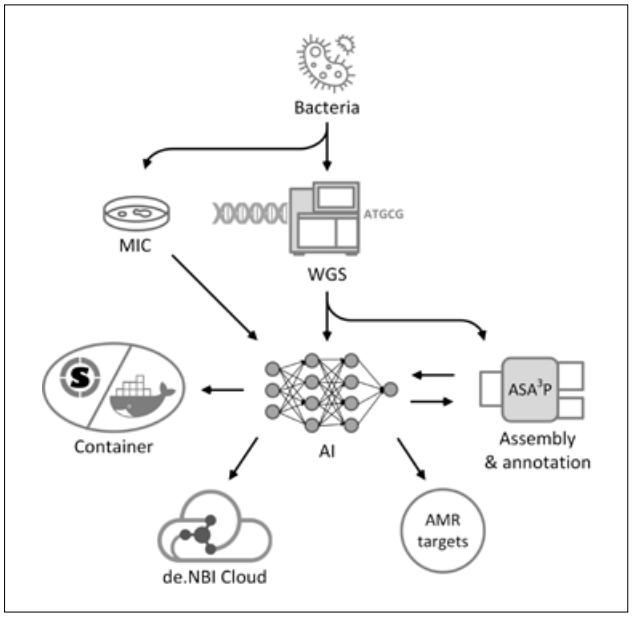


Abbildung 2: Der Datenfluss innerhalb des Deep-iAMR-Projekts. Das Projekt zielt darauf ab, verschiedene Omics-Datensätze mit klinischen und phänotypischen Informationen für einen großen, gut charakterisierten Satz von multiresistenten E.coli-Isolaten zu kombinieren. Diese Daten werden dann verwendet, um tiefe neuronale Netze zu trainieren. Schließlich wird unsere Software über benutzerfreundliche Containerisierungstechniken und skalierbare Cloud-Lösungen bereitgestellt. WGS = Gesamtgenomsequenzierung, MIC = Minimale Hemm-Konzentration, AMR = Antimikrobielle Resistenz (Quelle: O. Schwengers, K. Brinkrolf, S. Doijad, T. Chakraborty, D. Heider, A. Goesmann).

Weitere Informationen finden Sie auf Deep-iAMR:
https://www.gesundheitsforschung-bmbf.de/de/deep-iamr-identifizierung-von-neuen-antimikrobiellen-resistenz-targets-durch-deep-learning-10900.php

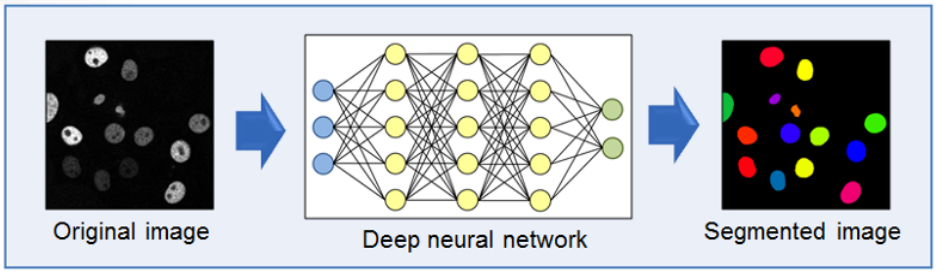


Abbildung 3: Tiefes neuronales Netzwerk für die Segmentierung von Zellmikroskopie-Bildern (Quelle: C. Krug, K. Rohr).

Deep Learning zur Analyse von Mikroskopiebildern und zur zellulären Phänotypisierung

Die automatisierte Analyse von Hochdurchsatz- und High-Content-Mikroskopie-Bilddaten ist wichtig, um zelluläre Prozesse aufzuklären. Allerdings ist die Analyse solcher Daten mit einer Reihe von Herausforderungen verbunden. In jüngster Zeit sind Deep Learning-Methoden aus dem Bereich der Künstlichen Intelligenz aufgekommen, die im Vergleich zu klassischen Bildanalysemethoden bessere Ergebnisse liefern.

Die Arbeitsgruppe Biomedical Computer Vision der Universität Heidelberg unter der Leitung von Karl Rohr entwickelt DL-Methoden für die computergestützte Analyse von Zellmikroskopie-Bildern. Ziel ist es, die automatisierte Quantifizierung von zellulären Phänotypen auf Einzelzellebene zu verbessern sowie großskalige Mikroskopiedaten effizient zu verarbeiten. Insbesondere wurden dabei Deep Learning-Methoden für die Zellsegmentierung entwickelt. Die Methoden kombinieren verschiedene Arten von neuronalen Netzwerkarchitekturen, wie z. B. faltende neuronale Netze und rekurrente neuronale Netze.

Die entwickelten DL Methoden zur Zellsegmentierung wurden zur Analyse von Hochdurchsatz- und High-Content-Mikroskopie-Bilddaten eingesetzt. Ein Beispiel für ein Segmentierungsergebnis für ein Zellmikroskopie-Bild ist in Abbildung 3 dargestellt. Es ist zu erkennen, dass Zellen mit hohem und niedrigem Bildkontrast gut segmentiert werden können. Die entwickelten Methoden wurden z. B. zur Segmentierung von Zellen in Gewebebildern verwendet, um anschließend die Länge der Telomere (Enden der Chromosomen) zu quantifizieren, was für die medizinische Diagnose genutzt werden kann. Diese Arbeiten wurden im Rahmen des BMBF-Projekts *„CancerTelSys - Identifizierung von Netzwerken für die Telomer-Erhaltung in Tumoren zur Diagnose, Prognose, Patientenstratifizierung und Vorhersage der Therapieantwort“* durchgeführt. Die Entwicklung und Anwendung der DL-Methoden profitierten dabei von der de.NBI-Compute-Infrastruktur und der Cloud.

Für weitere Informationen besuchen Sie bitte die Webseiten der Biomedical Computer Vision Group.
http://www.bioquant.uni-heidelberg.de/research/groups/biomedical_computer_vision.html

de.NBI-Cloud: Rechnerische Ressourcen für die Anwendung von Künstlicher Intelligenz in den Lebenswissenschaften

Auch wenn Machine und Deep Learning von vielen wichtigen Software-Programmen der Lebenswissenschaften genutzt werden, gibt es Anforderungen für die performante und erfolgreiche Anwendung. Zudem können die Methoden der Künstlichen Intelligenz durch den Einsatz von GPUs beschleunigt werden und gerade im DL vom Zugriff auf enorme Datenmengen profitieren. Wie bereits in den beiden beschriebenen Projekten angedeutet, bietet die de.NBI-Cloud (https://cloud.denbi.de/get-started/) hierzu optimale Compute-Ressourcen. GPUs sind in verschiedenen Varianten und Typen verfügbar und können je nach Anforderung der ML-Anwendung flexibel eingesetzt werden. Die de.NBI-Cloud speichert eine große Menge an wissenschaftlichen Daten, wie z. B. die Metagenom-Datensätze der Sequence Read Archive, was einen schnellen Zugriff ermöglicht und das Herunterladen der Daten von verschiedenen Ressourcen überflüssig macht. GPUs, virtuelle Prozessoren (vCPUs), Arbeitsspeicher (RAM) und Speicherlösungen sind für die Forschende in Deutschland kostenlos und helfen, die Entwicklung und Anwendung von Methoden der Künstlichen Intelligenz in den Lebenswissenschaften zu fördern (Abbildung 4).

Die de.NBI-Cloud stellt bereits Ressourcen für eine breite Palette von ML Anwendungen zur Verfügung. Die Themen der Bioinformatik reichen von der Bildanalyse über die Vorhersage von antimikrobiellen Resistenzen bis hin zur Modellierung von 3D-Genomanalysen, Fluoreszenzmikroskopie, Krebsdiagnose und Vorhersage von antiviralen Peptiden. Von den über 480 in der de.NBI-Cloud gerechneten Projekte nutzen insgesamt 55 Projekte Machine Learning Methoden im Rahmen ihres Forschungsvorhabens. Zahlreiche Projekte, die DL oder

ML nutzen, sind bereits veröffentlicht. Die komplett akademische de.NBI-Cloud ist aufgrund der föderativen Struktur über sechs Standorte verteilt (Abbildung 2). Mittels starker personeller Unterstützung wird auch unerfahrenen Forschenden der Zugang ermöglicht. Insgesamt wurden knapp 50 Millionen Euro in den Aufbau der de.NBI-Cloud investiert.

Danksagung

Die Arbeiten an diesem Artikel wurden unterstützt durch O. Schwengers, K. Brinkrolf, A. Goesmann (Deep-iAMR Projekt) und C. Krug sowie K. Rohr (Teilprojekt CancerTelSys).

Steckbrief Deutsches Netzwerk für Bioinformatik-Infrastruktur

Das Deutsche Netzwerk für Bioinformatik-Infrastruktur (de.NBI) 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. Dabei bietet de.NBI umfassende Bioinformatik-Services auf dem neuesten Stand der Technik für Anwender in der Grundlagenforschung und der angewandten Forschung aus den Lebenswissenschaften, der Wissenschaft, der Industrie und der Medizin an. Darüber hinaus fördert de.NBI die Zusammenarbeit der deutschen Bioinformatik-Community mit internationalen bioinformatischen Netzwerkstrukturen. Das Netzwerk zählt zurzeit mehr als 300 wissenschaftliche Mitglieder und besteht aus 40 Projekten, die in acht thematisch ausgerichteten Servicezentren angesiedelt sind. Die im de.NBI-Netzwerk aufgebaute Infrastruktur umfasst die Bereiche Service, Training und Compute. Allen Forschenden aus den Lebenswissenschaften steht die Nutzung der de.NBI-Cloud gebührenfrei zur Verfügung (siehe https://www.denbi.de/cloud). Ergänzt wird diese einmalige Rechnerressource durch die im Servicebereich angebotenen Analyseprogramme, Workflows und Datenbanken, deren Nutzung durch Trainingskurse unterstützt wird.

Weitere Informationen:
www.denbi.de

Kontakt:

Dr. Vera Ortseifen
Koordinatorin des Deutschen Knotens der europäischen Life Science Infrastruktur für biologische Informationen (ELIXIR)
Centrum für Biotechnologie
Universität Bielefeld
vera@cebitec.uni-bielefeld.de

Peter Belmann
de.NBI Cloud Governance
de.NBI Geschäftsstelle
pbelmann@cebitec.uni-bielefeld.de

Prof. Dr. Alfred Pühler
Koordinator Deutsches Netzwerk für Bioinformatik-Infrastruktur
de.NBI Geschäftsstelle
puehler@cebitec.uni-bielefeld.de

www.denbi.de

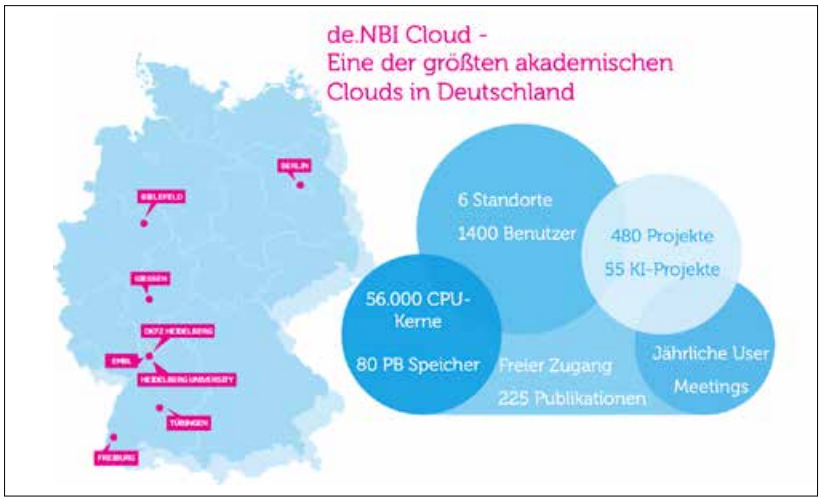


Abbildung 4: Informationen zur de.NBI Cloud
Standorte in Deutschland, technische Daten sowie Nutzung der Compute-Einrichtung (Quelle: de.NBI Geschäftsstelle).

Fotos: privat/Ortseifen, privat/Belmann, privat/Pühler

SmICS

Eine smarte Lösung zur automatisierten Ausbruchserkennung im Krankenhaus

von Antje Wulff, Claas Baier, Michael Marschollek und Simone Scheithauer für alle Mitglieder und Standorte des Use Case Infection Control des HiGHmed Projektes und die Infection Control Study Group

Hintergrund

Es ist ein seit Jahrhunderten bekanntes Phänomen: wo Menschen in Gruppen zusammenleben, kommt es zur Häufung von Infektionskrankheiten. Um sie einzugrenzen, ist es wichtig, die Wege nachzuvollziehen und wie es zu den Häufungen kommt. Je nach epidemiologischen Rahmenbedingungen kann man dabei von einem endemischen, epidemischen oder – wie im Falle der derzeitigen COVID-19 Situation – von einem pandemischen Geschehen sprechen. In Krankenhäusern – aber auch generell – wird zudem oft der Begriff Ausbruch für Erreger-und Infektionshäufungen verwendet. Das frühzeitige Erkennen solcher Ausbrüche ist eine zentrale Herausforderung.

Herausforderung

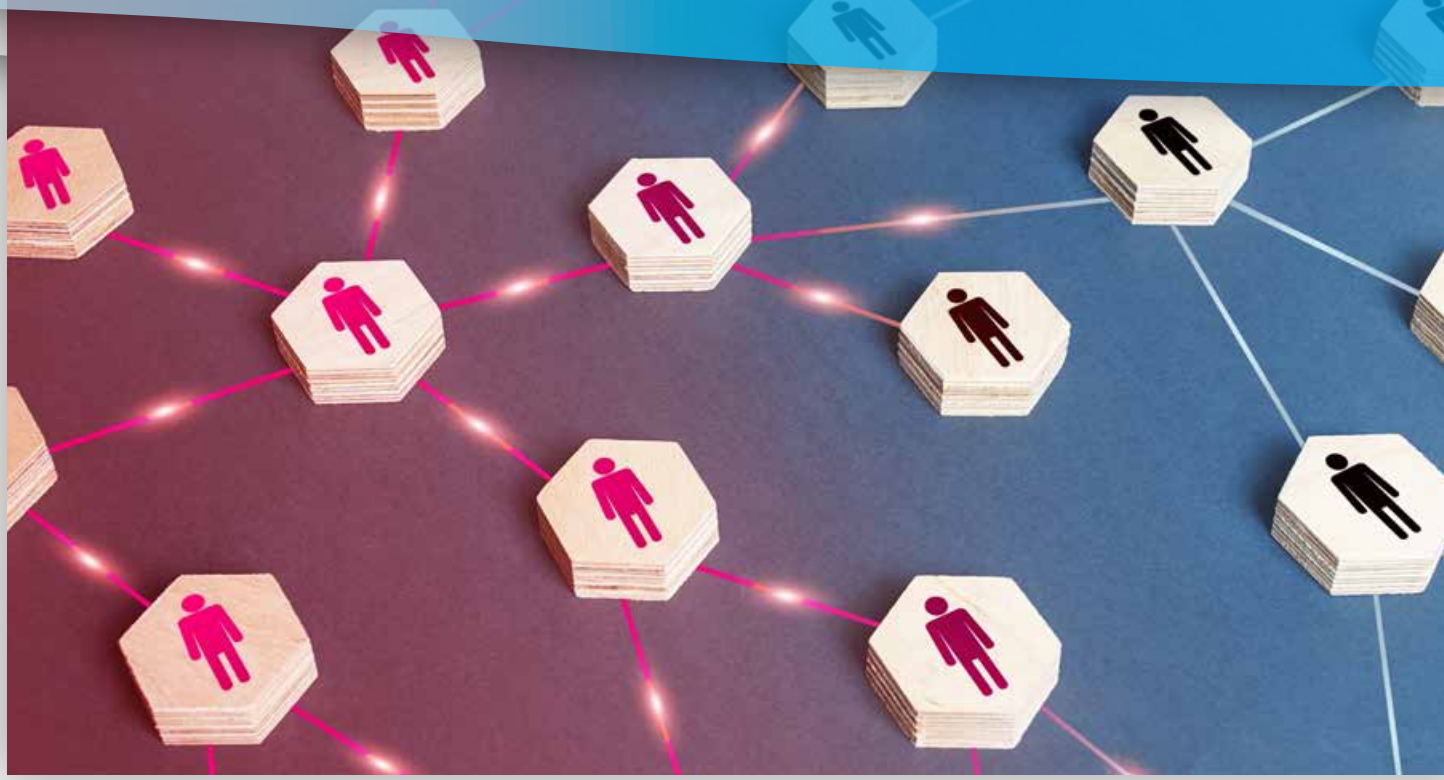
Für stationäre Gesundheitseinrichtungen wie Krankenhäuser oder Rehabilitationskliniken standen in den vergangenen Jahren Ausbruchsgeschehen, insbesondere verursacht durch die Verbreitung multiresistenter bakterieller Erreger (MRB) (David et al., 2019) , im Vordergrund. Das gilt auch weiterhin und parallel zur COVID-19 Pandemie. Ein Ausbruch mit MRB stellt alle Beteiligten im Krankenhaus vor enorme Herausforderungen. Zudem sind solche Ausbruchsgeschehen unabhängig vom Ort des Auftretens an das jeweils zuständige Gesundheitsamt meldepflichtig. Die Zusammenschau aller gemeldeten Ausbrüche durch die jährlichen Berichte des Robert-Koch Instituts lässt eine deutliche Unterschätzung der Ausbrüche in Krankenhäusern zumindest vermuten. Wir entwickeln eine smarte IT-Lösung, um ein frühzeitiges Warnsystem zu schaffen – quasi eine smarte Kontaktverfolgung, damit es gar nicht erst soweit kommt.

Die meisten durch Bakterien verursachten Ausbrüche sind durch MRB bedingt. Es handelt sich dabei um Erreger wie dem Methicillin-resistenten *Staphylococcus aureus* (MRSA) oder, in den letzten Jahren zunehmend, die Gruppe der multiresistenten Gram-negativen Erreger (MRGN) oder Vancomycin-resistente Enterokokken (VRE). Diese MRB stellen aus vielerlei Gesichtspunkten ein Problem in der Patientenversorgung dar – und das auf globaler Ebene, aber auch lokal in Deutschland und Europa.

Klinisches Leitmerkmal der bakteriellen Multiresistenz ist aus infektiologischer Sicht der Verlust der Wirksamkeit von den etablierten, gut wirksamen und gut verträglichen antibiotischen Subtanzen im Falle einer Infektion. Tritt eine Infektion mit einem solchen multiresistenten Erreger auf, geht dies häufig mit einem schlechteren klinischen Outcome und auch erhöhten Kosten für das Gesundheitssystem einher, da u. a. weniger wirksame und kostenträchtigere Reserveantibiotika eingesetzt werden müssen bzw. eine adäquate Therapie verzögert vorgenommen wird. Tritt eine Infektion im Krankenhaus (z. B. als Folge von medizinisch notwendigen Maßnahmen wie Operationen oder der Anlage von zentralen Gefäßkathetern) auf, spricht man von einer nosokomialen Infektion. Diese werden in der Mehrzahl in Deutschland nicht durch MRB ausgelöst. Oftmals handelt es sich bei MRB Nachweisen lediglich um sogenannte Kolonisationen (Besiedelungen) mit diesen Erregern auf den Schleimhäuten, z. B. im Darm (vor allem bei MRGN und VRE) und im Nasen-Rachen-Raum (MRSA). Diese Besiedlungen werden aufgrund fehlender Symptome nicht immer zeitnah erkannt, und es kann so zur Ausbreitung von solchen Erregern kommen. Die Besiedlung mit multiresistenten Erregern kann jedoch auch in eine behandlungspflichtige Infektion übergehen.

Zentrales Ziel des Use Case Infection Control

Im HiGHmed-Konsortium (https://www.highmed.org/) haben wir im Use Case Infection Control (https://www.highmed.org/highmed-use-case-infection-control) ein System zum frühzeitigen Erkennen von nosokomialen Transmissionsketten entwickelt und erforscht (Ausbruchsfrüherkennung). Es wird eine digitale Lösung entwickelt, welche die gehäufte Weiterverbreitung von multiresistenten (und perspektivisch suszeptiblen) bakteriellen Erregern automatisiert erkennt und visua-



Digitalisierung hilft, Erreger-Übertragungen frühzeitig zu erkennen und Transmissionsketten zu unterbrechen
(Foto: shutterstock © Andrii Yalanskyi)

lisiert (Baumgartl et al., 2021). Die Idee dahinter ist, dass durch sehr frühes Erkennen einer Transmissionskette ein Ausbruch möglicherweise vermieden, sicher aber in seiner Dimension reduziert werden kann, und zwar indem rechtzeitig geeignete Interventionen zur Unterbrechung umgesetzt werden können. Als Folge können Patientinnen und Patienten bestmöglich geschützt, Infektionen verhindert, Ressourcen bewusst eingesetzt und Kosten reduziert werden.

Technische Lösung und Plattform

SmICS, das Smarte Infection Control System (siehe Abbildung 1), ermöglicht eine automatisierte, Algorithmen-basierte Erkennung von Ausbrüchen. Wir nutzen dazu Daten, die sowieso in der Versorgungsroutine erhoben werden. In HiGHmed werden in enger Kooperation von universitären und außeruniversitären Einrichtungen und Institutionen zunächst an deutschen Universitätskliniken sogenannte medizinische Datenintegrationszentren (medical data integration centers, MeDICs; gesundhyte.de berichtete hierüber in der vorherigen Ausgabe 13, Dez 2020) aufgebaut. Diese Datenintegrationszentren erfüllen in dem Projekt mehrere Zwecke. Zum einen dienen Sie als Kommunikationsknoten für eine potentielle Vernetzung der partizipierenden Standorte. Zum anderen sind die MeDICs die lokalen Speicher- und Integrationsorte für administrative und versorgungsrelevante Daten, die an dem jeweiligen Standort während der Behandlung entstehen (Haarbrandt et al., 2018). Für die Erkennung von Ausbrüchen sind insbesondere Bewegungsdaten (d. h. Aufenthaltsorte der Patientinnen und Patienten im Krankenhaus) und die mikrobiologischen Befunde von entscheidender Relevanz. Ein Kernelement von SmICS ist dabei die Verknüpfung dieser beiden Datenquellen. So kann nachvollzogen werden, in welchem konkreten zeitlichen und örtlichen Kontext Erregernachweise auftreten. Algorithmenbasiert und durch standardisierte, zum Beispiel schwellwertbasierte Abfragen, können somit automatisiert potentielle Häufungs- oder Ausbruchsereignisse detektiert werden.

Alle relevanten Daten werden in den MeDICs in ihrer vollständigen Granularität mittels eines offenen Standards zur Repräsentation und Speicherung medizinischer Daten (openEHR (Beale, 2002)) modelliert (z. B. der mikrobiologische Befund (Wulff et al., 2021)) und stehen damit auf der Datenplattform der MeDICS für differenzierte Abfragen im Versorgungs- wie auch im Forschungskontext zur Verfügung. Auch die Berücksichtigung molekularer Typisierungsergebnisse – also die genaue Aufschlüsselung, ob es sich um den identischen Erreger handelt – im Kontext mikrobiologischer Befundergebnisse ist aktuell in der Umsetzung. Die robuste und einheitliche Modellierung mit dem openEHR Standard (semantische Interoperabilität) ermöglicht die Integration von Quelldaten aus unterschiedlichen, auch kommerziellen Laborinformationssystemen in eine harmonisierte, standardisierte und über Open Source-Schnittstellen zugängliche Datenplattform. Wir haben von Beginn an auf eine einheitliche Datenhaltung und konsequente Nutzung von standardisierten Abfragemöglichkeiten und Schnittstellen geachtet, daher kann die Applikation SmICS ohne wesentliche Modifikationen an dessen Programmierung an allen teilnehmenden Standorten genutzt werden.

Infektionspräventiver Mehrwert

Das frühzeitige Erkennen von sich entwickelnden Ausbrüchen ist von zentraler Bedeutung, um die Zahl der betroffenen Patienten zu begrenzen („Alert-System"). Zudem bietet die automatisierte Erkennung den Vorteil, auch schwer zu erkennende Transmissionsketten, welche bei händischer Nachverfolgung unerkannt bleiben, detektieren zu können. Die manuelle, allenfalls digital unterstützte, aber nicht automatisierte Ausbruchsuntersuchung ist zudem eine sehr zeitaufwendige Tätigkeit, die viele Ressourcen des Hygienefachpersonals bindet. Diese Ressource Zeit kann wesentlich effektiver zur Verifizierung/Falsifizierung der durch SmICS erzeugten Alerts eingesetzt werden. SmICS kann zudem einen entscheidenden Beitrag leisten, die Dunkelziffer der bislang nicht erkannten und/oder nicht gemeldeten Ausbrüche zu verringern, Patientinnen und Patienten vor Infektionen zu schützen und Personalressourcen in der Infektionsprävention effektiver einzusetzen. Es ist zu berücksichtigen, dass es zu gesteigerten Anfangsverdachtsfällen kommen wird, die durch entsprechend fachspezifische Expertise bewertet werden müssen. In der Summe kann so die Güte der Infektionsprävention ressourcenneutral optimiert werden.

Ausblick

Im Rahmen der Medizininformatikinitiative wurde und wird anhand von konkreten Anwendungsfällen im HiGHmed Projekt eine offene und konsequent auf internationale Standards bauende IT-Infrastruktur (Plattform) geschaffen, die für viele potentielle Anwendungszwecke in der vernetzten medizinischen Versorgung anwendbar ist. Im Use Case Infection Control haben wir uns zunächst auf ein System zur smarten, automatisierten Ausbruchserkennung fokussiert (SmICS) und dies in den Jahren 2018 bis 2021 entwickelt. Weitere Anwendungsbeispiele aus dem Bereich der Infektionsprävention, wie z. B. die automatisierte Surveillance von nosokomialen Infektionen oder syndromische Surveillance (basierend auf klinischen Symptomen und nicht auf Erregernachweisen), sind logische weitere Schritte. Insbesondere wenn man weitere Daten aus der Versorgung wie Arzneimittelgaben, Vitalparameter oder klinisch-chemische Laborbefunde in dem System berücksichtigt, ergibt sich eine Vielzahl von möglichen, weiteren Anwendungsfällen. Die technische Skalierbarkeit und Portierbarkeit der Plattform wurde beispielsweise bei der Entwicklung eines Tools zur Surveillance von COVID-19 Infektionen in Gesundheitseinrichtungen im Rahmen des B-FAST Projektes (https://www.umg.eu/forschung/corona-forschung/num/b-fast/) des Nationalen Forschungsnetzwerkes der Universitätsmedizin (NUM) demonstriert (siehe auch Abbildung 1).

Use Case Infection Control Study Group:

Universitätsmedizin Göttingen:
Simone Scheithauer, Martin Kaase, Patrick Fehling, Sabine Rey, Markus Suhr
Universitätsklinikum Heidelberg:
Vanessa M. Eichel, Nico T. Mutters, Klaus Heeg, Angela Merzweiler
Medizinische Hochschule Hannover:
Dirk Schlüter, Claas Baier, Michael Marschollek, Antje Wulff, Pascal Biermann
Charité Universitätsmedizin Berlin:
Petra Gastmeier, Michael Behnke, Luis Alberto Peña Diaz, Sylvia Thun, Roland Eils
WWU Münster/Universitätsklinikum Münster:
Alexander Mellmann, Hauke Tönnies, Martin Dugas, Michael Storck
TU Darmstadt:
Tatiana von Landesberger, Tom Baumgartl
Universitätsklinikum Schleswig-Holstein:
Benjamin Gebel, Cora Drenkhahn
Helmholtz Zentrum für Infektionsforschung, Braunschweig:
Thorsten Klingen, Stephan Glöckner
Robert-Koch Institut:
Benedikt Zacher, Tim Eckmanns
NEC Laboratories Europe GmBH:
Timo Sztyler, Brandon Malone

Das Forschungsprojekt in Kurzform

Das vom Bundesministerium für Bildung und Forschung im Rahmen der Medizininformatikinitiative geförderte HiGHmed Verbundprojekt (www.highmed.org) hat zum Ziel, interoperable, medizininformatische Strukturen zu generieren, um Patientendaten in digitaler Form für klinische Forschung und Lehre verfügbar zu machen. Damit soll langfristig und nachhaltig auch die klinische Patientenversorgung gestärkt werden. Indem sichere Datenintegrationszentren an den beteiligten universitären Standorten aufgebaut werden, zielt das Projekt darauf ab, wissenschaftlichem und ärztlichem Personal zielgerichtet digitale Ressourcen für Entscheidungsfindungen zur Verfügung zu stellen. Neben den universitären Standorten beteiligen sich auch viele weitere öffentliche und private Institutionen an dem HiGHmed Konsortium. Im Use Case Infection Control geht es um die Entwicklung und Erforschung eines digitalen Systems zur Ausbruchsfrüherkennung.

Referenzen:

David S, Reuter S, Harris SR, Glasner C, Feltwell T, Argimon S, *et al.* Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is driven by nosocomial spread. Nat Microbiol. 2019;4:1919–29. doi:10.1038/s41564-019-0492-8.

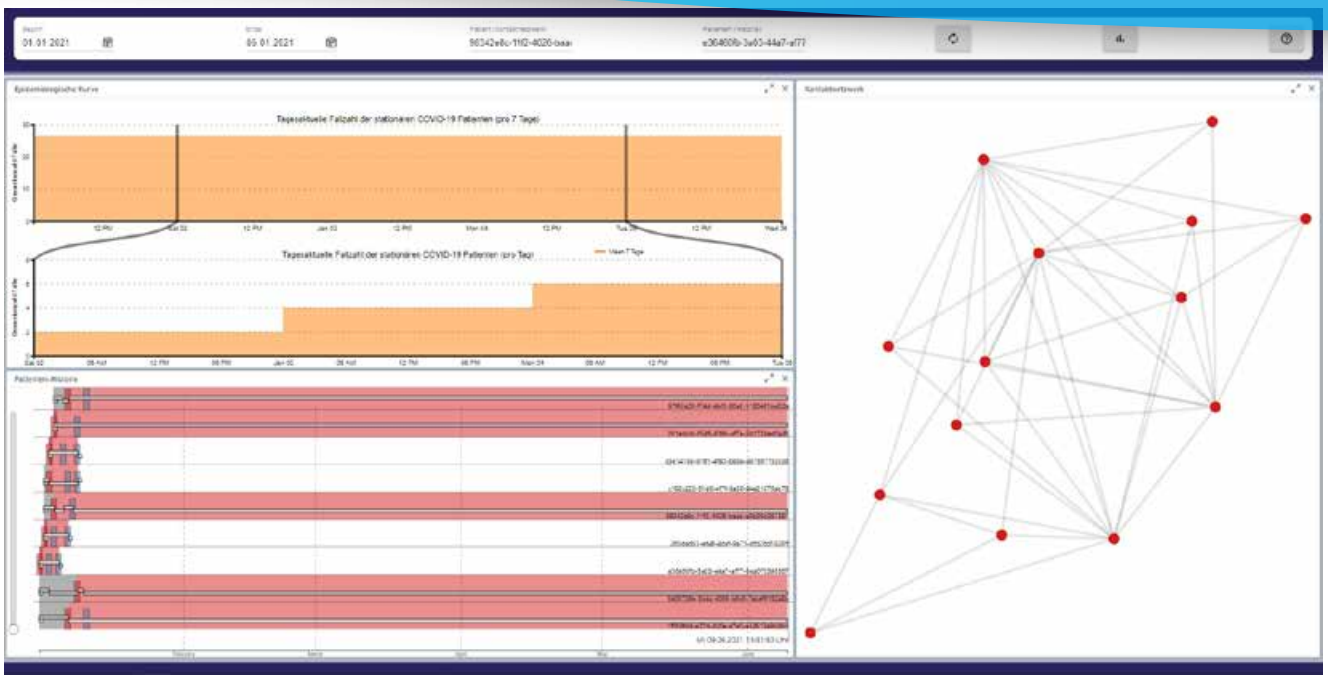


Abbildung 1: Zum besseren Verständnis von Häufungssituationen ist das Wissen um Patientenbewegungen, mikrobiologische Erregernachweise und Kontaktnetzwerken zentral. SmICS liefert eine entscheidende Hilfe bei der Visualisierung (hier am Beispiel von COVID-19) (Quelle: HiGHmed, Use Case Infection Control).

Baumgartl T, Petzold M, Wunderlich M, Hohn M, Archambault D, Lieser M, *et al.* In Search of Patient Zero: Visual Analytics of Pathogen Transmission Pathways in Hospitals. IEEE Trans Vis Comput Graph. 2021;27:711–21. doi:10.1109/TVCG.2020.3030437.

Haarbrandt B, Schreiweis B, Rey S, Sax U, Scheithauer S, Rienhoff O, *et al.* HiGHmed - An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. Methods Inf Med. 2018;57 S 01:e66–81. doi:10.3414/ME18-02-0002.

Beale T. Archetypes: Constraint-based Domain Models for Future-proof Information Systems, Revision: 2.2.1. In: Eleventh OOPSLA Workshop on Behavioral Semantics: Serving the Customer. Seattle, Washington, USA. Northeastern University, Boston; 2002. p. 16–32.

Wulff A, Baier C, Ballout S, Tute E, Sommer KK, Kaase M, *et al.* Transformation of microbiology data into a standardised data representation using OpenEHR. Sci Rep. 2021;11:10556. doi:10.1038/s41598-021-89796-y.

Kontakt:

Univ.-Prof.'in Simone Scheithauer
Direktorin des Instituts für Krankenhaushygiene und Infektiologie der Universitätsmedizin Göttingen, Klinische Leiterin des Use Case Infection Control im HiGHmed Projekt
simone.scheithauer@med.uni-goettingen.de

Univ.-Prof. Michael Marschollek
Geschäftsführender Direktor des Peter L. Reichertz Instituts für Medizinische Informatik (PLRI) der TU Braunschweig und der Medizinischen Hochschule Hannover, Leiter des PLRI-Standorts in Hannover
Ko-Leiter des Use Case Infection Control im HiGHmed Projekt
marschollek.michael@mh-hannover.de

Dr. rer. nat. Antje Wulff
Peter L. Reichertz Institut für Medizinische Informatik der TU Braunschweig und der Medizinischen Hochschule Hannover
wulff.antje@mh-hannover.de

Dr. med. Claas Baier
Institut für Medizinische Mikrobiologie und Krankenhaushygiene der Medizinischen Hochschule Hannover
baier.claas@mh-hannover.de

www.highmed.org

Fotos: Antje Wulff: Antje Wulff, Claas Baier: Claas Baier, Simone Scheithauer: umg/fskimmel, Michael Marschollek: Karin Kaiser/MHH

nationale forschungsdateninfrastruktur NFDI

Der gemeinsame Aufbau eines FAIRen Forschungsdatenmanagements in Deutschland

von Nathalie Hartl

Täglich werden neue Forschungsdaten erhoben, die uns helfen, die Welt besser zu verstehen, Probleme zu lösen und Innovationen zu entwickeln. Auch in bereits vorhandenen Datensätzen schlummert Wissen, das durch neue Verknüpfungen oder andere Forschungsfragen extrahiert werden kann. Doch aktuell sind Daten häufig nur eingeschränkt, zeitlich begrenzt und wenig standardisiert verfügbar. Ein Grund dafür ist die oft fehlende Integration eines nachhaltigen Datenmanagements. Vorhandenes Potenzial für die Nachnutzung der Daten wird dadurch verschenkt. Die Nationale Forschungsdateninfrastruktur (NFDI) setzt hier an und will das Forschungsdatenmanagement in Deutschland effektiver gestalten.

NFDI hat die Vision, dass relevante Daten aus Wissenschaft und Forschung nach den FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable) (Wilkinson *et al.,* 2016) zur Verfügung gestellt werden. Um dieses Ziel zu erreichen und die Ansprüche verschiedener Disziplinen zu berücksichtigen, ist NFDI als bundesweites, verteiltes und wachsendes Netzwerk konzipiert worden. Aktuell werden 19 NFDI-Konsortien, Zusammenschlüsse verschiedener Universitäten, Hochschulen und Forschungseinrichtungen innerhalb einer Disziplin, durch Bund und Länder gefördert.

Von der Empfehlung zum Verein

Damit alle Konsortien an einem Strang ziehen und zielgerichtet zum Aufbau einer Nationalen Forschungsdateninfrastruktur beitragen können, wurde der gemeinnützige Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V. mit Sitz in Karlsruhe gegründet. Hier laufen alle Fäden zusammen. Der Verein koordiniert das Zusammenspiel der Konsortien und widmet sich inhaltlichen und strategischen Fragestellungen, die nicht nur die Zukunft des Forschungsdatenmanagements in Deutschland, sondern auch die internationale Anbindung betreffen.

In einem 2016 veröffentlichten Papier (https://rfii.de/download/rfii-empfehlungen-2016) sprach sich der Rat für Informationsinfrastrukturen (RfII) dafür aus, eine koordinierte Forschungsdateninfrastruktur zu errichten. Bereits bestehende Strukturen sollten besser vernetzt und neue Potenziale erschlossen werden, um einen digitalen Wissensspeicher aufzubauen.

Der Vorschlag des RfII fand breiten Anklang in den Forschungscommunities. Bund sowie Länder schlossen sich der Empfehlung an und vereinbarten im November 2018 den Aufbau und die Förderung einer NFDI. Das weitgreifende Ziel ist, dass der Wissenschaftsstandort Deutschland durch ein besseres Forschungsdatenmanagement gestärkt wird und die gesamte Gesellschaft zukünftig von neuen Erkenntnissen, die aus bestehenden Datenschätzen extrahiert werden, profitieren kann.

Bis zu 30 Fachkonsortien sollen innerhalb von NFDI gefördert werden. Die Deutsche Forschungsgemeinschaft (DFG) steuert das Bewerbungsverfahren und das DFG NFDI-Expertengremium bewertet die Anträge und spricht eine Empfehlung an die Gemeinsame Wissenschaftskonferenz (GWK) aus, welche im Anschluss über die Förderung entscheidet.

Eröffnung der NFDI-Geschäftsstelle am 15.10.2020 in Karlsruhe
abgebildet sind von links nach rechts: Sabine Brünger-Weilandt (FIZ Karlsruhe), Holger Hanselka (KIT), Eva Lübke (NFDI-Direktorat), York Sure-Vetter (NFDI-Direktorat) und Frank Mentrup (Oberbürgermeister Karlsruhe) (Foto: Cynthia Ruf/KIT).

Effektives Forschungsdatenmanagement als Schlüssel zum Erfolg

Die Konsortien eint das Ziel, eine optimierte Dateninfrastruktur aufzubauen, um individuelle Fragestellungen besser bearbeiten zu können. Die Bedeutung für die Gesundheitsforschung möchten wir an drei Beispielen veranschaulichen.

NFDI4Health, die Nationale Forschungsdateninfrastruktur für personenbezogene Gesundheitsdaten, legt den Schwerpunkt auf Daten, die in klinischen, epidemiologischen und Public Health-Studien generiert werden. Das Konsortium möchte die Effektivität und Qualität in der Medizinforschung steigern und einen Beitrag dazu leisten, die Gesundheit der Bevölkerung zu verbessern. Unter anderem arbeiten beteiligte Einrichtungen daran, die Forschung im Rahmen der COVID-19-Pandemie leistungsfähiger zu gestalten. Ein Verfahren, um Gesundheitsdaten aus verschiedenen Studien und Datenbanken besser miteinander zu verknüpfen, soll etabliert werden. Dies kann den Erkenntnisgewinn zu Infektionswegen oder zum Verlauf von Krankheiten beschleunigen. Für den Austausch und die Verknüpfung von sensiblen personenbezogenen Daten müssen auch datenschutzrechtliche Fragestellungen bearbeitet werden.

Ein anderes Beispiel für die Relevanz einer funktionierenden Forschungsdateninfrastruktur liefert das Konsortium GHGA (German Human Genome-Phenome Archive). Immer mehr Genomdaten werden erhoben, um eine verbesserte Diagnostik und eine personalisierte Behandlung von Patient:innen zu ermöglichen. Die Verknüpfung möglichst vieler lokaler Datensätze und der Einsatz von KI kann wesentlich zur Erforschung von seltenen genetischen Erkrankungen beitragen. In der Onkologie soll es Mediziner:innen in der Zukunft möglich sein, beobachtete, tumorspezifische genetische Veränderungen mit ähnlichen Fällen in der GHGA Datenbank zu vergleichen. Um auf möglichst umfangreiche Datensätze zurückgreifen zu können, vernetzt sich GHGA mit dem europäischen Genomarchiv (EGA).

Ein drittes von vielen weiteren möglichen Beispielen kommt aus der Chemie: NFDI4Chem arbeitet daran, mit einem effektiven Forschungsdatenmanagement in Verbindung mit Machine-Learning-Strategien die Suche nach Wirkstoffen zu beschleunigen und Metastudien über Wirkstofffamilien hinweg zu erleichtern. Hier bestehen ebenfalls Anknüpfungspunkte an Medizin oder Strukturbiologie.

Herausforderungen bei der Bereitstellung von Daten

Eine verbesserte Auffindbarkeit, Zugänglichkeit und Langzeitarchivierung von Daten, wie sie von NFDI angestrebt wird,

bedeutet nicht, dass jeder immer uneingeschränkten Zugriff auf alle Daten hat. Vor allem sensible, personenbezogene Daten wie in der Medizin oder den Sozialwissenschaften, wirtschaftlich relevante Daten im Rahmen technischer Innovationen oder Daten mit urheberrechtlichen Bezügen wie in den Kulturwissenschaften müssen gut geschützt werden. NFDI widmet sich daher auch konsortialübergreifenden rechtlichen Fragestellungen. Ziel ist ein offener Zugriff auf Metadaten und Persistent Identifier, die unter anderem Informationen über die Voraussetzungen für den Datenabruf, die Provenienz sowie die Nutzungslizenzen enthalten. Ein effizientes Teilen von Forschungsdaten setzt voraus, dass einheitliche (maschinenlesbare) Formate genutzt werden, gemeinsam definierte Terminologien und Ontologien zugrunde liegen und Standards etabliert werden. Gerade deshalb verfolgt NFDI einen Community-Ansatz.

Nur wenn Wissenschaftler:innen auf breiter Front ein sinnvolles Datenmanagement in ihren Forschungsprozess integrieren und geteilte Standards zur Anwendung kommen, können Daten unkompliziert wiederverwendet und Potenziale bestehender Datensätze ausgeschöpft werden. Um Forscher:innen dazu zu motivieren, sich mit dem Thema Forschungsdatenmanagement auseinanderzusetzen und ihre Daten – natürlich immer unter Einhaltung rechtlicher Vorgaben – zu teilen, könnte ein Reputationssystem etabliert werden. Genau wie beim Veröffentlichen von wissenschaftlichen Beiträgen sollten Personen für das Teilen von Forschungsdaten honoriert werden.

Ein anderer Ansatz ist die Stärkung der Datenkompetenz bei (Nachwuchs-) Wissenschaftler:innen. Spezifische Aus- und Weiterbildungsangebote sollen den enormen Stellenwert eines qualitativ hochwertigen Forschungsdatenmanagements in das Bewusstsein rücken und Fähigkeiten im Bereich Data Literacy stärken.

Die geteilten Herausforderungen sollen als Querschnittsthemen in sogenannten konsortialübergreifenden Sektionen angegangen werden. In einem ersten Schritt ist die Einrichtung von vier Sektionen zu den Themen „(Meta)daten, Terminologien und Provenienz“, „gemeinsame Infrastrukturen“, „Training und Weiterbildung“ sowie „ethische und rechtliche Fragestellungen“ im Oktober 2021 geplant.

Projekte auf europäischer Ebene

Big Data in der Forschung ist eine große Herausforderung, aber eine noch größere Chance. NFDI möchte diese Chance nutzen und zusammen mit hunderten Wissenschaftler:innen Standards für ein optimiertes Forschungsdatenmanagement in Deutschland und darüber hinaus schaffen. Ähnliche Bestrebungen gibt es auch auf europäischer Ebene. Im Projekt European Open Science Cloud (EOSC), das von der Europäischen Union ins Leben gerufen wurde, arbeiten verschiedene Nationen daran, eine Cloud-Infrastruktur aufzubauen. Auch NFDI wird hier repräsentiert sein.

DER VEREIN

Seit Oktober 2020 besteht der NFDI-Verein. Das **NFDI-Direktorat** bildet den Vereinsvorstand, unterstützt von den Mitarbeiter:innen der in Karlsruhe angesiedelten Geschäftsstelle. Vereinsorgane sind das **Kuratorium** als administrativ-strategisches Kontrollgremium, der **Wissenschaftlichen Senat** als inhaltlich-strategisches Gremium, die **Konsortialversammlung**, welche disziplinübergreifend die inhaltlich-technischen Grundsätze gestaltet, sowie die **Mitgliederversammlung**. (https://www.nfdi.de/mitgliedschaft/)

Aktuell hat der Verein 188 Mitglieder (Stand 10.09.2021), darunter die Bundesrepublik Deutschland und die 16 Bundesländer als Gründungsmitglieder sowie zahlreiche Universitäten und wissenschaftliche Einrichtungen. Dank der umfassenden Vernetzung können Synergieeffekte genutzt und neue Möglichkeiten zur Zusammenarbeit erschlossen werden. Der Aufbau der Infrastruktur wird bis 2028 jährlich mit bis zu 90 Millionen Euro gefördert.

Immer mehr Forschungsdaten werden erhoben. Um diesen immensen Schatz zu nutzen, wurde NFDI ins Leben gerufen.
(Quelle: iStock © in future).

Das Verbundprojekt FAIR Data Spaces vereint die europäische Initiative Gaia-X, an der Wirtschaft und Wissenschaft gemeinsam mitwirken, und NFDI. Gefördert wird das Projekt vom Bundesministerium für Bildung und Forschung (BMBF). Gemeinsam soll ein cloudbasierter FAIRer Datenraum für den Wissenstransfer zwischen Wirtschaft und Wissenschaft aufgebaut werden.

NFDI engagiert sich also gleichzeitig auf nationaler sowie internationaler Ebene, um kompatible Lösungen mit dem globalen Wissenschaftssystem zu schaffen und Entwicklungen synergetisch voranzutreiben.

Kontakt:

Nathalie Hartl
Wissenschaftliche Referentin
PR/Kommunikation
Nationale Forschungsdateninfrastruktur (NFDI) e.V.
Karlsruhe
nathalie.hartl@nfdi.de

www.nfdi.de

Foto: NFDI

Referenzen:

Wilkinson, M. (2016): The FAIR Guiding Principles for scientific data management and stewardship. In: Sci Data 3, Artikelnr. 160018. https://www.nature.com/articles/sdata201618 (abgerufen am 02.09.2021).

Rat für Informationsinfrastrukturen (RfII) (2016): Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland. https://rfii.de/download/rfii-empfehlungen-2016/ (abgerufen am 16.07.2021).

Satzung des Vereins Nationale Forschungsdateninfrastruktur (NFDI) e.V.: https://www.nfdi.de/mitgliedschaft/ (abgerufen am 02.09.2021).

DIE KONSORTIEN

Die **neun Konsortien** DataPLANT, GHGA, KonsortSWD, NFDI4BioDiversity, NFDI4Cat, NFDI4Chem, NFDI4Culture, NFDI4Health und NFDI4Ing erhalten seit Oktober 2020 eine Förderung. In einer zweiten Runde wurde die Förderung weiterer zehn Konsortien ab Oktober 2021 beschlossen. BERD@NFDI, DAPHNE4NFDI, FAIRmat, MaRDI, NFDI4DataScience, NFDI4Earth, NFDI4Microbiota, NFDI-MatWerk, PUNCH4NFDI und Text+ werden in den kommenden Monaten in die bereits bestehenden Strukturen integriert.

GAIA-X als enabler für einen gesundheitsdatenraum

Das 2019 gestartete deutsch/französische Projekt GAIA-X geht ins dritte Jahr. Eine Momentaufnahme.

von Harald Wagener

GAIA-X steht für die Strategie, in Europa konkurrenzfähige digitale Ökosysteme aufzubauen, die auf den Prinzipen der Kooperation und Transparenz sichere und Datenschutz gewährende Anwendungen ermöglichen. Ziel ist, als Gegenentwurf zu den großen Hyperscalern aus USA und Asien, die Grundlage für verteilte und föderierte Infrastrukturen aufzubauen, die es vielen unabhängigen Teilnehmenden an diesem Ökosystem ermöglichen, innovative und neue Verwertungsmöglichkeiten für Daten zu schaffen, ohne dabei die Souveränität über diese Daten aufzugeben.

Wie ist GAIA-X organisiert?

GAIA-X ist in mehreren Gruppen organisiert, die sich gegenseitig überlappen und ergänzen: Als offizielle Körperschaft gibt es die **GAIA-X Association**, ein Non-Profit nach belgischem Recht. Die Association hat die Definitionshoheit über die offizielle GAIA-X Architektur und die Prozesse und Regeln für die Teilnahme am GAIA-X-Ökosystem. Darüber hinaus stellt sie eine quelloffene Referenzimplementierung der GAIA-X Federation Services zur Verfügung, die als Schnittstelle zwischen den Infrastruktur- und Service-Anbietern auf der einen Seite und den Anwendern auf der anderen Seite dienen.

Die Interessen der Mitgliedsstaaten und der dort ansässigen Firmen und Institutionen werden in den **GAIA-X Hubs** gebündelt. Hier werden auf regionaler Ebene Anforderungen an GAIA-X zusammengefasst. Vertreter:innen der GAIA-X Hubs besetzen ein Representative Board, das mit der GAIA-X Association im direkten Austausch steht und so sicherstellt, dass die Interessen und Anforderungen der Hubs in die Entscheidungen und Veröffentlichungen der Association einfließen können. Im Querschnitt dazu gibt es auf regionaler und auf europäischer Ebene domänen-spezifische Arbeitsgruppen (**Domain Working Groups**), die neben den allen Domänen gemeinen Anforderungen an eine föderierte Infrastruktur, die spezifischen Anforderungen zum Beispiel im Bereich Gesundheit oder Mobilität auf Basis konkreter Use Cases gebündelt an die Association weiterleiten.

Als Klammer um diese Gruppen gibt es die **GAIA-X Community**, die in Work Packages in Ergänzung zu Teilgruppen aus der Association Vorschläge für technische und regulatorische Standards erarbeitet, auf Basis derer sich das GAIA-X Ökosystem entwickeln soll.

GAIA-X und die Gesundheitsbranche

Auch im Gesundheitsbereich wird der digitale Fortschritt nur mittels Cloud-Technologien realisierbar sein. Zukünftig müssen Gesundheitsdaten aus verschiedenen Quellen verknüpft und verarbeitet werden können. Dabei müssen auf der einen Seite die hohen Sicherheits- und Datenschutzansprüche gewährleistet bleiben, auf der anderen Seite müssen die bis heute überwiegend isolierten Lösungen aus ihren abgeschotteten Systemen gelöst werden, um mit integrativen Methoden bessere Diagnoseverfahren und individualisierte Therapie-Ansätze zu ermöglichen.

Die Domäne Gesundheit ist von Anbeginn des GAIA-X Projektes in allen Bereichen stark vertreten. Schon die ersten Papiere zu GAIA-X wurden von Vertretern der Gesundheit wesentlich

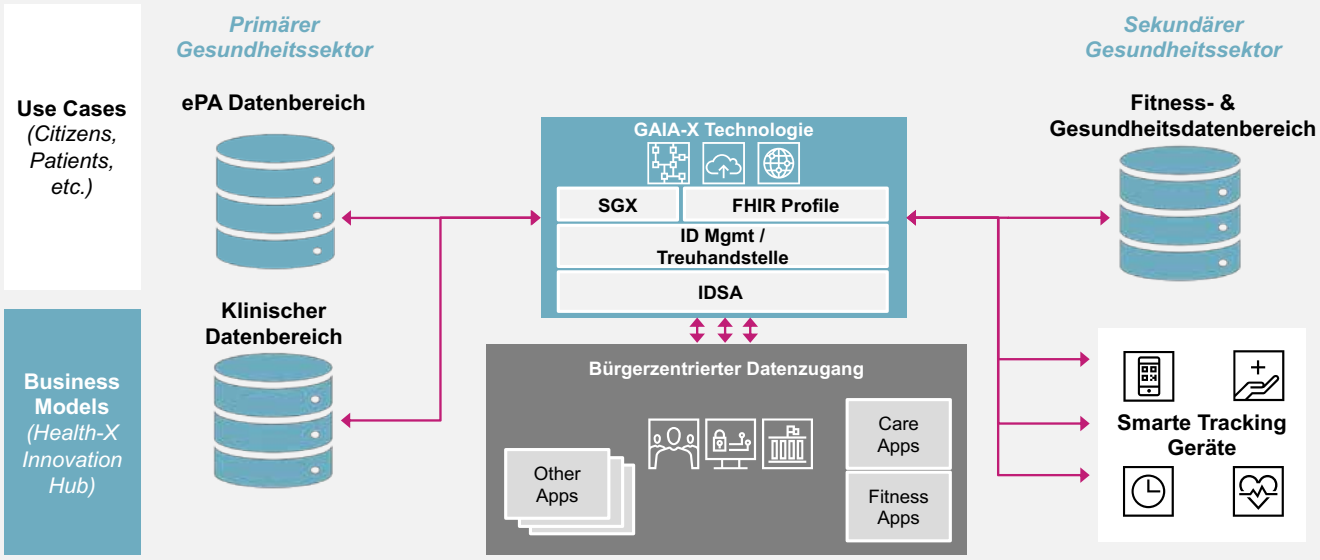


Abbildung 1: HEALTH-X dataLOFT technische Realisierung. (Quelle: BIH, LANGE und PFLANZ).

mitgestaltet. Darüber hinaus war die Gesundheits-Domäne mit über 20 Use Cases in der frühen Phase von GAIA-X wesentlich an der Ausdifferenzierung konkreter Anforderungen an ein GAIA-X Ökosystem beteiligt. Hierbei konnten auch praktische Erfahrungen mit föderierten Cloud-Systemen wie zum Beispiel de.NBI Cloud, ELIXIR und Europen Open Science Cloud (EOSC) einfließen, die in den Lebenswissenschaften auf nationaler und europäischer Ebene viele Teilaspekte einer möglichen Umsetzung bereits praktisch zeigen konnten. Heute stehen die GAIA-X Association und die EOSC Association im engen Austausch, um zukünftig eine noch bessere Kollaboration zu ermöglichen und zu eruieren, welche Schnittstellen notwendig sind, um die Interoperabilität zwischen EOSC und GAIA-X gewährleisten zu können.

GAIA-X ganz praktisch

Bis auf wenige Ausnahmen sind die Use Cases von GAIA-X bisher theoretisch betrachtet worden. In der Working Group MVG der GAIA-X Community werden nun in Piloten, die als Hackathons organisiert sind, Teilbereiche der existierenden GAIA-X-Spezifikationen auf praktische Anwendbarkeit geprüft. Auch gibt es erste kleinere geförderte Projekte im Kontext von GAIA-X wie zum Beispiel „Smart Health Connect“, das die Echtzeit-Sammlung von Gesundheitsdaten von Patient:innen mit Herzkrankheiten behandelt, die mittels KI-Methoden ausgewertet werden und Versorgern bei der Auswahl individueller Therapieschritte unterstützen soll. Dieses Projekt ist auch Teil des im August 2021 durchgeführten Pilot-003 Hackathons der MVG, der veteilte KI-Analyse von Daten behandelt.

Im Frühjahr 2021 veranstaltete das BMWi einen Förderwettbewerb mit dem Ziel Projekte zu identifizieren, die als Leuchtturmprojekte für die Umsetzung von GAIA-X in den verschiedenen Domänen förderungswürdig sind. Einer der elf Gewinner dieses Wettbewerbs ist das Projekt HEALTH-X dataLOFT unter Konsortialleitung der Charité, das sich zum Ziel gesetzt hat, die Transformation des Gesundheitssektors mit den Bürger:innen als aktive Partner:innen anstelle von passiven Leistungsempfänger:innen voranzutreiben. Die Hoheit über die Nutzung der Daten liegt über ein Data Wallet in den Händen der Bürger:innen. So wird eine sichere und vertrauenswürdige Nutzung der Daten im europäischen Gesundheitsdatenraum gewährleistet. Zweiter wesentlicher Eckpfeiler der HEALTH-X dataLOFT Idee ist die Verbindung von Daten aus dem primären und sekundären Gesundheitsmarkt, zum Beispiel über den Mechanismus der Datenspende. So werden Ökosysteme generiert, die modulare Plattformdienste wie zum Beispiel KI-Analysen anbieten, und die Verbindung von Selbsttrackingdaten mit neuen Diagnose- und Therapieansätzen wird ermöglicht. Kurz: Die Idee der datenbasierten „Real World Evidence“ Forschung wird nutzbare Realität mit unmittelbarem Nutzen für die Bürger:innen in Prävention und Versorgung.

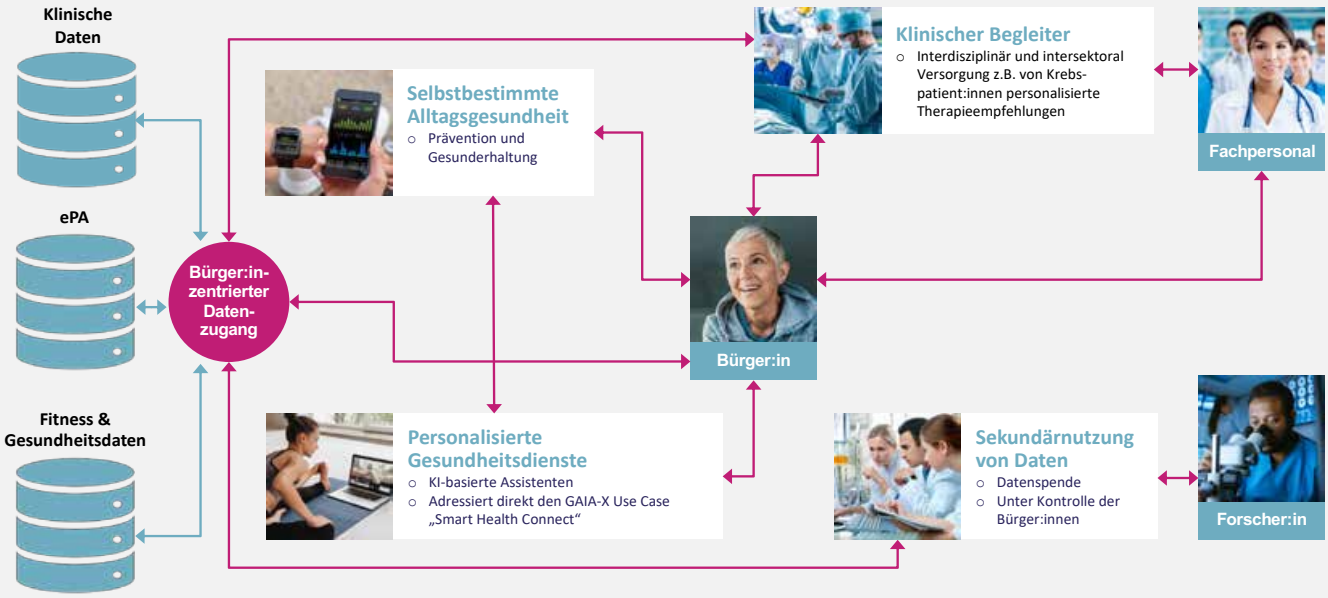


Abbildung 2: HEALTH-X dataLOFT bietet eine Vielfalt von Use Cases basierend auf einheitlichen Datenquellen, Technologien und Prozessen (Quelle: BIH, LANGE und PFLANZ, shutterstock © Alexander Raths, shutterstock © Andrey Popov, shutterstock © Arsenii Palivoda, shutterstock © ESB Professional, shutterstock © Gorodenkoff, iStock © ljubaphoto).

Gezeigt wird dies anhand verschiedener Use Cases aus den Themenfeldern Selbstbestimmte Alltagsgesundheit, Klinische Begleiter, personalisierte Gesundheitsdienste, und Sekundärnutzung von Daten. Diese Use Cases führen die Daten aus klinischer Versorgung, der elektronischen Patientenakte, und die selbst erhobenen Fitness- und Gesundheitsdaten über den Bürger:innen zentrierten Datenzugang zusammen, sodass diese entsprechend der Use Cases selbstbestimmt in jeder Phase des Lebens die bestmögliche datenbasierte Gesundheitsversorgung erhalten können.

Kontakt:

Harald Wagener
Zentrum für Digitale Gesundheit
Gruppenleiter Cloud
BIH at Charité, Berlin
Harald.Wagener@charite.de

www.hidih.org/research/health-data

GAIA-X GANZ PRAKTISCH: DIE HACKATHONS

Die Hackathons sind eine Veranstaltungsreihe aus der „Minimal Viable GAIA“ (MVG) Gruppe, deren Zweck es ist, basierend auf Beiträgen aus der Community und Anforderungen des Technical Committees, Technologien und Services konkret in kurzer Zeit umzusetzen und auf Machbarkeit und Nutzen zu prüfen.

GAIA-X Hackathons auf GitLab:
https://gitlab.com/gaia-x/gaia-x-community/gx-hackathon

Foto: privat/Wagener

FAIRe datennutzung: erfahrungen aus verbundprojekten

Anforderungen und schrittweise Umsetzung in datengetriebenen Forschungsverbünden

von Matthias Ganzinger und Matthieu-P. Schapranow

Ein neues Verbundprojekt wurde bewilligt und die Mitarbeitenden sind hochmotiviert mit dem Projekt loszulegen. Doch wie geht man am besten die gemeinsame Nutzung von Projektdaten an? Wir geben einen Überblick über die wichtigsten Aspekte, auf die man in datengetriebenen Verbundprojekten von vornherein achten sollte.

Viele Projekte erfordern die gemeinsame Nutzung von Daten, entweder Bestandsdaten, die von Partnerinstitutionen eingebracht werden, oder im Rahmen des Projektes neu erhobene Daten. Da das Datenmanagement meist gleichzeitig mit den übrigen Projektteilen startet, ergeben sich gerade in einer solchen Anfangsphase viele Fragen, die einer schnellen Antwort bedürfen, um die Daten effizient für die Forschenden im Projekt bereitzustellen.

An dieser Stelle setzt die Arbeit der im Jahr 2014 als Teil des e:Med-Forschungsprogramms zur Etablierung der Systemmedizin gegründeten Projektgruppe „Datenmanagement und Bioinformatik“ an. Sie ermöglichte den für das Datenmanagement Verantwortlichen, sich im Rahmen mehrerer Workshops über die Ansätze und Erfahrungen der einzelnen Verbundprojekte auszutauschen. Dabei wurden die in Abbildung 1 gezeigten Herausforderungen identifiziert, die offenbar von projektübergreifender Bedeutung waren und daher eine gemeinsame Bearbeitung sinnvoll erscheinen ließen.

Im Folgenden gehen wir auf die Integration verschiedener Datenquellen und die Anwendung der FAIR-Prinzipien zur Verwertung der Ergebnisse ein. Weitere Details dazu finden sich im Buchkapitel „Biomedical and Clinical Research Data Management“ des Buches „Systems Medicine: Integrative, Qualitative and Computational Approaches“, welches die Arbeit der Arbeitsgruppe ausführlicher vorstellt.

Integration heterogener Datenquellen

Die Verwendung von Realdaten ist extrem wichtig für die biomedizinische Forschung, denn nur sie spiegeln die aktuelle Wirklichkeit umfassend wider und weichen daher auch nicht selten von einer etablierten Lehrmeinung ab. Gleichzeitig sind Datenschutzbedenken bei ihrer Verwendung zu adressieren. Dazu ist es generell zielführend, vorab ein positives Votum der zuständigen Ethikkommission für das Forschungsvorhaben einzuholen. Bei der Verwendung prospektiver, also noch zu erhebender, Daten sind laienverständliche Aufklärungsgespräche und die Einholung persönlicher Einwilligungen zur Datenakquise und -verwendung der Weg der Wahl, entsprechend EU-DSGVO und Good Clinical Practices.

In manchen Fällen ergeben sich aber im Verlauf eines Projekts Verwendungsmöglichkeiten für die erhobenen Daten, die zu Beginn noch nicht bekannt waren, so dass eine entsprechende Aufklärung nicht durchgeführt werden kann. Ist das Einholen

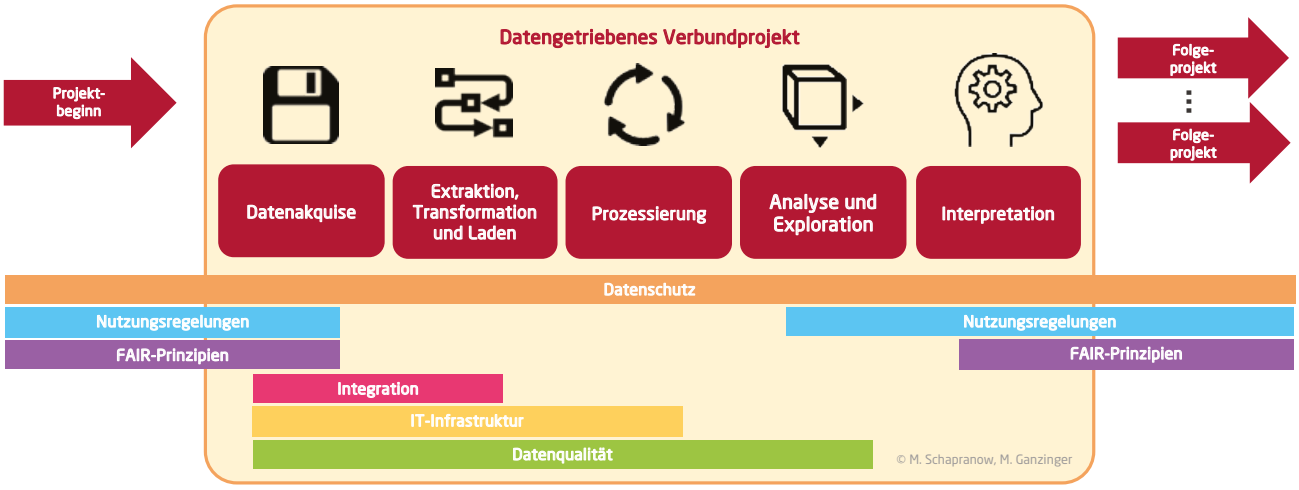


Abbildung 1: Herausforderungen für datengetriebene Verbundprojekte von der Projektidee bis zur Nutzung der Ergebnisse durch Folgeprojekte: Datenschutz, Nutzungsregelungen, FAIR-Prinzipien, Datenintegration, projektinterne IT-Infrastruktur, und Datenqualität (Quelle: M.-P. Schapranow / M. Ganzinger).

einer neuerlichen Einwilligung nicht möglich, z. B. weil es sich um retrospektive Daten aus einer anderen Studie handelt, bieten sich Anonymisierungsverfahren an. Sie werden dazu eingesetzt, um den Personenbezug aus den Daten zu verfremden bzw. entfernen, damit kein Rückschluss auf einzelne Individuen mehr möglich wird. Für die meisten Forschungsvorhaben ist dies jedoch völlig ausreichend, da meist Aussagen zu ganzen Patientenkohorten statt einzelner Individuen getätigt werden. Generell sollte man nach dem Prinzip der Datensparsamkeit verfahren, das heißt nur die Daten zu erheben bzw. extrahieren, die für die Beantwortung der Forschungsfrage erforderlich sind. Im Zweifel gilt: je weniger, desto besser.

Gemeinsame Datennutzung und die FAIR-Prinzipien

Viele öffentlich geförderte Projekte haben schon heute die Anforderung, ihre Ergebnisse für Folgeprojekte zur Verfügung zu stellen. Hierzu haben sich in den letzten Jahren die sogenannten FAIR-Prinzipien als De-facto-Standard im wissenschaftlichen Umfeld etabliert. Dabei steht die Abkürzung **FAIR** für:

- **F**indable: Die Daten sollen auffindbar sein, wie über Suchmaschinen oder Datenrepositorien, und eindeutig identifizierbar, z. B. über einen Digital Object Identifier (DOI).
- **A**ccessible: Der technische Zugang zu den Daten soll eine möglichst geringe Hürde darstellen.
- **I**nteroperable: Die Daten sollen standardisierten Definitionen entsprechen, damit die Zusammenführung mit Daten weiterer Quellen unterstützt wird.
- **R**eusable: Daten sollen von umfangreichen Metadaten begleitet werden, um das Auffinden und deren Verwendung zu erleichtern.

Das Datenmanagement: Eine Schritt-für-Schritt-Anleitung

Insbesondere die Integration verschiedener heterogener Datenquellen stellt erfahrungsgemäß eine große Herausforderung in der Verbundforschung dar. Abbildung 1 zeigt die einzelnen Prozessschritte für das Datenmanagement bei datengetriebenen Verbundprojekten.

Der erste Schritt ist die **Datenakquise**: sie umfasst sowohl das Identifizieren bestehender Datenquellen bei der Verwendung bereits verfügbarer Daten, als auch die Definition von Maßnahmen zur Datenerfassung, falls Daten erst erhoben werden müssen. Dazu gehören die Auswahl von Sensoren oder Messgeräten genauso wie die Definition von Qualitätskriterien, wie der Messgenauigkeit. In der biomedizinischen Forschung ist

die Verarbeitung personenbezogener Daten oftmals nicht zu vermeiden. Dabei muss auf die angemessene Umsetzung der gesetzlichen Anforderungen zum Datenschutz geachtet werden, etwa der Regelungen, die sich aus der EU-Datenschutzgrundverordnung ergeben.

Im zweiten Schritt erfolgt die **Extraktion** der Daten aus ihrer Quelle oder dem Messgerät, die **Transformation** in ein vorher definiertes Datenformat bzw. -modell, sowie das **Laden** der Daten in ein für das Verbundprojekt genutztes Datenrepositorium. Dieser sogenannte ETL-Prozess ist ein etabliertes Vorgehen bei der Aufbereitung von Daten für die Analyse in Wirtschaftsunternehmen. Stammen die Daten aus heterogenen Quellen ohne einheitliches Datenmodell, so erfolgt im Rahmen des ETL-Prozesses auch die *Harmonisierung* der Daten. Für diesen, im Aufwand oft unterschätzten Teil der Verarbeitung, muss, unter Beteiligung aller betroffenen Disziplinen, ein gemeinsames Datenmodell für den Verbund entwickelt werden. Daten aus relationalen Modellen, die sich gut in Datenbanktabellen darstellen lassen, sollten bevorzugt in ein gemeinsam genutztes Datenbankmanagementsystem integriert werden, da dieses nicht nur die Integrität der Daten sicherstellen kann, sondern auch standardisierte Schnittstellen, wie die Structured Query Language (SQL), zur Datenabfrage bereitstellt.

Als nächster Schritt schließt sich die **Prozessierung** der Daten an. Dazu werden sie mittels geeigneter Algorithmen überprüft und für Analysen vorbereitet. Dabei können auch Schritte zur Bereinigung der Daten oder zur Interpolation fehlender Daten erforderlich sein; dies hängt jeweils von der geplanten Nutzung der Daten im konkreten Anwendungsfall ab. Es schließt sich die **Analyse und Exploration** der Daten an, bei der spezielle Analyseverfahren zum Einsatz kommen, um Einblicke in die Daten zu erhalten. Dazu zählen zum Beispiel Verfahren des maschinellen Lernens, um Daten semi-automatisch zu klassifizieren oder um aus bestehenden Daten fehlende Einträge zu errechnen. Abschließend folgt die **Interpretation** der Analyseergebnisse im jeweiligen Anwendungsfall. Hierzu sollten Expert:innen des jeweiligen Fachbereichs, wie der Biologie oder der Medizin, hinzugezogen werden, um beispielsweise die Ergebnisse im Kontext des aktuellen Stands der Wissenschaft einzuordnen.

Ausblick: Föderierte Lernverfahren

Eine Alternative für datengetriebene Verbundprojekte stellt der Einsatz föderierter IT-Infrastrukturen dar. Statt eines einzelnen zentralen Systems, das Dienste für alle Verbundpartner zur Verfügung stellt, werden Dienste von mehreren Partner:innen dezentral bereitgestellt. Diese Dienste können aber wie ein zentraler Dienst von den Nutzer:innen verwendet werden. Da oft auf bereits bestehende, dezentrale IT-Systeme zurückgegriffen werden kann, ist es mitunter möglich, Zeit für den Aufbau und die Verwaltung einer zentralen IT-Infrastruktur einzusparen. Gleichzeitig können Daten am Ort ihres Entstehens verbleiben, was gerade bei sensiblen personenbezogenen Daten einen Vorteil gegenüber einer zentralen Lösung darstellen kann. In diesem Fall werden die für die Analysen der Daten erforderlichen Algorithmen zu den Daten transportiert und nur aggregierte Analyseergebnisse zurückgegeben.

Ein konkretes Beispiel stellt ein föderiertes, auf KI-Lernverfahren basierendes Prognosemodell dar. Ziel des Modells könnte sein, Hochrisikopatienten nach einer Nierentransplantation zu identifizieren, um Abstoßungsreaktionen oder gar Organversagen zu verhindern. Hierzu wird am Klinikstandort *A* auf lokalen Trainingsdaten ein klinisches Prognosemodell *V1* trainiert. Nun wird dieses Modell an einen weiteren Klinikstandort *B* gesandt, an dem Daten einer anderen Kohorte transplantierter Patienten zur Verfügung stehen. Testet man das Modell *V1* auf den Daten des Standorts *B*, ist die Prognosequalität schlechter als am ursprünglichen Standort, da das Modell einige der hiesigen Sonderfälle bisher nicht kannte. Wendet man nun das sogenannte *Transfer Learning* an und optimiert das Modell *V1* weiter mit den Daten von Standort *B*, so erhält man eine neue Modellversion *V2*. Dieses Vorgehen könnte man nun mit

beliebig vielen Standorten wiederholen, in der Hoffnung, dass sich die Prognosequalität durch die zusätzlichen Trainingsdaten verbessert.

Fazit

Schon heute stehen für datengetriebene Verbundprojekte eine Reihe von Ansätzen zur Verfügung. Ihre Auswahl kann jedoch nicht pauschal erfolgen, sondern sollte jeweils unter Beachtung der konkreten Projektanforderungen getätigt werden. Unabhängig davon, welche Methode letztlich gewählt wird, stellt die Zusammenarbeit über den eigenen Projektverbund hinaus eine wertvolle wissenschaftliche Bereicherung dar. Dabei geht es um die Verwendung der eigenen Ergebnisse ebenso wie um Maßnahmen des Datenmanagements in Folgeprojekten.

Steckbrief Forschungsprojekt:

Die Projektgruppe **„Datenmanagement und Bioinformatik“** wurde im Jahr 2014 als Teil des e:Med Forschungsprogramms zur Etablierung der Systemmedizin gegründet. Die Projektgruppe ermöglicht den projektübergreifenden Austausch zum Thema Datenmanagement und zur gemeinsamen Verwendung von einheitlichen Verfahren zur Optimierung der Verwendung von Daten sowohl projektintern als auch deren Verwertung nach dem Projektende.

Referenzen:

Ganzinger M, Glaab E, Kerssemakers J, Nahnsen S, Sax U, Schaadt NS, Schapranow M-P, Tiede T. Biomedical and Clinical Research Data Management. In: Wolkenhauer O, editor. Systems Medicine: Integrative, qualitative and computational approaches. Elsevier Academic Press; 2021. p. 532–43.

E:MED PROJEKTGRUPPEN – QUERSCHNITTSTHEMEN IM FOKUS

Die e:Med Projektgruppen (PG) zu inhaltlichen und methodischen Querschnittsthemen der Systemmedizin bieten die Chance zu inhaltsgetriebener Vernetzung und offenem, direktem Austausch. Die PG sind eine Eigeninitiative aus der Wissenschaft heraus auf Aufruf des e:Med Projektkomitees. Jede PG wird durch Wissenschaftler*innen geleitet, meist im Team, und steht wiederum allen e:Med Wissenschaftler*innen offen.

Die Ziele der thematisch fokussierten Projektgruppen sind:

1. aktuelle wissenschaftliche Inhalte und Technologien aus dem Bereich der Systemmedizin Forschungsverbund-übergreifend diskutieren und ggf. spezifische Initiativen starten und
2. inhaltsgetriebene Vernetzung der Mitglieder ermöglichen und den Zugang zu wissenschaftlich relevanten Informationen und deren Austausch gewährleisten.

Themen der Projektgruppen:

Datenmanagement und Bioinformatik, Datensicherheit und Ethik, Bildverarbeitung, Multiskalige sequenzbasierte Analysewerkzeuge in der Systemmedizin, MS- und NMR-basierte Omics, Modellierung von Krankheitsprozessen, *In vitro* und *in vivo* Modellsysteme.

Die PG sind ein Beispiel für den Mehrwert vernetzter interdisziplinärer Forschung, die durch e:Med, die Forschungs- und Förderinitiative des BMBF zur Etablierung der Systemmedizin in Deutschland, möglich geworden ist.

www.sys-med.de

Viele Projekte erfordern die gemeinsame Nutzung von Daten, entweder Bestandsdaten, die von Partnerinstitutionen eingebracht werden, oder im Rahmen des Projektes neu erhobene Daten. Symbolfoto: Gruppe von Wissenschaftler*innen analysiert gemeinsam Projektdaten (Quelle: shutterstock © puhhha).

Schapranow M-P, Perscheid C, Wachsmann A, Siegert A, Bock C, Horschig F, Liedke F, Brauer J, Plattner H. A Federated In-memory Database System for Life Sciences. In: Castellanos M, Chrysanthis PK, Pelechrinis K, editors. Real-Time Business Intelligence and Analytics, Springer International Publishing, 2019, ISBN: 978-3-030-24123-0

Kontakt:

Dr. Matthias Ganzinger
Wissenschaftlicher Mitarbeiter
Institut für Medizinische Informatik,
Universitätsklinikum Heidelberg
matthias.ganzinger@med.uni-heidelberg.de

https://www.klinikum.uni-heidelberg.de/personen/dr-sc-hum-matthias-ganzinger-3732

Dr.-Ing. Matthieu-P. Schapranow
Leiter der Arbeitsgruppe „In-Memory Computing for Digital Health“ und Scientific Manager Digital Health Innovations,
HPI Digital Health Center, Hasso-Plattner-Institut für Digital Engineering gGmbH,
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 Universitätsklinikum Heidelberg

Bundesministerium
für Bildung
und Forschung

Meldungen aus dem BMBF

Schnelleres Aufspüren von Antibiotika-Resistenzen

Über 50.000 Menschen erkranken in Deutschland jährlich an antibiotikaresistenten Erregern. Und die Resistenzen nehmen stetig zu. Antibiotika als eine der bislang stärksten Waffen im Kampf gegen ansteckende Krankheiten stehen damit kurz davor, wirkungslos zu werden. Die Weltgesundheitsorganisation WHO schätzt, dass in 30 Jahren zehn Millionen Menschen jährlich an resistenten Keimen sterben, wenn die Entwicklung so ungehindert weitergeht. Allerdings ist es mitunter nicht einfach, die Resistenzen aufzuspüren. Die Betroffenen erhalten dadurch häufig zunächst wirkungslose Antibiotika. „Gerade bei schweren Fällen kann hier wertvolle Zeit verloren gehen und die gefährlichen Keime infizieren weitere Menschen“, erklärt Professor Dr. Alexander Goesmann von der Justus-Liebig-Universität Gießen.

Mit Methoden des maschinellen Lernens wollen er und sein Forschungsteam Resistenzen schneller aufspüren und langfristig Angriffspunkte für neue Antibiotika identifizieren. Das Bundesministerium für Bildung und Forschung (BMBF) fördert das interdisziplinäre Team aus Gießen und Marburg dabei mit rund einer Million Euro im Rahmen der Fördermaßnahme „Computational Life Sciences“.

Resistenzen sind komplex

Multiresistente Erreger schnell und sicher zu identifizieren ist essenziell, wenn es darum geht, den Infizierten effektiv zu helfen und eine Besiedelung weiterer Personen zu vermeiden. Moderne Genanalysen können schnell Klarheit schaffen, wenn ein einzelnes Gen die Resistenz vermittelt. „Die Resistenzen sind jedoch häufig deutlich komplexer. Es gibt verschiedene Abstufungen und häufig sind weitere Gene beteiligt“, erklärt Goesmann. So können einzelne Mutationen harmlos, aber in Kombination mit weiteren Mutationen gefährlich sein. Um diese komplexen Zusammenhänge zu identifizieren, setzt das Forschungsteam auf Methoden des maschinellen Lernens. „Die Algorithmen können riesige Datensätze analysieren und die Zusammenhänge aufspüren“, sagt Professor Dr. Dominik Heider von der Philipps-Universität Marburg. Langfristig wollen die Wissenschaftlerinnen und Wissenschaftler auch bei solch komplexen Resistenzmechanismen einen schnellen und einfachen Test auf Resistenzen möglich machen.

Für die Erhebung der Daten kooperiert das Team von Professor Goesmann mit der Abteilung für medizinische Mikrobiologie in Gießen unter der Leitung von Professor Dr. Chakraborty. Sein Team nimmt regelmäßig Proben von Patientinnen und Patienten, die mit multiresistenten Erregern besiedelt sind, und untersucht diese mit molekularbiologischen Methoden. Neben der Genomsequenz der Bakterien interessiert sein Forschungsteam vor allem, wie die Bakterien auf Antibiotika reagieren. Hierzu überprüfen sie, welche Gene die Bakterien nach Gabe von Antibiotika an- und ausschalten, aber auch, wie sich das Wachstum der Bakterien unter bestimmten Konzentrationen verschiedener Antibiotika verändert. Dafür lassen die Forschenden die Bakterien im Labor unter standardisierten Bedingungen wachsen, behandeln sie zu definierten Zeitpunkten mit einer festgelegten Menge an Antibiotika und analysieren, wie sich das Wachstum verändert.

Ansatzpunkte für neue Antibiotika

Mithilfe der Algorithmen lernt dann ein Computerprogramm aus diesen Daten, welche Gene vorhanden und angeschaltet sein müssen, damit die Bakterien trotz der

Antibiotikaresistenzen sind eine große Gefahr. Mithilfe von KI können Kliniken sie zukünftig schneller erkennen und entsprechende Maßnahmen ergreifen.

Bildnachweis: Adobe Stock / eremit08

Zugabe einer bestimmten Gruppe von Antibiotika weiterwachsen. So kann dann auch trotz der komplexen Muster schnell klassifiziert werden, ob der untersuchte Bakterienstamm bereits resistent gegen bestimmte Antibiotika ist oder nicht. Für den Einsatz in der Praxis wollen die Bioinformatikerinnen und Bioinformatiker auch verstehen, warum das Programm die Entscheidung trifft. Letztlich gilt es also, herauszufinden, was sich auf molekularer und genetischer Ebene bei den Bakterien verändert hat. „Wie gefährlich die Bakterien sind, entscheiden nicht nur die Resistenzmechanismen, sondern auch andere Eigenschaften, wie ihr Potenzial zur Verbreitung."

In Zukunft könnte die computergestützte Analyse eine schnellere und sicherere Diagnostik als bisher ermöglichen. Die behandelnden Ärztinnen und Ärzte erhalten dann bereits kurz nach Einlieferung der Betroffenen in die Klinik eine Einschätzung, ob einer oder mehrere der Erreger resistent sind, ob wirksame Antibiotika verfügbar sind und ob eine Isolierung der Betroffenen angebracht ist. Um einen sicheren Einsatz bei den Patientinnen und Patienten zu ermöglichen, muss das Forschungsteam das Verfahren jedoch noch weiter validieren und zertifizieren. „Wir schätzen, dass wir in fünf bis zehn Jahren so weit sind", so Goesmann. „Darüber hinaus hoffen wir allerdings, dass unsere Analysen Ansatzpunkte für neuartige Antibiotika liefern und wir dazu beitragen können, neue Waffen im Kampf gegen multiresistente Erreger zu entwickeln."

Weitere Informationen finden Sie unter:

YouTube-Film zum Projekt:
youtube.com/watch?v=UCv8QEKwXts

Bessere Versorgung durch digitale Innovationen

Die digitale Zukunft der Medizin hat längst begonnen: So können tragbare Sensoren beispielsweise die Vitaldaten zur Herzgesundheit von Patientinnen und Patienten mit Herzschwäche auch im häuslichen Umfeld erfassen und direkt an die Klinik übermitteln – Vorboten kritischer Entwicklungen können Behandelnde so frühzeitig erkennen und gezielt dagegen vorgehen. Und intelligente Smartphone-Apps helfen Ärztinnen und Ärzten längst, auch in Notfallsituationen die bestmöglichen Therapieentscheidungen zu treffen. Zentraler Wegbereiter solcher Innovationen in Deutschland ist die Medizininformatik-Initiative (MII) des Bundesministeriums für Bildung und Forschung (BMBF). In vier Konsortien, an denen sich alle deutschen Unikliniken beteiligen, hat die MII zukunftweisende Dateninfrastrukturen aufgebaut und innovative IT-Lösungen entwickelt, die schon heute die Versorgung der Patientinnen und Patienten vieler Unikliniken verbessern.

Transfermodelle für die Medizin der Zukunft

Für die meisten Menschen ist jedoch nicht die Uniklinik, sondern die Hausarztpraxis oder das regionale Krankenhaus die erste Anlaufstelle für die medizinische Versorgung. Daher gilt es nun, die von den Unikliniken geleistete Pionierarbeit zur Digitalisierung der Medizin möglichst in alle Bereiche des Gesundheitssystems einfließen zu lassen: von der ambulanten Versorgung in der Hausarztpraxis über den stationären Aufenthalt im örtlichen Krankenhaus bis zur Versorgung in Pflege- und Rehabilitationseinrichtungen. Dieser Transfer ist eine anspruchsvolle Aufgabe. Dafür zunächst modellhafte Lösungen zu entwickeln und zu optimieren, damit sie zukünftig allen Menschen zu Gute kommen können – das ist die Aufgabe der sechs vom BMBF geförderten Digitalen FortschrittsHubs Gesundheit. Bis 2025 stellt das BMBF für diese Leitinitiative seiner Digitalstrategie rund 50 Millionen Euro bereit. Die FortschrittsHubs zahlen zudem auf die Strategie Künstliche Intelligenz der Bundesregierung ein.

Im Fokus: Innovationen, von denen die Menschen vielfältig profitieren

Die Digitalen FortschrittsHubs Gesundheit fokussieren auf Anwendungsfälle, bei denen der Datenaustausch zwischen regionalen Akteuren, Unikliniken und Forschungseinrichtungen vielen Menschen zu Gute kommt. So widmen sich mehrere FortschrittsHubs der Krebsmedizin und verfolgen die Ziele der Nationalen Dekade gegen Krebs, die das BMBF 2019 ins Leben rief. Auch das Pandemiemanagement liegt im Themenspektrum der FortschrittsHubs. Die hier entwickelten und erprobten Lösungen sollen helfen, das Gesundheitssystem gegen künftige Krisen noch besser zu wappnen. Ebenso vielfältig wie die medizinischen Themen sind auch die technischen Lösungen, die dabei zum Einsatz kommen – das zeigt die folgende Auswahl:

Plattformen für den Datenaustausch sollen die Erforschung, Diagnostik und Therapie von Krankheiten durch die intelligente Verknüpfung und den Austausch von Informationen verbessern – und zwar über die Grenzen von Klinik, ambulanter Versorgung, Rehabilitation und Pflege hinweg. Zugleich gilt es, die Kommunikation zwischen Behandelnden unterschiedlicher Professionen – darunter Medizin, Psychologie und Physiotherapie – sowie zwischen den Behandelnden und

ihren Patientinnen und Patienten zu fördern. Denn der Austausch und die Verfügbarkeit von Daten sind wichtige Voraussetzungen, um Patientinnen und Patienten personalisiert und – etwa im Verlauf einer Krebserkrankung – auch langfristig von der Klinik bis zur Rehabilitation und darüber hinaus bestmöglich zu versorgen.

Künstliche Intelligenz soll Ärztinnen und Ärzten beispielsweise in zeitkritischen Notfallsituationen helfen, schnell die richtigen Maßnahmen einzuleiten und Leben zu retten – etwa bei der Versorgung von Schlaganfallpatienten im Rettungseinsatz.

Telemedizin und mobile Sensoren, die Gesundheitsdaten an behandelnde Ärztinnen und Ärzte übermitteln, sollen die Versorgung auch in ländlichen Regionen verbessern. Die dabei erfassten Daten werden auch helfen, Volkskrankheiten wie psychische Erkrankungen und Krebs noch besser zu verstehen und Therapien zu optimieren.

Wie Digitale FortschrittsHubs funktionieren

Ausgangspunkt eines Hubs ist das Datenintegrationszentrum einer Uniklinik. Diese Zentren werden derzeit als IT-Infrastrukturen an fast allen Unikliniken im Rahmen der Medizininformatik-Initiative aufgebaut. Dieses vernetzt sich mit regionalen Partnern – darunter Krankenhäuser, Arztpraxen, Rehabilitations- und Pflegeeinrichtungen sowie Rettungsdienste. Auch Forschungseinrichtungen und Krankenkassen sind Partner der Hubs. Sie alle teilen und nutzen ihre Daten gemeinsam. Wissenschaft, IT, Versorger, Ärztinnen und Ärzte, Pflegepersonal, Patientinnen und Patienten arbeiten dabei eng zusammen.

Weitere Informationen finden Sie unter:

medizininformatik-initiative.de

KI verbessert Therapie von COPD

Erste Symptome wie Husten und Atemnot beim Treppensteigen ignorieren die Betroffenen häufig; dabei ist die atemwegsverengende Lungenerkrankung COPD (chronic obstructive pulmonary disease) eine ernst zu nehmende Erkrankung. Unbehandelt entstehen irreparable Lungenschäden und die Krankheit verschlimmert sich stetig. Im späteren Verlauf haben viele Patientinnen und Patienten bereits Atemnot, wenn sie nur auf dem Sofa sitzen. Mit Medikamenten lassen sich die Symptome abmildern und eine Verschlimmerung zumindest verlangsamen.

Bislang gibt es jedoch keine Methode, um den Krankheitsverlauf ganz aufzuhalten oder gar umzudrehen. Problematisch ist, dass die Erkrankung bei den Patientinnen und Patienten sehr unterschiedlich verläuft. Insbesondere weitere Beeinträchtigungen wie Diabetes oder Osteoporose spielen eine große Rolle. Je nach Begleiterkrankung und gesundheitlichem Zustand der Erkrankten sind daher jeweils andere Therapieansätze notwendig. Hier setzt das Team um den Marburger Lungenspezialisten Professor Dr. Bernd Schmeck an. Mit Hilfe von Methoden, die sich auf maschinelles Lernen stützen, wollen sie Patientinnen und Patienten besser klassifizieren und so eine personalisierte Behandlung ermöglichen. Dabei unterstützt sie das Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Fördermaßnahme ERACoSysMed mit knapp 600.000 Euro.

Das Team um Schmeck verwendet Daten von mehreren Tausend Patientinnen und Patienten. Sein Marburger Kollege Professor Dr. Claus Vogelmeier konnte eine Studiengruppe mit Daten von über 2.700 Behandelten über einen Zeitraum von mehr als fünf Jahren aufbauen. Neben klinischen Daten zum Krankheitsverlauf erfassen die Medizinerinnen und Mediziner auch speziellere Daten aus Laboruntersuchungen, so etwa bestimmte Proteine und molekulare Marker, um noch mehr über die Erkrankung zu erfahren. „Bei der Fülle an Daten brauchen wir einen Ansatz mit künstlicher Intelligenz – für einen Menschen ist das ohne dieses Hilfsmittel nicht mehr erfassbar“, erklärt Schmeck.

Daten helfen heilen! Nicht nur an den Unikliniken, auch in Hausarztpraxen, im Krankenhaus um die Ecke, in Pflege- und Rehabilitationseinrichtungen sollen digitale Innovationen künftig helfen, die Versorgung der Menschen zu verbessern.

Bildnachweis: Adobe Stock/lenetsnikolai

Begleiterkrankungen im Blick behalten

Der Algorithmus lernt anhand Tausender Kombinationen aus Labordaten und Krankheitsverläufen, Übereinstimmungen bei verschiedenen Patientinnen und Patienten zu erkennen und Gruppen mit ähnlichen Symptomen und Verläufen zu bilden. Diese Ergebnisse wiederum vergleichen die Forscherinnen und Forscher im Rahmen einer europäischen Zusammenarbeit mit Daten aus einer Studiengruppe in den Niederlanden. Insbesondere wollen sie so herausfinden, welche Parameter und Daten für den Krankheitsverlauf relevant sind und welche nicht. Ihr Ziel ist es, die Patientinnen und Patienten möglichst früh einer Gruppe zuordnen zu können und so eine optimale Behandlung zu ermöglichen.

Gerade bei einer Erkrankung wie COPD, die vorwiegend Menschen ab dem 50. Lebensjahr trifft und häufig mit weiteren gesundheitlichen Problemen wie Herz-Kreislauf-Erkrankungen, Diabetes, Osteoporose oder auch psychischen Erkrankungen einhergeht, ist es wichtig, alle Aspekte im Blick zu behalten. Wenn Erkrankte in die Klinik kommen, zielt die Behandlung in der Regel auf das augenscheinlich größte Problem und konzentriert sich darauf, zunächst die Lungenfunktion zu verbessern. In einigen Fällen ist es jedoch notwendig, zuerst die Begleiterkrankung zu behandeln, um eine Verbesserung des Gesamtzustands zu erreichen: So ist es möglich, dass die Patientinnen und Patienten aufgrund von Osteoporose Schmerzen bei Bewegungen haben und deshalb weniger atmen. Auch hier erhofft sich das Team eine Unterstützung durch die neue Software. Nach Auswertung der vorliegenden Daten könnte das System beispielsweise eine Knochendichtemessung vorschlagen und die behandelnden Ärztinnen und Ärzte auf die genannte Problematik mit der Osteoporose aufmerksam machen.

Erste Tests der Software in einem Jahr

Damit ihre Software schnell den Weg in die Anwendung finden kann, kooperieren die Forscherinnen und Forscher bereits jetzt mit einer Firma, die Krankenhaussoftware entwickelt und einsetzt. Um im klinischen Alltag konkret Hilfestellung geben zu können, sollte die Software möglichst einfach, unkompliziert und ohne großen Zeitaufwand bedient werden können. Idealerweise sollte die Software im Rahmen der normalen Krankenhaus-IT installiert werden und auf die vorhandenen Patientendaten zugreifen, um den behandelnden Ärztinnen und Ärzten gleich vor Ort Vorschläge hinsichtlich Diagnostik und Therapie geben zu können. „Bereits in einem Jahr soll die Software für erste Tests in der Klinik eingesetzt werden“, so Schmeck. Langfristig soll es eine größer angelegte Testphase geben und dann auch eine Zulassung als Medizinprodukt erfolgen.

Auch Untersuchungsdaten aus dem Labor fließen in den komplexen Computeralgorithmus zur besseren Klassifizierung der COPD-Patientinnen und -Patienten ein.

Bildnachweis: Philipps-Universität Marburg

Die Unterteilung in Patientengruppen könnte neben der Unterstützung bei Therapieentscheidungen weitere Möglichkeiten bieten, etwa bei der Testung neuer Medikamente. Denn was einem Betroffenen hilft, führt beim nächsten mit anderen Begleitumständen noch lange nicht zu einer Besserung. Die Forscherinnen und Forscher erhoffen sich durch die Erkenntnisse der Software somit neben einem verbesserten Verständnis der Erkrankung auch neue Ansätze für die Entwicklung neuer Therapien.

Weitere Informationen finden Sie unter:

gesundheitsforschung-bmbf.de/de/systemmedizin-9458.php

BMBF-Newsletter:

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

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

twitter.com/BMBF_Bund

facebook.com/bmbf.de

instagram.com/bmbf.bund

news aus dem BIH

künstliche intelligenz auf der intensivstation

Die x-cardiac GmbH hat ihre Zulassung als Medizinproduktehersteller in der EU und zugleich die Zulassung für ihr erstes Medizinprodukt „x-c-bleeding“ erhalten. Die auf Künstlicher Intelligenz basierende Software kann in kurzer Zeit das Risiko postoperativer Nachblutungen von Patient*innen vorhersagen. Spezialisten am Deutschen Herzzentrum Berlin (DHZB) haben die Software entwickelt und wurden dabei vom Berlin Institute of Health in der Charité (BIH) mit verschiedenen Förderprogrammen unterstützt.

Die x-cardiac GmbH hat mit der K.I.-Software x-c-bleeding das erste zugelassene Medizinprodukt zur Vorhersage postoperativer Nachblutungen auf den Markt gebracht. Das Medizin-Start-up entwickelt KI-basierte Software, mit deren Hilfe Ärzt*innen postoperative Komplikationen nach schweren Herzoperationen erkennen können, noch bevor eine kritische Lage erreicht ist. Das Team von x-cardiac hat die Software mithilfe von anonymisierten Daten von knapp 50.000 Patient*innen am DHZB „trainiert“ und im realen Klinikbetrieb auf den DHZB-Intensivstationen erprobt.

Das x-cardiac System in Betrieb auf der Intensivstation IPS 1 des Deutschen Herzzentrums Berlin. Foto: Külker/DHZB

Professor Alexander Meyer, Facharzt für Kardiologie und Informatiker am DHZB und Geschäftsführer von x-cardiac, hat die Software entwickelt. „Wir sehen den Bedarf an Systemen, die Entscheidungen am Point of Care unterstützen. Außerdem stehen die Kliniken vor großen Herausforderungen bei der Digitalisierung. Mit x-c-bleeding können wir ein ausgereiftes und praxistaugliches Angebot machen. Die im Krankenhauszukunftsgesetz (KHZG) verankerten Fördermöglichkeiten leisten hier einen wichtigen Beitrag.“

Meyer war während der Entwicklung und Testphase von x-c-bleeding als Stipendiat des BIH Charité Clinician Scientist Programms gefördert worden. Das Validation Fund SPARK-BIH Programm und das Digital-Health-Accelerator-Programm des BIH haben schließlich den Weg zur Anwendung und Vermarktung der Software begleitet und gemeinsam mit dem DHZB die Gründung der x-cardiac GmbH unterstützt.

Über x-cardiac:

Die Systeme von x-cardiac beruhen auf sogenannten „rekurrenten neuronalen Netzwerken“, also künstlicher Intelligenz, die unter Verwendung der gespeicherten und anonymisierten Daten von knapp 50.000 Patient*innen am DHZB entwickelt wurde, um Nachblutungen und akutes Nierenversagen frühzeitig erkennen zu können. Diese lebensbedrohlichen Komplikationen können nach Operationen am Herzen oder den herznahen Gefäßen auftreten. Je früher sie erkannt werden, desto größer ist die Aussicht auf erfolgreiche Behandlung.

Die Systeme wurden in den Intensivstationen des DHZB seit April 2018 im realen Klinikbetrieb erprobt und werden nun in die zertifizierten Medizinprodukte „x-c-bleeding“ und „x-c-renal-injury“ überführt. Diese erfüllen alle Anforderungen zur Förderfähigkeit als klinisches Entscheidungsunterstützungssystem im Rahmen des Krankenhauszukunftsgesetzes (KHZG) des Bundesgesundheitsministeriums, das die Digitalisierung in deutschen Krankenhäusern mit insgesamt bis zu 4,3 Milliarden Euro fördert.

warum kinder nicht so schwer an COVID-19 erkranken

Kinder sind ebenso häufig mit dem neuen Coronavirus SARS-CoV-2 infiziert wie Erwachsene, haben aber ein sehr niedriges Risiko, schwer an COVID-19 zu erkranken. Ein Team aus Wissenschaftler*innen des Berlin Institute of Health in der Charité (BIH), der Charité –Universitätsmedizin Berlin, des Universitätsklinikums in Leipzig sowie des Deutschen Krebsforschungszentrums (DKFZ) in Heidelberg hat mit Einzelzellanalysen die Ursache hierfür herausgefunden. Sie konnten zeigen, dass das kindliche Immunsystem in den oberen Atemwegen wesentlich stärker aktiv ist als bei Erwachsenen und damit besser gewappnet im Kampf gegen das Virus. Ihre Ergebnisse haben die Forscher*innen im Fachjournal Nature Biotechnology veröffentlicht.

Seit Beginn der Pandemie ist das BIH Team um Irina Lehmann, Leiterin der AG Molekulare Epidemiologie am BIH, und Roland Eils, Direktor des Zentrums für Digitale Gesundheit am BIH, den COVID-19 Krankheitsmechanismen auf der Spur. Basierend auf Einzelzellanalysen aus dem Nasen-Rachen-Raum von Erwachsenen haben die BIH Forscher*innen die an schweren Erkrankungen beteiligten Zellen und Signalwege identifiziert und darauf aufbauend eine klinische Studie in die Wege geleitet, die eine neue Therapie für schwerkranke Patient*innen erprobt. Aktuell rücken jedoch Kinder immer mehr in das Zentrum des Interesses, da diese noch nicht durch Impfungen geschützt sind und sich derzeit viel häufiger anstecken. „Wir wollten deshalb vergleichende Einzelzellanalysen bei Kindern und Erwachsenen durchführen, um daraus zu lernen, wie der Schutz gegen COVID-19 funktionieren kann“, sagt Roland Eils.

Das Team um Professor Markus Mall, Direktor der Klinik für Pädiatrie mit Schwerpunkt Pneumologie Immunologie und Intensivmedizin an der Charité, hatte für diese Untersuchungen Proben aus der Nasenschleimhaut von gesunden und von mit SARS-CoV-2 infizierten Kindern gesammelt. „Die meisten der infizierten Kinder hatten nur leichte Symptome wie Schnupfen oder leicht erhöhte Temperatur“, erklärt Markus Mall. In den Proben führten die BIH Forscher*innen Einzelzell-Transkriptom-Analysen durch, sie untersuchten, welche Gene in welchen Zellen wie häufig abgelesen wurden. Ihre Ergebnisse verglichen sie mit entsprechenden Proben von nicht-infizierten und infizierten Erwachsenen.

Es zeigte sich, dass die Zellen der Nasenschleimhaut von gesunden Kindern bereits in erhöhter Alarmbereitschaft waren und vorbereitet für den Kampf gegen SARS-CoV-2. Für eine schnelle Immunantwort gegen das Virus müssen sogenannte Mustererkennungsrezeptoren aktiviert werden, die das Erbgut des Virus, die Virus-RNA, erkennen und eine Interferon-Antwort einleiten. In den untersuchten kindlichen Zellen war dieses Mustererkennungssystem deutlich stärker ausgeprägt als bei Erwachsenen, so dass das Virus, sobald es in der Zelle ankommt, schnell erkannt und bekämpft werden kann.

***„Wir wollten verstehen, warum die Virusabwehr bei Kindern offenbar so viel besser funktioniert als bei Erwachsenen“*, erklärt Professorin Irina Lehmann, Leiterin der AG Molekulare Epidemiologie am BIH.** Foto: D. Aussenhofer

Um ihre kleinen Patienten über das Prozedere der Behandlung aufzuklären, hat sich die AG Spuler ein Comic ausgedacht. Copyright: AG Spuler/Charité

neue muskeln für die blase

Etwa sieben Jungen werden jedes Jahr in Deutschland geboren, deren Harnröhre und Blasenschließmuskel unvollständig ausgebildet sind. In einer klinischen Studie wollen Wissenschaftler*innen vom ECRC, dem gemeinsamen Experimental and Clinical Research Center der Charité – Universitätsmedizin Berlin und des Max-Delbrück-Centrums für molekulare Medizin in der Helmholtz-Gemeinschaft (MDC), nun prüfen, ob den Kindern mithilfe einer Transplantation ihrer eigenen Muskelstammzellen geholfen werden kann. Das BIH hat das Vorhaben mit seinem Spark-BIH-Programm auf dem Weg vom Labor in die Klinik mit einer Million Euro unterstützt.

Bei der Epispadie führt eine vorgeburtliche Entwicklungsstörung zu einer abnormalen Lage und Spaltbildung der Harnröhre und einem unvollständig ausgebildeten Blasenschließmuskel. „Diese Kinder sind oft lebenslang inkontinent, was mit einer hohen psychologischen Belastung für die Betroffenen und deren Familien einhergeht“, erklärt Professorin Simone Spuler vom ECRC, Spezialistin für Stammzell- und Muskelforschung. „Wir haben deshalb überlegt, wie wir ihnen mit unserer Expertise helfen können.“

Die Wissenschaftler*innen um Simone Spuler hatten eine Methode entwickelt, mit der sie regenerationsfähige Muskelstammzellen aus Muskelgewebe isolieren können. „Wir nehmen eine Gewebeprobe aus dem Oberschenkel und isolieren daraus Muskelstammzellen. Diese vermehren wir anschließend auf ein Vielfaches und spritzen sie direkt in die Defektstelle des Blasenschließmuskels.“ Bei Ratten führte das tatsächlich dazu, dass sich ein neuer Schließmuskel bildete, der auch funktionstüchtig war.

„Doch trotz dieser ermutigenden Ergebnisse konnten wir nicht sofort eine klinische Studie mit Betroffenen beginnen“, erklärt die Medizinerin. „Denn nur Zellen, die in einem pharmazeutischen Herstellungsverfahren, der so genannten „good manufacturing practice“ kurz GMP, produziert werden, dürfen im Menschen angewandt werden.“ Die Auflagen der Zulassungsbehörden, in Deutschland das Paul-Ehrlich-Institut, verlangen, dass nur speziell dafür akkreditierte Labore die Tierversuche für eine klinische Studie durchführen dürfen. „Ungefähr 300 km östlich von Chicago, mitten in Michigan, gab es ein solches Labor“, berichtet Simone Spuler. „Die Vorbereitungen, die Einarbeitung und die Durchführung der Versuche waren unglaublich zeit- und kostenintensiv. Das hätten wir ohne die Unterstützung durch das BIH-Spark-Programm nicht geschafft!“ Eine Million Euro stellte das BIH den Muskelforscher*innen um Simone Spuler zur Verfügung.

Nachdem die Ergebnisse in den USA gezeigt hatten, dass die transplantierten Muskelstammzellen die Inkontinenz bei den Ratten beheben konnte und die Sicherheit des Zellproduktes bestmöglich bestätigt werden konnte, steht nun der klinischen Studie nichts mehr im Weg: 21 betroffene Jungen im Alter zwischen drei und siebzehn Jahren sollen demnächst an den Universitätskliniken Ulm und Regensburg behandelt werden.

Stefanie Grosswendt
Foto: © BSIO

female independence award für Stefanie Grosswendt

Dr. Stefanie Grosswendt ist Nachwuchsgruppenleiterin im Fokusbereich Single-Cell-Ansätze für die personalisierte Medizin des BIH, des MDC sowie der Charité. Sie entwickelt neue Methoden, mit denen sich das Zusammenspiel einzelner Zellen während der Entwicklung untersuchen lässt. Anwenden kann sie ihre Forschungsansätze beim Neuroblastom, einer Krebserkrankung des frühen Kindesalters, bei der während der Entwicklung einzelne Zellen entarten. Für ihre Forschung erhielt die Wissenschaftlerin den Female Independence Award der Berlin School of Integrative Oncology (BSIO).

Der Fokus von Stefanie Grosswendts Forschung liegt auf der Embryonalentwicklung und dem Neuroblastom, einer der häufigsten Krebserkrankungen im frühen Kindesalter. „Das Neuroblastom entsteht schon im Mutterleib“, sagt Grosswendt. „In diesen frühkindlichen Tumoren befinden sich Zellen in verschiedenen Entwicklungsstadien. Die von uns entwickelten Single-Cell-Ansätze sollen dabei helfen, diese Krebserkrankung genauer zu verstehen und damit langfristig die Diagnose- und Behandlungsmöglichkeiten zu verbessern.“

Grosswendt gehört zu den vier Nachwuchsgruppenleiter*innen, die das BIH, das MDC und die Charité international für den gemeinsamen Fokusbereich „Single-Cell-Ansätze für die personalisierte Medizin“ rekrutiert haben. Die Gruppen sind am MDC in Mitte, und somit am Berliner Institut für Medizinische Systembiologie (BIMSB), angesiedelt.

Einzelzelltechnologien ermöglichen es, im großen Maßstab nachzuvollziehen, welche Gene in einzelnen Zellen gerade aktiv sind – zum Beispiel in einem Tumor. „Für die Mobilität und Plastizität der Neuroblastom-Zellen sind zelluläre Programme verantwortlich, die denen ähneln, die während der Embryonalentwicklung aktiv sind. Um sie zu entdecken, nutzen wir die Single-Cell-Technologien“, berichtet Grosswendt. „Außerdem entwickeln wir neue Methoden, mit denen wir untersuchen können, wie Zellen miteinander interagieren und kommunizieren.“ Vor allem will sie die Möglichkeiten der Einzelzellanalyse für klinische Fragestellungen nutzbar machen. Sie und ihr Team kooperieren daher eng mit Professorin Angelika Eggert, Direktorin der Klinik für Pädiatrie mit Schwerpunkt Onkologie und Hämatologie der Charité.

Stefanie Grosswendt hat am MDC im Bereich der Systembiologie 2015 bei Professor Nikolaus Rajewsky promoviert. Anschließend ging sie in die USA nach Boston, wo sie in der Gruppe von Professor Alexander Meissner die Entwicklung von Mausembryonen mittels Single-Cell-Technologien untersuchte. Ihm folgte sie 2018 ans Max-Planck-Institut für Molekulare Genetik in Berlin. Seit Dezember 2020 ist sie Leiterin der BIH-Nachwuchsgruppe „Vom Zellstatus zur Funktion“.

Der BSIO Female Independence Award ist mit 25.000 Euro dotiert und soll als Labor-Starthilfe dienen; Preis und Preisgeld teilt sich Stefanie Grosswendt mit Dr. Soulafa Mamlouk vom Institut für Pathologie der Charité.

mTOMADY

Ein digitales Gesundheitsportemonnaie für Menschen ohne Krankenversicherung

Firmenporträt mTOMADY

von Stefanie Seltmann

Mehr als eine Milliarde Menschen in Ländern mit niedrigem und mittlerem Einkommen haben keinen Zugang zur medizinischen Grundversorgung, weil die ausreichende finanzielle Absicherung fehlt. Ärzte der Charité – Universitätsmedizin Berlin haben eine digitale Lösung für dieses Problem entwickelt und Ende 2020 das Unternehmen mTOMADY gegründet: Über die Mobiltelefon-Infrastruktur kann Geld sicher und effizient für medizinische Behandlungen eingezahlt, angespart und abgerufen werden. Das BIH Charité Digital Clinician Scientist Programm sowie das Digital Health Accelerator Programm des Berlin Institute of Health (BIH) in der Charité hat sie dabei unterstützt. Die Gesundheitsplattform mTOMADY ist zunächst in Madagaskar im Einsatz, weitere afrikanische Länder sollen folgen.

Die Assistenzärzte der Klinik für Neurologie mit Experimenteller Neurologie der Charité, Dr. Julius Emmrich und Dr. Samuel Knauß, arbeiten seit Jahren ehrenamtlich in Entwicklungsländern, wie zum Beispiel in Madagaskar. Dort leben 93 Prozent der Einwohner*innen unter der Armutsgrenze und weniger als fünf Prozent der Madagassen verfügen über ein Bankkonto. Diese Zustände sorgten dafür, dass die beiden Ärzte bei ihrer ehrenamtlichen Arbeit mit zahlreichen dramatischen Situationen konfrontiert wurden. „Beispielsweise starb ein Patient, der sich bei einem Autounfall schwer verletzte, weil seine Familie die Kosten für die Behandlung nicht im Voraus bezahlen konnte“, erzählt Samuel Knauß. Und Julius Emmrich erinnert sich: „Wir haben oft erlebt, dass sich schwer kranke oder verletzte Patient*innen aus Angst vor den hohen Behandlungskosten weigerten, ins Krankenhaus zu gehen.“

Mobile Money für die Gesundheit

In den letzten Jahren hat Madagaskar die Digitalisierung stark vorangetrieben und große Teile des Landes mit Mobilfunkmasten ausgestattet, wodurch nun sogar die ländlichsten Gebiete über mobiles Netz verfügen. Mittlerweile hat mehr als die Hälfte der Einwohner*innen Madagaskars ein Mobiltelefon und jährlich kommen über eine Million neue Nutzer*innen hinzu. Diese zunehmende Digitalisierung im Land sorgte für einen enormen Aufschwung von Mobile Money. Hierbei wird Geld sicher, ähnlich wie eine SMS, über das Mobilfunknetz vom Sender zum Empfänger transferiert.

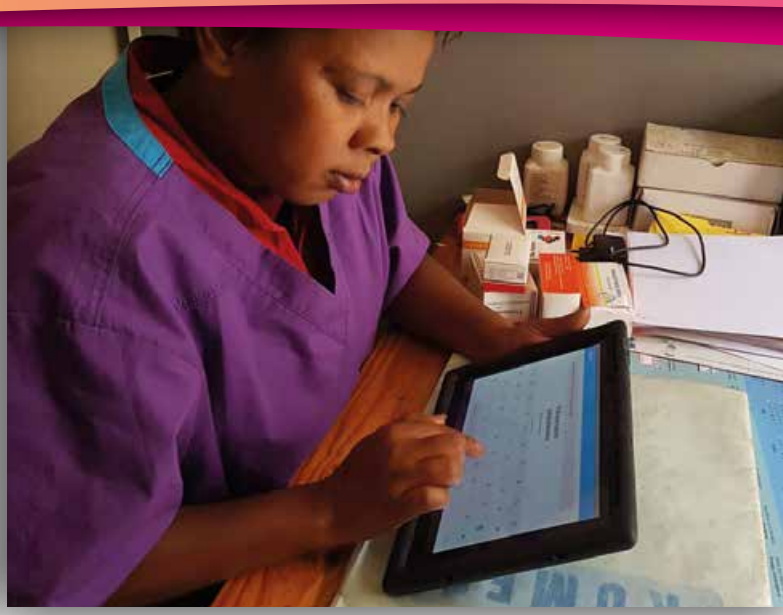
Die herausfordernden Zustände im Gesundheitswesen, mit denen Emmrich und Knauß bei ihrer Arbeit konfrontiert wurden, sowie die neuen digitalen Entwicklungen in Madagaskar, brachten die beiden auf die Idee, die Gesundheitsplattform mTOMADY zu gründen, was auf madagassisch „gesund und stark“ bedeutet. mTOMADY funktioniert wie ein digitales Gesundheitsportemonnaie. Um es zu nutzen, benötigt man lediglich einen Zugang zur örtlichen Mobiltelefon-Infrastruktur. Hierfür reicht eine einfache SIM-Karte. Bargeld kann man an sogenannten Mobile-Cashpoints, die man in Madagaskar an fast jeder Straßenecke findet, in Mobile Money umwandeln. Die Nutzer*innen können effizient und sicher Geld über das Mobile Money-System auf ein Konto transferieren, das für medizinische Behandlungen vorgesehen ist.

Gesunde und glückliche Kinder in Ejeda, einer Stadt und Gemeinde im Südwesten Madagaskars (Fotos: © Ärzte für Madagaskar e.V.).

Eine Besonderheit von mTOMADY ist, dass man das Geld, das man aufs Gesundheitskonto geladen hat, ausschließlich für Gesundheitsdienstleistungen nutzen kann. So ist sichergestellt, dass die Nutzer*innen ein finanzielles Polster für den Notfall ansparen, das in prekären Situationen aufgrund von Krankheiten oder Unfällen Leben retten kann. Zusätzlich ist es Nutzer*innen möglich, eine Krankenversicherung über die Plattform abzuschließen. Ebenso können Spenden auf eine digitale Spendenplattform eingezahlt werden, von denen alle Mitglieder von mTOMADY profitieren.

BIH unterstützt Entwicklung und Ausgründung

Julius Emmrich und Samuel Knauß sind beide seit 2019 Teilnehmer des BIH Charité Digital Clinician Scientist-Programms, was ihnen ermöglichte, während ihrer Facharztweiterbildung zum Neurologen mTOMADY zu entwickeln. Der Digital Health Accelerator des BIH förderte das Projekt von Juli 2018 bis Dezember 2020 mit rund 1 Mio. Euro. Das Team erhielt Mentoring durch erfahrene Expert*innen in Themen wie Mobiltechnologie, Software- und Produktentwicklung, Versicherung sowie Hilfe, die Ausgründung vorzubereiten, die als mTOMADY gGmbH im Dezember 2020 erfolgte. „Wir freuen uns, dass mit unserer Unterstützung mit mTOMADY eine digitale Gesundheitslösung für strukturschwache Länder bereitsteht und wir so einen Beitrag für die gesellschaftliche Wertschöpfung vor Ort leisten“, sagt Thomas Gazlig, Leiter von BIH Charité Inno-



Gesundheitspersonal in Madagaskar: (links) eine Fortbildung, und (rechts) mTOMADY im Einsatz in einer Klinik (Fotos: © Ärzte für Madagaskar e.V.).

vation, dem gemeinsamenTechnologietranfer von Charité und BIH.

Im Januar 2020 war mTOMADY Preisträger beim Ideenwettbewerb „Neue Ideen für globale Gesundheit“ des Global Health Hub Germany. Aktuell arbeitet das Gesundheitsministerium von Madagaskar zusammen mit mTOMADY daran, das Gesundheitsportemonnaie mTOMADY zu einem integralen Bestandteil der Gesundheitsversorgung des Landes weiterzuentwickeln.

Zu Beginn schwangere Frauen unterstützt

Angefangen hat mTOMADY mit der Unterstützung von schwangeren Frauen im zentralen Hochland. „In Madagaskar herrscht eine hohe Mutter- und Säuglingssterblichkeit, und nur wenige Geburten werden professionell durchgeführt“, begründet Julius Emmrich diese Wahl. Samuel Knauß ergänzt: „Wenn man bei der Unterstützung schwangerer Frauen ansetzt, wird ein wichtiger Grundpfeiler für die Zukunft des Landes gesetzt.“ Mittlerweile steht mTOMADY auch anderen Patientengruppen zur Verfügung. Außerdem sind Emmrich und Knauß gerade dabei, mehrere Krankenkassen in Madagaskar in das System zu integrieren, die dafür sorgen sollen, dass das System noch effizienter und die Patient*innen noch umfangreicher unterstützt werden. Mittlerweile arbeiten im Team von mTOMADY Mitglieder aus neun verschiedenen Nationen. Bald wollen sie ihr System auch in Ghana und Uganda eingeführen.

Im BIH Podcast berichten Samuel Knauß und Julius Emmrich über ihre Erfahrungen in Madagaskar und die Gründung von mTOMADY: https://www.bihealth.org/de/aktuell/wie-bezahlt-man-den-arzt-in-madagaskar

Firmenportrait:

- **Gründung 2020 durch deutsch madagassisches Gründerteam**
- **25 Mitarbeiter*innen in Antananarivo und Berlin**
- **>200.000 registrierte Nutzer*innen**
- **> 45.000 Behandlungen**

www.mtomady.de

Wenn jemand das Projekt unterstützen möchte, freuen sich die Organisatoren sehr über Spenden an den Verein „Ärzte für Madagaskar e.V.“ https://www.aerzte-fuer-madagaskar.de/

Kontakt:

Dr. Samuel Knauß und Dr. Julius Emmrich
Klinik für Neurologie mit Experimenteller Neurologie
Charité – Universitätsmedizin Berlin
samuel.knauss@charite.de
julius.emmrich@charite.de

Samuel Knauß (Mitte) und Julius Emmrich (rechts) mit Gesundheitspersonal vor einem Gesundheitszentrum in Antananarivo, Madagaskar (Foto: © Ärzte für Madagaskar e.V.).

„komplexe fragestellungen kann man nur durch interdisziplinarität beantworten“

Interview mit Dana Westphal und Michael Seifert

Interdisziplinärer Juniorverbund forscht an neuen Therapieansätzen für Melanom-Hirnmetastasen

Dana Westphal und Michael Seifert leiten einen der sieben Juniorverbünde, die seit 2020 im Rahmen des e:Med Förderkonzeptes vom BMBF gefördert werden. Im Interview mit gesundhyte.de sprechen die Biologin und der Bioinformatiker über interdisziplinäre Zusammenarbeit sowie Ziele und Erfolge ihres Juniorverbundes.

gesundhyte.de: *Der schwarze Hautkrebs, das Melanom, gilt als die gefährlichste Form aller Hautkrebserkrankungen. Was macht ihn so gefährlich?*

Dr. Dana Westphal: Jedes Jahr erkranken allein in Deutschland etwa 21 Menschen pro 100.000 Einwohnern an einem Melanom. Leider steigt die Häufigkeit stetig an, vor allem bei jüngeren Menschen. Melanome sind so gefährlich, weil sie im Vergleich zum hellen Hautkrebs sehr viel häufiger in andere Körperregionen metastasieren. Hinzu kommt, dass etwa die Hälfte der Patientinnen und Patienten mit einem metastasierten Melanom Hirnmetastasen entwickeln, wodurch sich ihre Überlebenschancen deutlich verschlechtern. Frühzeitige und regelmäßige Vorsorgeuntersuchungen sind daher besonders wichtig, um bösartige Veränderungen rechtzeitig zu erkennen und diese operativ zu entfernen bevor es zur Metastasierung kommt.

Sie forschen seit 6 Jahren am Melanom. Was möchten Sie mit dem e:Med-Junioverbund MelBrainSys herausfinden?

Dr. Dana Westphal: Glücklicherweise gibt es seit 2011 effektive Therapien für das metastasierte Melanom. Jedoch stellen Hirnmetastasen, ganz besonders wenn diese Beschwerden verursachen, immer noch eine therapeutische Herausforderung dar. Mit unserem Projekt möchten wir herausfinden, warum sich Hirnmetastasen von den anderen Organmetastasen unterscheiden und Ansatzstellen für neue Hirn-spezifische Therapien identifizieren.

Dr. Michael Seifert: Unsere Forschung baut auf unserem Vorläuferprojekt auf. In diesem Projekt konnten wir mithilfe von Patientinnen und Patienten aus der Region einen Datensatz von Methylom- und Transkriptom-Daten aufbauen. Dieser Datensatz enthält für jede Patientin und für jeden Patienten jeweils ein Metastasepaar bestehend aus einer Hirn- und einer zugehörigen Nichthirn-Metastase. Diese patientenspezifischen Metastasepaare ermöglichen es uns, Unterschiede zwischen Hirn- und Nichthirn-Metastasen viel genauer zu bestimmen. Bisher wurden dazu hauptsächlich Hirn- und Nichthirn-Metastasen von unterschiedlichen Patienten untersucht. Das macht

Dana Westphal und Michael Seifert koordinieren gemeinsam den e:Med-Juniorverbund MelBrainSys zur Erforschung neuer Therapieansätze für Melanom-Hirnmetastasen (Quelle: Universitätsklinikum Dresden/I. Starke, D. Westphal und M. Seifert /privat).

unseren Datensatz so wertvoll für unsere Forschung und hilft uns hoffentlich, Kandidatengene und veränderte Signalwege für neue Therapieansatzstellen zu ermitteln.

gesundhyte.de: *Sind Sie in das erste Jahr Ihres Projektes gut gestartet oder gab es Anlaufschwierigkeiten?*

Dr. Dana Westphal: Wie viele andere Wissenschaftlerinnen und Wissenschaftler sind wir auch von der Corona-Krise betroffen. Unser Labor hat drei Monate schließen müssen. Aber wir haben die Zeit genutzt, das beschriebene Kollektiv von Patientenproben zu charakterisieren und für die Untersuchungen vorzubereiten.

gesundhyte.de: *Was haben sie im ersten Jahr bereits erreichen können?*

Dr. Michael Seifert: Wir können mittlerweile die Daten der individuellen Metastasepaare – bestehend aus Hirn- und Nicht-Hirnmetastasen - mit den von uns entwickelten Modellen schon detailliert analysieren. Es ist uns zudem gelungen, einen großen öffentlichen Datensatz so aufzubereiten, dass wir die Expressions- und Methylierungsdaten nutzen können, um ein genomweites, melanomspezifisches Geninteraktionsnetzwerk abzubilden. Durch dieses Netzwerk und mithilfe unserer Patientendaten können wir jetzt herausfinden, wie gewisse Veränderungen von Genen auf Signalwege oder Immunsignaturen wirken.

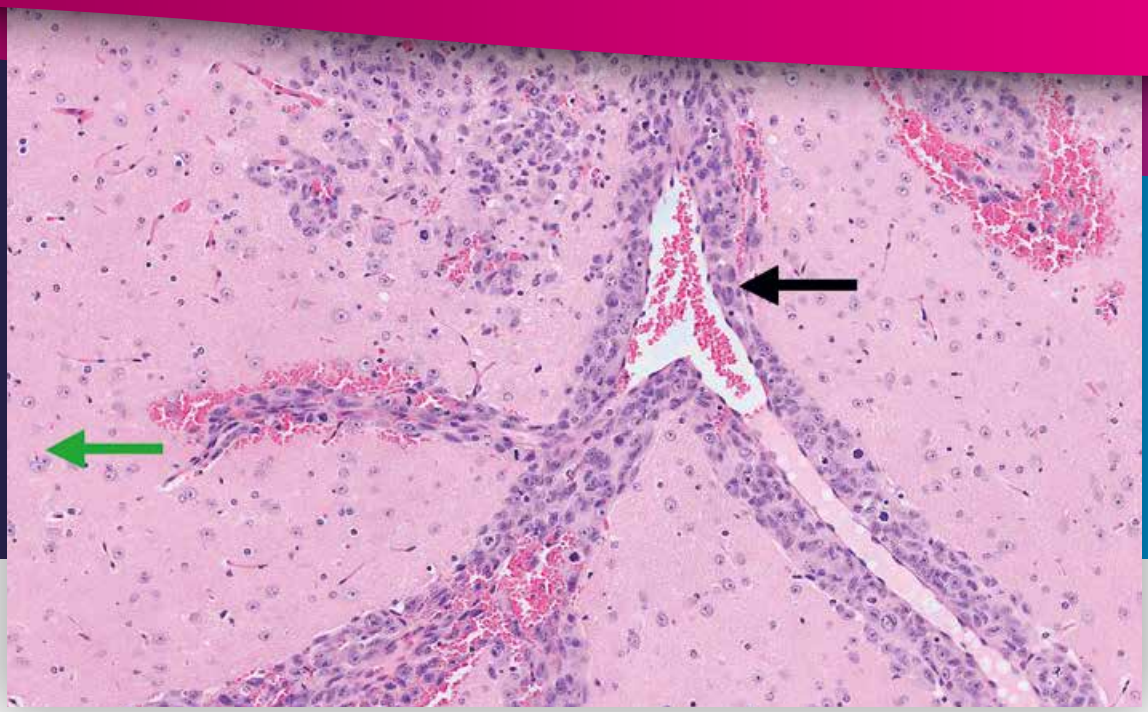
gesundhyte.de: *Welche Erkenntnisse erhoffen Sie sich von Ihrer Arbeit?*

Dr. Dana Westphal: Ziel ist es, zelluläre Mechanismen aufzudecken, die sich zwischen Hirnmetastasen und anderen Organmetastasen unterscheiden, denn hier könnte ein Angriffspunkt für neue Behandlungsmöglichkeiten liegen, die spezifisch gegen Hirnmetastasen wirken. Das wäre ein toller Erfolg.

Dr. Michael Seifert: Ein weiteres längerfristiges gemeinsames Ziel unserer Arbeit ist es, auf Basis der erhobenen molekularen Daten Biomarker zu identifizieren, um Medizinerinnen und Mediziner mit Hilfe eines Klassifikationssystems bei Therapieentscheidungen zu unterstützen.

gesundhyte.de: *Wie ist Ihr Verbundprojekt aufgebaut?*

Dr. Dana Westphal: Neben Dr. Michael Seifert und mir arbeiten auch Dr. Rebekka Wehner und Dr. Matthia Karreman an dem Projekt. Frau Wehner ist Immunologin am Institut für Immunologie hier in Dresden, Frau Karreman ist Neuroonkologin am DKFZ in Heidelberg. Unsere Zusammenarbeit läuft sehr gut, jeder bringt seine Expertise mit in das Projekt ein.



Hämatoxylin-Eosin-Färbung von Tumorgewebe: Melanomzellen (violett) wachsen entlang der Blutgefäße (schwarzer Pfeil) und dringen in das umliegende Hirngewebe (grüner Pfeil) ein (Quelle: Dr. med. Matthias Meinhardt, Institut für Pathologie, Universitätsklinikum Dresden).

Dr. Michael Seifert: Interdisziplinarität ist für einen Juniorverbund das zentrale Element. Ich als Bioinformatiker könnte nicht arbeiten, wenn ich nicht die molekularen Daten von den Projektpartnerinnen bekommen würde. Und meine drei Projektpartnerinnen könnten keine Signalwege oder Kandidatengene untersuchen, wenn sie die Analysen aus meinem Teilprojekt nicht zur Verfügung hätten. Wir bearbeiten eine sehr komplexe Fragestellung, die man nur durch Interdisziplinarität beantworten kann, weil keiner von uns Experte in allen Disziplinen sein kann.

gesundhyte.de: *Hat es gedauert, bis diese Teamarbeit etabliert war? Oder hat das von Anfang an funktioniert?*

Dr. Michael Seifert: Frau Westphal und ich haben schon im Vorgängerprojekt gut zusammengearbeitet. Dabei haben wir festgestellt, dass wir die Patientinnen und Patienten noch wesentlich individueller betrachten und dafür unsere Untersuchungen systematisch erweitern müssen. Für die Suche nach neuen Therapieansätzen sind deswegen weitere Partner notwendig gewesen. Schon in der Antragsphase haben wir gemerkt, dass die Chemie zwischen uns stimmt. Ich habe aber auch gelernt, dass es viele Dinge gibt, die einem selbst so klar sind, dass man sie gar nicht mehr erläutert. Daraus können auch Missverständnisse entstehen.

Dr. Dana Westphal: Es ist wichtig, viel miteinander zu reden. Wir treffen uns daher regelmäßig, um von Anfang an die Projektausrichtung zu diskutieren und in die richtigen Bahnen zu lenken.

gesundhyte.de: *Warum sind Sie eigentlich Naturwissenschaftler geworden?*

Dr. Michael Seifert: Schon in der Schule hat mich Informatik und Biologie interessiert, auch medizinische Themen fand ich immer sehr interessant. Als ich mein Abitur gerade hatte, erlebte die Bioinformatik in Deutschland durch die Förderung von Bioinformatikzentren einen großen Schub, der auch ganz neue berufliche Perspektiven eröffnete. Deswegen habe ich mich für das Studium entschieden. Mir ist es wichtig, mit meiner Arbeit einen Beitrag zur Entwicklung von besseren Behandlungsmöglichkeiten leisten zu können.

Dr. Dana Westphal: Ich bin Wissenschaftlerin geworden, weil meine Biologielehrerin das Fach so gut rübergebracht hat. Deswegen habe ich Biologie studiert. Dann hat es mich auf die andere Seite der Erde verschlagen und ich habe in Neuseeland promoviert. Nach weiteren fünf Jahren in Australien bin ich nach Dresden zurückgekommen. Was mir an meiner aktuellen Stelle hier am Universitätsklinikum Dresden so gefällt, ist die Patienten-nahe Forschung und das Gefühl, mit meiner Arbeit etwas zum Wohle der Patientinnen und Patienten beitragen zu können.

Das Gespräch führte Marco Leuer.

MelBrainSys:

Erforschung von therapeutischen Angriffspunkten von Melanom-Hirnmetastasen

Das maligne Melanom, der sogenannte schwarze Hautkrebs, ist eine bösartige Krebserkrankung, deren Auftreten in den letzten Jahren stark zugenommen hat. Das Melanom bildet besonders häufig bereits in frühen Krankheitsstadien Metastasen aus. Dabei entwickelt fast die Hälfte der Patienten im Verlauf der Erkrankung auch Metastasen im Gehirn. Diese lassen sich mit bisher existierenden Therapien deutlich schlechter behandeln als Metastasen in anderen Organen. Welche molekularen Faktoren dabei eine wichtige Rolle spielen und wie sich diese zukünftig für wirksamere Behandlungsansätze nutzen lassen könnten, ist bisher kaum erforscht.

Neben bereits bekannten spezifischen Genveränderungen von Melanom-Hirnmetastasen trägt wahrscheinlich auch eine epigenetische Reprogrammierung zu einer verminderten Wirkung oder Resistenz gegen bisher existierende Behandlungsstrategien bei. Die Erforschung des komplexen Zusammenspiels dieser molekularen Ebenen ist sehr schwierig und wurde in der Vergangenheit noch dadurch erschwert, dass bisher hauptsächlich Hirn- und Organmetastasen (z. B. Lunge, Leber, Niere) von verschiedenen Melanompatienten für eine Untersuchung zur Verfügung standen.

Genau an dieser Stelle setzt der vom BMBF geförderte e:Med-Juniorverbund MelBrainSys an. Im Rahmen des Projektes kann auf einen besonderen Datensatz von Genexpressions- und Methylomprofilen aus Hirn- und Organmetastasen der jeweils gleichen Patienten zurückgegriffen werden, um tiefgreifendere Erkenntnisse über molekulare Veränderungen auf Basis von patientenspezifischen Analysen zu gewinnen. Dabei sollen mit Hilfe eines systemmedizinischen Ansatzes, welcher sowohl bioinformatische als auch laborexperimentelle Analysen beinhaltet, Kandidatengene und Signalwege identifiziert werden, die Hirn- und Organmetastasen unterscheiden. Aus diesen umfassenden Analysen sollen schlussendlich potenzielle Ansatzpunkte für zukünftige hirnspezifische Therapieansätze ermittelt werden.

Die interdisziplinäre Zusammenarbeit der Bereiche Dermatoonkologie, Immunologie, Neuroonkologie und Bioinformatik ist für die erfolgreiche Realisierung des Juniorverbundes von entscheidender Bedeutung. Nur so kann das notwendige Wechselspiel von Laborexperimenten und computerbasierten Analysen aussagekräftig verknüpft werden, um eine gezielte Untersuchung von Kandidatengenen und Signalwegen in Zellkulturexperimenten und *in-vivo* Modellen zu ermöglichen. Die im Rahmen des Juniorverbundes erzielten Ergebnisse können wichtige Grundlagen für zukünftige klinische Studien zur besseren Behandlung von Melanom-Hirnmetastasen legen.

Kontakt:

Dana Westphal, PhD
Juniorgruppenleiter AG Hirnmetastasen
Klinik und Poliklinik für Dermatologie
Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus Dresden
TU Dresden
dana.westphal@uniklinikum-dresden.de

Dr. Michael Seifert
Gruppenleiter Bioinformatik Core Unit
Institut für Medizinische Informatik und Biometrie (IMB)
Medizinische Fakultät Carl Gustav Carus
TU Dresden
michael.seifert@tu-dresden.de

INCOME – integrative, kollaborative modellierung in der systemmedizin

Per Hackathon in eine Kultur des *Data and Model Sharing*

von Nina Fischer, Wolfgang Müller, Dagmar Waltemath, Olaf Wolkenhauer und Jan Hasenauer

Die Systemmedizin ist ein interdisziplinärer Ansatz, bei dem Ärzt*innen und klinische Forscher*innen mit Expert*innen aus den Bereichen Biologie, Biostatistik, Informatik und Mathematik zusammenarbeiten, um Diagnostik, Prävention und Behandlung von Krankheiten zu verbessern. Dafür müssen großskalige, heterogene Datensätze experimentell erhoben und mittels ganzheitlicher Modelle zusammengeführt werden. In der Realität beschränken sich jedoch die meisten Projekte auf einzelne zelluläre Prozesse und entwickeln hierfür maßgeschneiderte Modelle. Im BMBF-geförderten e:Med Projekt INCOME haben wir mit gezielten Vernetzungsaktivitäten die Zusammenarbeit und den Austausch zwischen Forschungsgruppen gefördert. Forscher*innen und Entwickler*innen schufen gemeinsam in zahlreichen INCOME-Meetings eine Kultur des „Data and Model Sharing“ und entwickelten so Wege zur besseren technischen Verknüpfung und Wiederverwendbarkeit bestehender Simulationsmodelle. Die Meetings boten damit eine Plattform zur Stärkung der Community.

Mathematische Modelle werden in der Systemmedizin verwendet, um krankheitsbezogene biochemische Prozesse für Diagnostik, Prognostik und therapeutische Entscheidungen zu untersuchen. Diese Modelle können in offenen Standards wie der Systems Biology Markup Language (SBML) formuliert und in Modelldatenbanken wie BioModels hinterlegt werden. Standards wie die Simulation Experiment Description Language (SED-ML) zielen auf eine eindeutige Verknüpfung von Daten und Modellen sowie die Beschreibung verschiedener Experimentalbedingungen ab. Mit der Entwicklung und Etablierung der Standards und Datenbanken haben sich die Verfügbarkeit und Wiederverwendbarkeit von Modellen und die Reproduzierbarkeit von Forschungsergebnissen massiv verbessert. Dennoch haben viele Forschungsgruppen und auch weit verbreitete Software Tools diese Standards noch nicht adaptiert und verwenden stattdessen eigene wenig standardisierte Formate, die unter anderem oft keine rigorose Annotation gestatten. Dementsprechend sind gemäß aktueller Erhebungen weiterhin nur ca. 50% der veröffentlichten Modellierungsstudien reproduzierbar (Tiwari *et al.*, 2021).

Neben Modellen selbst sind auch die zur Parametrisierung und Validierung verwendeten Datensätze oft schwer zugänglich. Sie werden nicht veröffentlicht, nicht ausreichend annotiert oder nicht eindeutig an das Modell geknüpft. Die genannten Faktoren stehen der Wiederverwendung und Erweiterung bestehender Modelle und Datensätze im Wege und kompromittieren somit den Erfolg von Forschungsprojekten.

Die stark begrenzte Wiederverwendbarkeit und der Zeitaufwand für die Entwicklung qualitativ hochwertiger Modelle sind zwei Gründe dafür, dass viele mechanistische Modelle nur einzelne Signalwege abbilden und Wechselwirkungen mit anderen Signalwegen oft ignorieren. Es ist an der Zeit, die

Abbildung 1: Impressionen aus zwei INCOME Veranstaltungen.
INCOME 2019 in Berlin und INCOME 2021 virtuell. (Fotos: Jan Hasenauer).

bestehenden Infrastrukturen zu nutzen, Forschende mit den vorhandenen Ressourcen vertraut zu machen und Software-Entwickler*innen auf verfügbare Bibliotheken und Werkzeuge für die standard-konforme Modellierung explizit aufmerksam zu machen.

Wichtigste Ziele des INCOME Projektes

Ziele des Verbundprojekts INCOME waren

1. **die Vernetzung von Forschungsgruppen zu verbessern und die Verwendung von Standards voranzutreiben,**
2. **eine integrierte Datenbank für Modelle und Daten zu etablieren, und**
3. **die kollaborative Entwicklung großskaliger Modelle zu vereinfachen.**

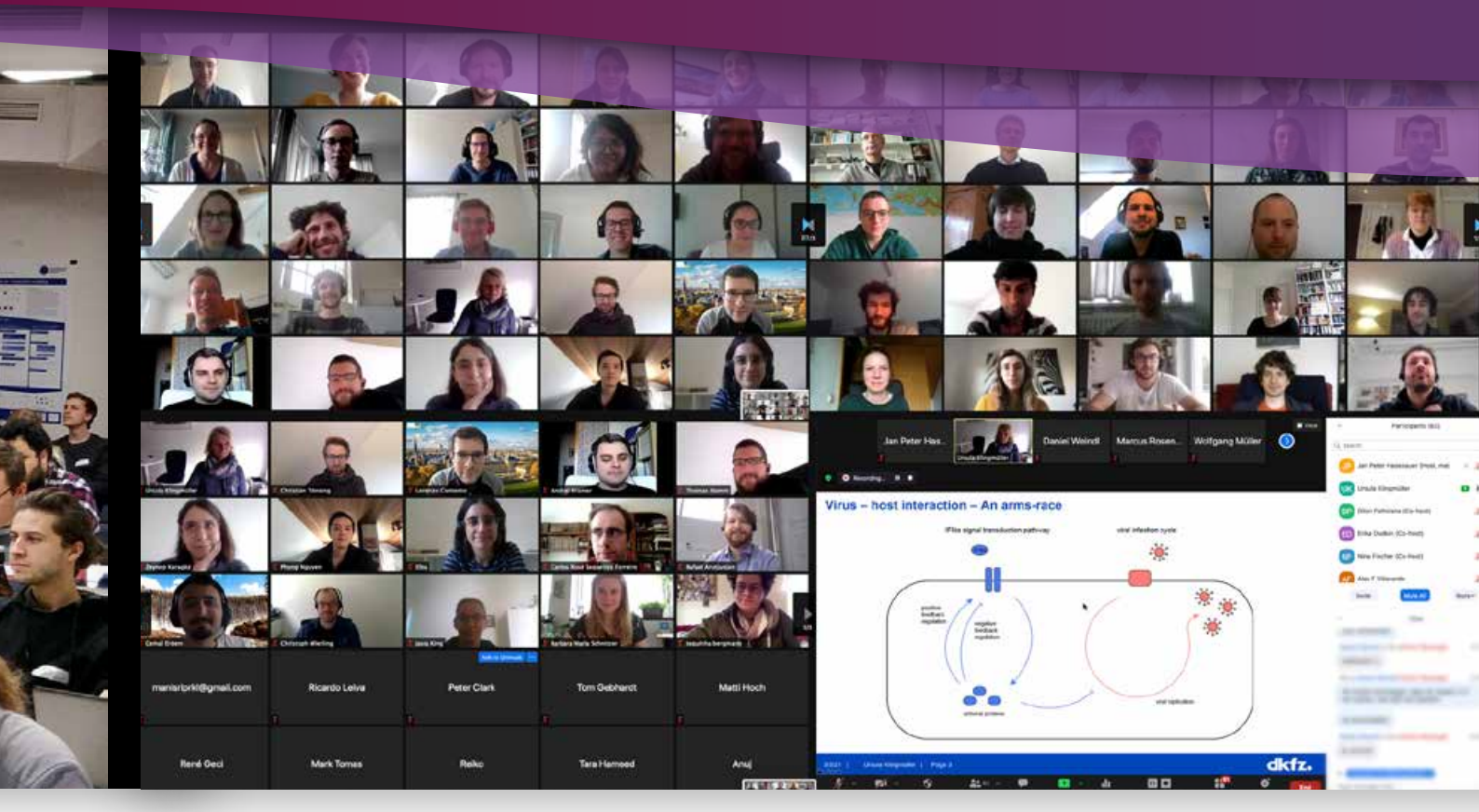
An der Umsetzung dieser Ziele wurde an mehreren Standorten in Deutschland gearbeitet. INCOME (2017 – 2021) wurde von Jan Hasenauer (Helmholtz Zentrum München und Uni Bonn) koordiniert und mit Wolfgang Müller (HITS Heidelberg), Olaf Wolkenhauer (Uni Rostock) und Dagmar Waltemath (Universitätsmedizin Greifswald, vormals Uni Rostock) bearbeitet.

Training, Networking und Community Building

INCOME hat insgesamt fünf Veranstaltungen durchgeführt: drei Konferenzen mit integrierten Hackathons, und zwei separate Hackathons (siehe Abbildung 1).

Die Konferenzen informierten über aktuelle Ergebnisse im Feld der Modellbildung, der Methodenentwicklung und der Anwendung. Eine Vielzahl von Vorträgen hat zudem der nächsten Generation von Systembiologen*innen und Systemmediziner*innen Standards nähergebracht und deren Nutzen aufgezeigt. Podiumsdiskussionen sowie Gruppen- und Einzelgespräche wurden dazu genutzt, offene Probleme anzusprechen und Lösungsansätze zu erarbeiten. Die Konferenzen hatten einen hohen Vernetzungscharakter und ermöglichten somit Nachwuchswissenschaftler*innen einen Austausch mit erfahrenen Wissenschaftler*innen der Community.

Die Hackathons boten ergänzend zu den Konferenzen Raum und Zeit für das Arbeiten an gemeinsamen Projekten. Um die Verwendung von Standards deutlich zu vereinfachen und sowohl Reproduzierbarkeit als auch Wiederverwendbarkeit zu verbessern, wurde Support von Standards in vielen verfügbaren Software Tools angeboten (z.B. Copasi, Data2Dynamics, Dmod und pyPESTO). Darüber hinaus haben Entwickler*innen sich Zeit genommen, neue Software vorzustellen und Nutzer*innen bei Problemen zu beraten. Im Gegenzug erhielten Sie direktes Feedback zu eigenen Tools und konnten die Weiterentwicklung an den Bedürfnissen der User ausrichten.



Die Veranstaltungen lockten Teilnehmer aus über 20 Ländern an. Diese Internationalität und die ebenfalls vorhandene Interdisziplinarität haben maßgeblich zur Entwicklung einer starken Community an der Schnittstelle von Modellierung, Methodenentwicklung und Standardisierung beigetragen. Die kontinuierlich steigende Teilnehmerzahl der Veranstaltungen spiegelt das wachsende Interesse hieran und allgemein an einer intensiven Zusammenarbeit wider (siehe Abbildung 2). Viele Aspekte der INCOME Veranstaltungen fanden direkt Eingang in die Entwicklungen des globalen Standardisierungsnetzwerks COMBINE (www.co.mbine.org).

Lücken schließen

INCOME hat neben der Implementierung von Standards in existierende Workflows auch offene Fragen identifiziert und adressiert. So wurde beispielsweise bereits bei der ersten Veranstaltung im Oktober 2018 erkannt, dass es keine passende Lösung gibt, um Parametrisierungsprobleme zu formulieren. Um diese Lücke zu schließen, wurde während der Hackathons, aber auch in der Zeit zwischen den Treffen, das Parameter Estimation table format (PEtab) entwickelt (siehe Abbildung 3). PEtab erlaubt eine Standardisierung auf einer neuen Ebene, und verbessert damit die Wiederverwendbarkeit von Daten und Modellen für neue Studien, z.B. die Entwicklung ganzheitlicher Modelle.

An der Entwicklung von PEtab hat eine Vielzahl von Forschungsgruppen mitgearbeitet und so ist der neue Standard bereits nach kürzester Zeit in acht Software Tools für Tausende von Nutzern zugänglich (Schmiester *et al.*, 2021). Dies wäre ohne die Ausbildung einer starken, thematisch fokussierten Community und die gemeinsamen Hackathons nicht möglich gewesen.

Ressourcen schaffen

INCOME hat dazu beigetragen, Community Ressourcen zu schaffen, um die Systembiologie und die Systemmedizin als Forschungsgebiete voranzubringen. So wurde ein zuvor etabliertes Repositorium für Datensätze und Modelle (Hass *et al.*, 2019) in PEtab überführt und kontinuierlich erweitert. Dies erleichterte nicht nur die Wiederverwendung, sondern ermöglicht auch ein realistisches Benchmarking von Methoden. Eine Untersuchung zur Skalierbarkeit und Robustheit von existierenden numerischen Lösungsverfahren wurde bereits abgeschlossen (Städter *et al.*, 2021), und mehrere weitere Studien zu Optimierungs- und Stichprobengenerierungsverfahren sollten

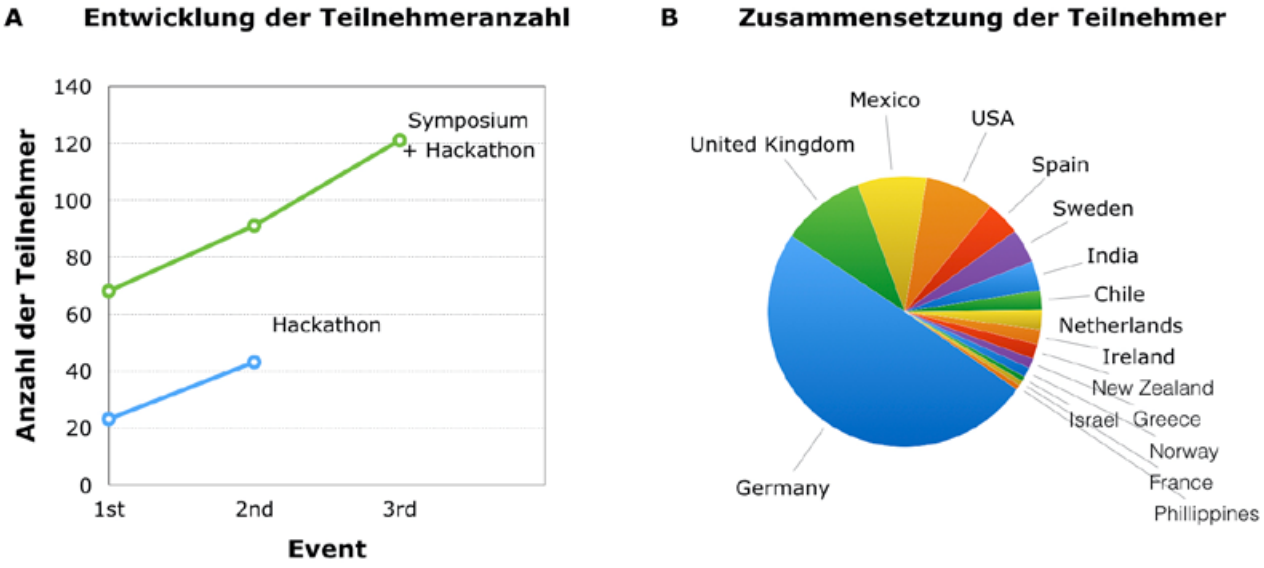


Abbildung 2: A Statistik aus Anmeldedaten (Anzahl der Teilnehmer) der 5 Veranstaltungen von 2018-2021, B Statistik aus Anmeldedaten der virtuellen Veranstaltung 2021. (Quelle: Jan Hasenauer).

zeitnah verfügbar sein. Darüber hinaus hat das INCOME Konsortium als Teil einer großen, weltumspannenden Community unter der Leitung von Marek Ostaszewski (Luxembourg Centre for Systems Biomedicine) an einem Modell für SARS-CoV-2 Infektionen gearbeitet (Ostaszewski *et al.*, 2020). Die daraus resultierende COVID-19 Disease Map ist wohl die mit Abstand umfangreichste Beschreibung der molekularen und zellulären Prozesse während der Infektion.

Alle geschaffenen Ressourcen sind frei verfügbar und können auf Grund ihres hohen Grades an Standardisierung einfach in weiteren Projekten genutzt werden. Dies ermöglicht einen Wissenstransfer und vereinfacht eine ganzheitliche Betrachtung von Krankheiten.

Aber wo geht es jetzt hin?

INCOME hat allen Konsortiumsmitgliedern aufgezeigt, wieviel mehr man als Teil einer starken Community erreichen kann, und wie wichtig im Umkehrschluss Community Building ist. Die Trainings während der Konferenzen, der Austausch und die gemeinsame Arbeit während der Hackathons, sowie die zahlreichen Diskussionen haben dazu beigetragen, Modell-, Software- und Standardentwicklung näher zueinander zu bringen. Dies war ein Prozess. Einzelne Treffen wären nicht ausreichend gewesen und es ergab sich ein substanzieller Mehrwert durch den wiederholten Austausch, wiederkehrende Teilnehmer, die wiederum in ihren Forscherkreisen für die Meetings geworben hatten. Wir gehen fest davon aus, dass das Wissen, welches durch die INCOME Veranstaltungen in einzelne Forschungseinrichtungen getragen worden ist und wird, die Wiederverwendbarkeit von Modellen deutlich verbessert hat und weiterhin steigern wird. Künftig sollen noch viel mehr Forscher*innen erreicht werden! Wir planen daher die Veranstaltungsreihe auch über das Projektende hinaus fortzusetzen (und arbeiten an der Realisierbarkeit).

Wir möchten uns bei allen Teilnehmer*innen der Konferenzen und Hackathons bedanken, die den Erfolg von INCOME möglich gemacht haben.

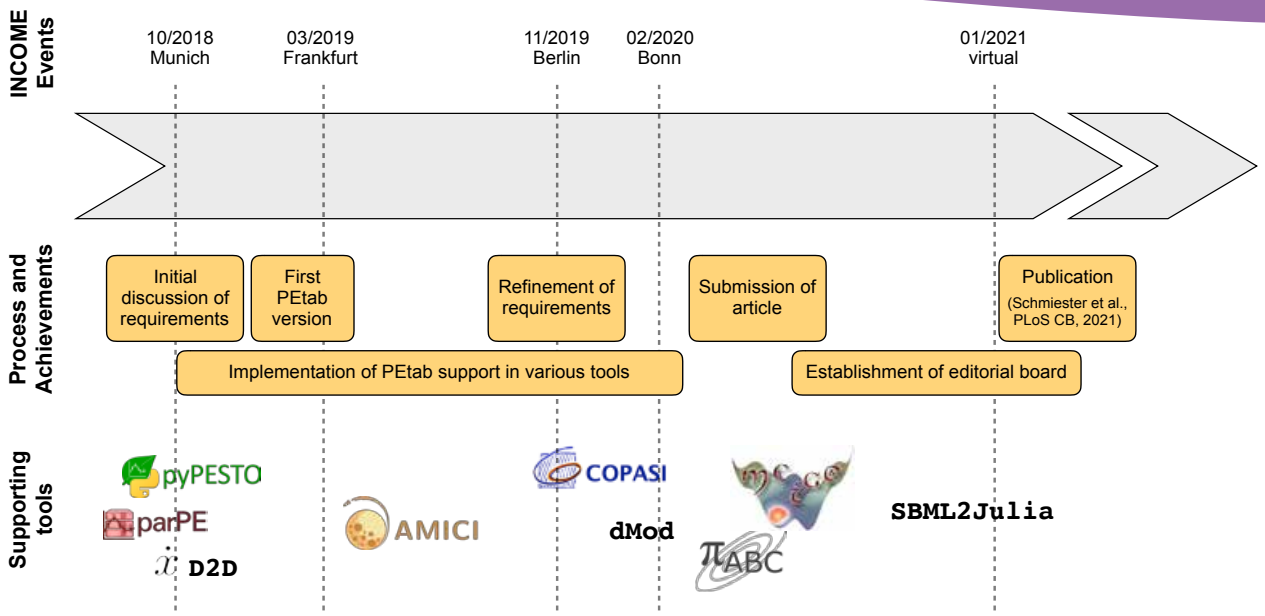


Abbildung 3: Entwicklung des PEtab Standards im Rahmen der INCOME Konferenzen (Quelle: Jan Hasenauer).

Referenzen:

Tiwari *et al.* (2021). Reproducibility in systems biology modelling. Molecular Systems Biology, 17:e9982.

Schmiester *et al.* (2021). PEtab - Interoperable specification of parameter estimation problems in systems biology. PLOS Computational Biology, 17(1):e1008646.

Hass *et al.* (2019). Benchmark problems for dynamic modeling of intracellular processes. Bioinformatics, 35(17):3073-3082.

Städter *et al.* (2021). Benchmarking of numerical integration methods for ODE models of biological systems. Scientific Reports, 11:2696.

Ostaszewski *et al.* (2020). COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. Scientific Data, 1:136.

Kontakt:

Prof. Dr. Jan Hasenauer
Life and Medica Sciences (LIMES) Institute
University of Bonn
jan.hasenauer@uni-bonn.de

https://www.mathematics-and-life-sciences.uni-bonn.de/de

E:MED VERNETZUNGSFONDS

Die **e:Med Vernetzungsfonds-Projekte** zeigen, wie in Hackathons, Workshops und Summer Schools wertvolle Ressourcen wie Software-Tools, Repositorien und Standards erarbeitet werden können. Ein Erfolg dieser Vorhaben sind konkrete translationale Anwendungen in klinischen Studien, wie zu Magenkrebs, Brustkrebs und entzündlichen Darmerkrankungen. Sechs Verbünde vernetzen jeweils gezielt mehrere e:Med Vorhaben. Die Wissenschaftler*innen bringen ihre Expertise zu relevanten Querschnittsthemen ein und erzeugen so durch ihre interdisziplinäre Zusammenarbeit in den Bereichen Klinik, Grundlagen, Bioinformatik und Modellierung zusätzlich entscheidenden Mehrwert. Diese Initiative hat das BMBF im Rahmen seiner Förderinitiative e:Med mit Mitteln aus einem Vernetzungsfonds ermöglicht.

www.sys-med.de

e:Med juniorverbund LeukoSyStem

Interview mit Simon Haas,
Einzelzell-Multi-Omics Leukämieforscher, e:Med Verbundleiter, BIH, Charité, MDC Berlin und DKFZ, HI-STEM Heidelberg

Simon Haas forscht wie er joggt, immer neue Wege erkundend. Seine wissenschaftliche Neugier hat das Lehrbuchwissen über Blutbildung umgeworfen, hierfür hat er weltweit als einer der ersten neuartige Einzelzell-Technologien eingesetzt. Wo eine Methode fehlt, entwickelt er sie, jetzt Einzelzell-Multi-Omics, um räumlich aufgelöst Leukämiestammzellen zu enttarnen – und therapeutisch anzugehen.

gesundhyte.de: *Sie leiten den* ***Juniorverbund LeukoSyStem*** *in e:Med, der großen Fördermaßnahme des BMBF zur Systemmedizin. Dieses Projekt ist ja nicht das erste, das Sie leiten, was ist denn das Ziel von LeukoSyStem?*

Dr. Simon Haas: Unser Hauptziel in diesem Verbund ist, Leukämien (Blutkrebs) besser zu verstehen, indem wir die Stammzellen in den Fokus nehmen, um **mit Einzelzellansätzen neue Therapiemöglichkeiten zu entwickeln.** Leukämien werden häufig immer noch mit der klassischen Chemotherapie behandelt. Diese kann man sich vorstellen, wie einen Bulldozer, der zwar effektiv Krebszellen bekämpft, aber auch gesunde Zellen schädigt. Dies führt zu starken Nebenwirkungen. Darüber hinaus bleiben bei der Chemotherapie häufig Krebszellen zurück und dies sind oft sogenannte Krebsstammzellen. Sie sind nur ein Teil des Tumors, aber diese wenigen Zellen können den kompletten Tumor neu bilden. Genau diese Krebsstammzellen sind häufig das Reservoir für einen Rückfall nach einer erfolgreichen Therapie.

Wir zielen also darauf ab, therapeutische Ansätze zu entwickeln, die speziell Leukämie- und besonders Leukämiestammzellen abtöten, aber gesunde Zellen unberührt lassen. Diese Therapien sollen weniger Nebenwirkungen haben, und gleichzeitig den Rückfall verhindern, denn der ist die Haupttodesursache in den meisten Leukämien. Dafür entwickeln wir sogenannte Einzelzell-Technologien. Aus einzelnen Zellen erhalten wir eine Vielzahl von Informationen und können insbesondere verstehen, welche Zelle gesund ist und welche erkrankt. So können wir spezifisch Zielstrukturen auf den Krebszellen identifizieren, diese Zellen unschädlich machen und gesunde Zellen am Leben lassen. Das ist, im Großen und Ganzen, die Gesamtidee des Vorhabens.

> „Leukämien werden häufig immer noch mit der klassischen Chemotherapie behandelt. Diese kann man sich vorstellen, wie einen Bulldozer, der zwar effektiv Krebszellen bekämpft, aber auch gesunde Zellen schädigt."

Das klingt einleuchtend und sehr effektiv. Welche Methoden verwenden Sie dafür? Haben Sie diese selbst entwickelt?

Der Bereich der Einzelzell-Methodik ist eine relativ neue Forschungsrichtung, die erst in den letzten Jahren entwickelt wurde. Die Mitglieder von LeukoSyStem waren seit Anbeginn des Feldes mit dabei und haben viele Methoden selbst entwickelt. Sie decken das Methodenspektrum vom molekularen Experiment über Datenanalyse bis zur Medizin ab.

Ihre Publikationen vermitteln den Eindruck einer beeindruckenden, zielgesteuerten Reise, war das so, oder hat jede gefundene Antwort hundert neue Fragen eröffnet, aus denen Sie gewählt haben?

Simon Haas entwickelt Einzelzell-Multi-Omics, um räumlich aufgelöst Leukämiestammzellen zu enttarnen – und therapeutisch anzugehen. (Quelle: © Felix Petermann, MDC).

Es ist natürlich eine Kombination aus beidem, man hat im Idealfall ein Ziel vor Augen, bekommt aber tagtäglich neue Ergebnisse und orientiert sich basierend hierauf neu. Man sieht, was mit einer neuen speziellen Methode überhaupt möglich ist und macht, was zuvor noch nicht möglich war. **Eindeutig lebt man in der Wissenschaft von unerwarteten Ergebnissen**. Es ist etwas ganz Tolles, diese zu verfolgen und so vielleicht zu noch viel spannenderen Erkenntnissen zu gelangen, als die Ursprungsfrage zu Beginn vermuten ließ.

Gehört die Entdeckung, dass die Mikroumgebung so wichtig ist, zu den Überraschungen, oder war das zu erwarten?

Dass die zelluläre Umgebung bei Krebserkrankungen eine wichtige Rolle spielt, ist schon länger bekannt. Insbesondere das Immunsystem ist ständig damit beschäftigt dem Krebs entgegenzuwirken. Wir haben Methoden entwickelt, die es erlauben, diese Mikroumgebung genauer zu betrachten, zunächst um **das gesunde Stammzellsystem zu verstehen. Jetzt wenden wir diese auch an, um Leukämien besser zu verstehen**.

Was sind das für Methoden und warum funktionieren sie besonders gut?

Durch Einzelzellmethoden können wir ein Gewebe in die einzelnen Zellen zerlegen und dadurch verstehen, welche Komponenten das gesunde oder das Krebsgewebe ausmachen. Insbesondere hat unser Konsortium verschiedene **Einzelzell-Multi-Omics Methoden** entwickelt, welche es erlauben, aus einzelnen Zellen mehrere Informationen gleichzeitig zu bestimmen. So können wir z. B. die Aktivität von Tausenden von Genen, DNA-Schäden (z. B. Mutationen) und Zelloberflächenmolekülen gleichzeitig in einzelnen Zellen detektieren. Das ist besonders vielversprechend, weil viele zielgerichtete Therapien häufig an den Zellstrukturen der Oberfläche wirken. Die von uns entwickelten Methoden setzen wir nun ein, um ganz gezielt molekulare Targets zu identifizieren, um Krebszellen spezifisch angreifen und beseitigen zu können.

Eine Einschränkung der Einzelzell Methoden ist allerdings, dass die räumliche Information verloren geht. Man versteht zwar die einzelnen Zellen, die das Grundgerüst eines Gewebes bilden, kann aber nicht mehr nachvollziehen, wo diese räumlich zu verorten sind. Klassischerweise hat man die räumliche Anordnung von Zellen in Geweben durch Mikroskopie untersucht, wobei man aber nur wenige Parameter gleichzeitig betrachten kann. In den letzten Jahren haben wir und andere daher sogenannte **„Spatial Omics“ Methoden** entwickelt, mit denen man auch räumlich aufgetrennt Tausende molekulare Informationen visualisieren kann. Die Idee ist, diese zwei Welten, die räumlichen und die Einzelzell-Omics Methoden, miteinander zu verbinden. So können wir besser verstehen, wie ein Organ oder wie Krebs aufgebaut und strukturiert ist.

„Die von uns entwickelten Methoden setzen wir nun ein, um ganz gezielt molekulare Targets zu identifizieren, um Krebszellen spezifisch angreifen und beseitigen zu können.“

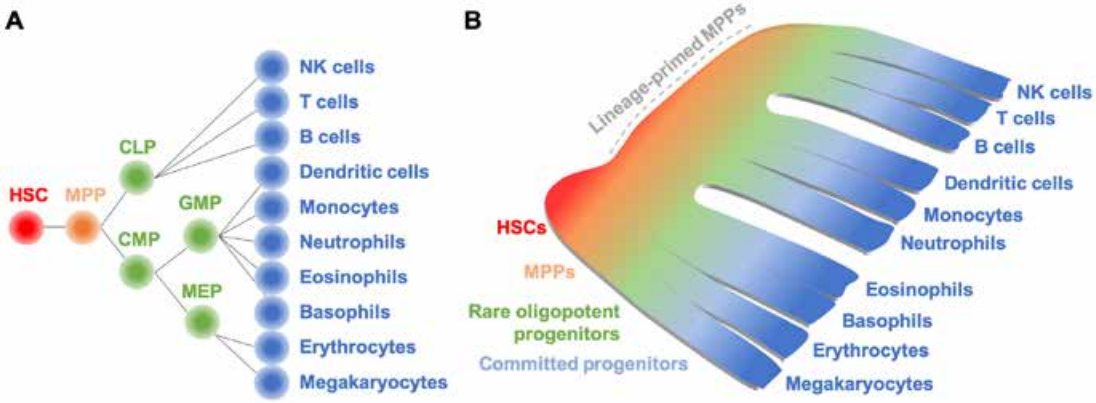


Abbildung 1: A) Klassisches Modell der Blutbildung (Hämatopoese), B) Revidiertes Modell der Blutbildung
(Quelle: https://www.sciencedirect.com/science/article/abs/pii/S0301472X20302678).

Das ist wirklich wahnsinnig spannend! Tragen diese neuartigen Methoden auch zu einem besseren Verständnis von physiologischen Prozessen bei? In welchen Methoden sehen Sie das größte Potential?

Wir konnten vor kurzem zeigen wie sich Blutstammzellen entwickeln, die für die lebenslange Neubildung von Blut- und Immunzellen verantwortlich sind. Durch die weltweit erstmalige Charakterisierung dieser seltenen Zellen mittels Einzelzell-Multi-Omics Technologien konnten wir nicht nur ein genaues molekulares und zelluläres Verständnis dieses wichtigen Prozesses erlangen, sondern auch ein **neues Modell zur Blutbildung** beitragen, das mittlerweile als Standardmodell anerkannt ist. Es gab ein Textbuch-Verständnis davon, wie Blut und Immunzellen im Knochenmark gebildet werden, in Zellpopulationen, die sich von einer Stammzelle ausgehend nach und nach verzweigen, wie ein Baum. Bevor wir die Einzelzell-Methoden dafür verwendet haben, konnte sich das niemand genauer anschauen. In Wirklichkeit ist es KEIN schrittweiser Prozess, sondern wir konnten zeigen, dass a) sich die Stammzellen schon sehr früh entscheiden, in welche Blutlinien sie gehen, und b) diese Differenzierung ein kontinuierlicher Fluss ist. Dieses Modell hat jetzt sozusagen das alte Modell abgelöst.

„Berlin ist momentan ein extrem heißes Pflaster, viel Umbruch, spannender Forschungsstandort, neue Institutionen."

Durch die Anwendung der von uns entwickelten *Spatial Omics* Methoden konnten wir darüber hinaus die Knochenmarksumgebung, in der sich Blutstammzellen befinden, mit einer enorm hohen Auflösung untersuchen. Dabei haben wir mehrere **neue Zelltypen entdeckt**, die die Stammzellen in ihrer Funktion unterstützen.

Das größte **Potential für die Zukunft** sehe ich in den Einzelzell-Multi-Omics Analysen, mit denen wir Hunderte von **Oberflächenmarkern** analysieren können, auf Grund des großen **therapeutischen Potentials** der Methode. Wir arbeiten derzeit daran, sie auch im klinischen Kontext anzuwenden.

*Das wäre ein entscheidender Schritt! Sind die Bedingungen hierfür in **Berlin** an der Charité besonders günstig? Sie sind ja erst seit gut einem Jahr in Berlin, was waren Ihre Gründe für diese Wahl?*

Berlin ist momentan ein extrem heißes Pflaster, viel Umbruch, spannender Forschungsstandort, neue Institutionen. Hier bin ich mit drei Institutionen assoziiert: Mit dem BIH, das sich auf die medizinische Translation spezialisiert hat, mit der Charité, einem der größten Universitätskrankenhäuser Europas, und dem MDC, das in der Entwicklung von Einzelzelltechnologien federführend ist. In dem Fokusschwerpunkt „Single-Cell-Ansätze für die personalisierte Medizin", der von Prof. Dr. Angelika Eggert und Prof. Dr. Nikolaus Rajewsky initiiert wurde, finden die Institute eine Plattform, auf der sie gemeinsam Einzelzell-Methoden entwickeln. In der Charité bin ich auch Teil der Kliniken für Hämatologie, Onkologie und Tumorimmunologie. Dies ermöglicht eine enge Interaktion mit den Klinikern und

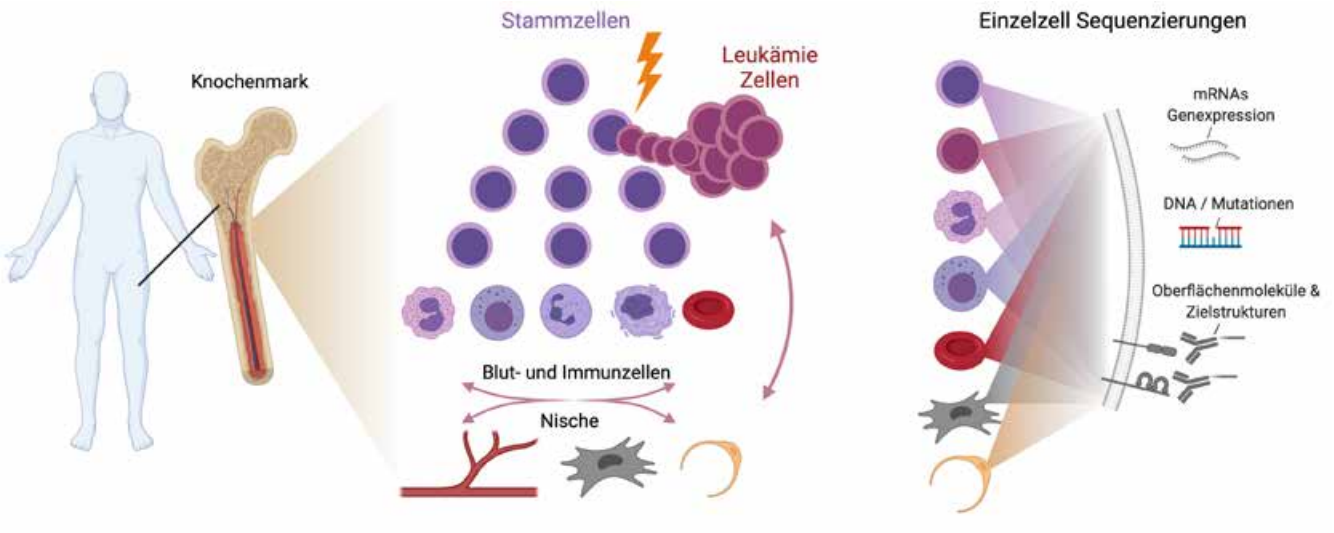


Abbildung 2: Einzelzell-Sequenzierungen ermöglichen das Studium von Leukämien. Gesunde hämatopoetische Stammzellen im Knochenmark produzieren Blut- und Immunzellen. Die zelluläre Umgebung im Knochenmark (Nische) ist für die Blutbildung von großer Bedeutung. Leukämien entwickeln sich aus hämatopoetischen Stamm- oder Vorläuferzellen durch die Anhäufung von DNA-Schäden (z.B. Mutationen). Einzelzell-Sequenzierungen ermöglichen das genaue Verständnis der molekularen Prozesse, die sich während der Transformation von gesunden Stammzellen in Leukämiezellen abspielen. Auch die Interaktion zwischen den Leukämiezellen und der zellulären Knochenmarksumgebung kann untersucht werden.
(Quelle: © Simon Haas, Created with BioRender.com).

die direkte Integration der neu entwickelten Methoden in den klinischen Alltag. Was mir in Berlin besonders gefällt, ist die **Energie und Willensbereitschaft, gemeinsam Bedeutendes zu erreichen**, auch über die Institute hinweg mit großer Interdisziplinarität.

Die **Förderung des BMBF über e:Med** hat zu einem guten Start in Berlin beigetragen. Vor allem die spezielle **Förderung für junge Wissenschaftler** ist wertvoll, um kreativen Ideen Platz zu geben. Die Systemmedizin wird in allen Bereichen der modernen Medizin wichtig sein und einen bedeutenden Beitrag für eine bessere medizinische Versorgung leisten.

„Es ist sehr befriedigend zu sehen, dass die eigene Arbeit das Verständnis in einem Feld vorantreibt."

*Sie sind in der **Heidelberger** Wissenschaftsszene „aufgewachsen", waren dort bis vor Kurzem beim Deutschen Krebsforschungszentrum (DKFZ) und HI-STEM tätig. Welche Rolle hat das für Sie gespielt?*

Ich habe in Großbritannien, den USA und Deutschland studiert und geforscht. Die Heidelberger Forschungsumgebung hat mich aber sicher am meisten geprägt. Heidelberg ist ein hervorragender Forschungsstandort, besonders in der Biomedizin und Krebsforschung, insbesondere durch das DKFZ, das EMBL und die Uniklinik. Diesem Umfeld habe ich unglaublich viel zu verdanken. Schon während des Studiums wurde ich dort in Forschungsaktivitäten eingebunden, wie z.B. in das von Prof. Dr. Roland Eils koordinierte iGEM Programm im Bereich der synthetischen Biologie. Am DKFZ habe ich unter Dr. Mareike Essers meine Doktorarbeit absolviert, und im von Prof. Dr. Andreas Trumpp geleiteten *Heidelberg Institute for Stem Cell Technology* (HI-STEM) habe ich meine erste eigene Forschungsgruppe geleitet. Auch derzeit leite ich dort noch eine Forschungsgruppe und habe viele Kooperationen mit dem Standort Heidelberg. Das wird sicher auch in Zukunft so bleiben.

Simon Haas in seinem Labor am BIMSB-MDC/BIH
(Quelle: @ Thomas Rafalzyk, BIH).

Sehen Sie das Übertragen Ihrer Arbeit in die Klinik?

Definitiv. Es ist kein leichter Weg, aber das ist unser Ziel. **Wir arbeiten daran, bis es klappt** – und das mit Hochdruck, vor allem in enger Zusammenarbeit mit den Kliniken der Hämatologie und Onkologie der Charité und der Universitätsklinik Heidelberg. Ich bin sehr zuversichtlich, dass man viele Aspekte der Grundlagenforschung nutzen kann, um die Prognose, Diagnose und Therapie von Patienten erheblich zu verbessern. Wenn ein Krebspatient derzeit in die Klinik kommt, wird er mit einer Batterie von unterschiedlichen diagnostischen Assays untersucht. Basierend auf den Einzelzell-Multi-Omics Analysen versuchen wir diagnostische Tests zu entwickeln, die alle Tests in einem verbinden, um so möglichst präzise Vorhersagen geben zu können, welcher Patient am besten auf welche zielgerichtete Therapie anspricht. Noch Zukunftsmusik, aber hoffentlich nicht mehr lange.

„Ich bin sehr zuversichtlich, dass man viele Aspekte der Grundlagenforschung nutzen kann, um die Prognose, Diagnose und Therapie von Patienten erheblich zu verbessern."

Sie sind sehr kreativ, umtriebig ...was treibt Sie in Ihrer Arbeit täglich an?

Ich liebe es Neuland zu betreten, neue Wege zu gehen. Das Motivierende an der Wissenschaft ist, dass man täglich vor neuen Herausforderungen steht, dadurch immer wieder Neues lernt und sich zusammen mit der Forschung und den eigenen Entdeckungen weiterentwickelt. Es ist sehr befriedigend zu sehen, dass **die eigene Arbeit das Verständnis in einem Feld vorantreibt**. Wenn der Beruf des Wissenschaftlers auch mit vielen Hürden verbunden ist, ist es auch ein großer Segen, dass man die eigene Forschungsausrichtung selbst bestimmen kann.

Kommen wir noch mal zurück auf Ihre neue Wahlheimat Berlin, wie nehmen Sie sie wahr und was machen Sie, falls Sie mal Freizeit haben sollten?

Ich liebe Berlin. Es ist einfach super divers, steht niemals still und es gibt immer etwas Neues. Auch Berlins Geschichte ist natürlich besonders spannend und an vielen Stellen gut sichtbar. Ansonsten mache ich viel Sport, von meiner Wohnung in Mitte aus kann man gut joggen gehen und Berlin in alle Richtungen erkunden. Da uns die COVID-19 Pandemie seit ich in Berlin bin im Griff hält, freue ich mich darauf, in Zukunft noch stärker in das Berliner Kulturleben eintauchen zu können, als das derzeit möglich ist.

Das Gespräch führte Dr. Silke Argo.

Kontakt:

Dr. Simon Haas
BIH, Charité, MDC Berlin
DKFZ, HI-STEM gGmbH Heidelberg
Simon.haas@bih-charite.de

https://www.bihealth.org/de/forschung/arbeitsgruppe/blutkrebs-stammzellen-praezisionsmedizin

„we love data!“

Porträt der BIH-Professorin Claudia Langenberg

von Stefanie Seltmann

Professorin Claudia Langenberg leitet seit September 2020 die Abteilung für Computational Medicine am Berlin Institute of Health (BIH) in der Charité. Die Expertin für genetische Epidemiologie und Fachärztin für Public Health untersucht die Grundlagen von Stoffwechselerkrankungen wie Typ-2-Diabetes. Dazu nutzt sie große Datenmengen aus internationalen Patienten- und Bevölkerungsstudien. Das BIH rekrutierte die deutsch-britische Forscherin von der University of Cambridge, wo sie die Molekulare Epidemiologie an der Medical Research Council (MRC) Epidemiology Unit leitete.

Die glatten blonden Haare umrahmen das schmale Gesicht, durch die große schwarze Brille blitzen hellblaue Augen: Claudia Langenberg spricht im Zoom-Interview schnell und mit leicht britischem Akzent, obwohl sie in München geboren ist. *„20 Jahre London gehen eben nicht spurlos an einem vorbei“*, lacht die zierliche Medizinerin, die sich mit einem schwergewichtigen Thema befasst: Sie untersucht die Ursachen von Übergewicht, Insulinresistenz oder Typ-2-Diabetes. *„Diese häufigen chronischen Stoffwechselerkrankungen haben sowohl umweltbedingte als auch genetische Ursachen“*, sagt die Fachärztin für Public Health, was soviel wie Öffentliche Gesundheit bedeutet und damit die Gesundheit der gesamten Bevölkerung betrifft. So spielt etwa beim Übergewicht das Überangebot an hochkalorischer Nahrung sowie Bewegungsmangel eine Rolle, aber gleichzeitig erhöhen auch viele Genvarianten das Risiko, zuzunehmen oder das Fett ungünstig zu verteilen. *„Wir brauchen riesige Studien mit zehn- oder hunderttausenden Teilnehmerinnen und Teilnehmern, um den Einfluss einzelner Gene eindeutig feststellen zu können. Schon allein, weil wir Millionen von Genvarianten testen“*, erklärt Claudia Langenberg den Kern ihrer Forschung. *„Dazu brauchen wir leistungsfähige Software und Computer. Trotz der relativ kleinen Effekte einzelner Varianten, können uns diese helfen, die Ursachen metabolischer Erkrankungen zu entschlüsseln. So können wir abschätzen, welche Erfolgsaussichten neue oder bereits vorhandene Behandlungen haben.“* *„We love data“*, kommentiert sie das neue Fachgebiet „Computational Medicine“ oder datenbasierte Medizin.

„Wir brauchen riesige Studien mit zehn- oder hunderttausenden Teilnehmerinnen und Teilnehmern, um den Einfluss einzelner Gene eindeutig feststellen zu können“

Foto: ThomasRafalzyk/BIH

Der Regierende Bürgermeister von Berlin, Michael Müller, eröffnet die Ausstellung „Berlin – Hauptstadt der Wissenschaftlerinnen" mit einigen der portraitierten Frauen, unter ihnen auch Claudia Langenberg, ganz links (Foto: BIH/Konstantin Börner).

„Architektur des Metabolismus"

In den letzten Jahren interessierte sich Claudia Langenberg insbesondere für die Metabolite und Proteine, Stoffwechselprodukte im menschlichen Blut. Bei jedem Menschen ist die Zusammensetzung anders, und darauf haben genetische Variationen großen Einfluss. *„Mittlerweile können wir Hunderte bis Tausende dieser Moleküle nachweisen und im großen Stil messen, das hätte man sich vor ein paar Jahren noch nicht in dieser Dimension vorstellen können"*, schwärmt die Wissenschaftlerin. Diese Arbeit hat es dem Team in Zusammenarbeit mit Kolleg*innen anderer Gruppen ermöglicht, einen Atlas der *„Architektur des Metabolismus"* zu erstellen. *„Es ist ungewöhnlich, dass man als Epidemiologin Wissen für die nächste Generation von Biochemiebüchern schafft"*, freut sich Claudia Langenberg.

> „Mittlerweile können wir Hunderte bis Tausende dieser Moleküle nachweisen und im großen Stil messen, das hätte man sich vor ein paar Jahren noch nicht in dieser Dimension vorstellen können."

„Besonders spannend ist es allerdings, Veränderungen, die zu Stoffwechselerkrankungen führen, von subtilen Messunterschieden bei Gesunden zu unterscheiden. Viele der verantwortlichen Gene, die wir gefunden haben, waren schon bekannt", berichtet Langenberg. *„Sie rufen schwerwiegende „monogene" Erkrankungen hervor, die sehr selten sind und oft schon im Kindesalter auffallen. Offensichtlich führen dieselben Gene auch in der gesunden Allgemeinbevölkerung zu Veränderungen im Stoffwechsel. Unsere Aufgabe ist jetzt herauszufinden, ob und wie diese sich klinisch bemerkbar machen, weil wir so neue Angriffspunkte für zielgerichtete Therapien identifizieren können"*, sagt Langenberg, was ganz zum Motto des BIH „Aus Forschung wird Gesundheit" passt.

Rolle der Gene auch bei COVID-19

„Ich bin sehr gerne nach Berlin gekommen", sagt Claudia Langenberg. *„Einerseits können uns die Translations-Hubs und Core Facilities des BIH großartig dabei unterstützen, die großen Datenmengen auszuwerten. Andererseits finde ich am BIH und an der Charité international renommierte Kooperationspartnerinnen und -partner."* So kollaboriert Claudia Langenberg schon länger mit Professor Markus Ralser, dem Direktor des Instituts für Biochemie an der Charité, unter anderem, um Proteinbestandteile im Blut von COVID-19-Patient*innen zu analysieren. *„Wir konnten die Ergebnisse der Berliner Studie nutzen, um genetische Einflussfaktoren auf COVID-19-relevante Proteine in unseren Studien aus Cambridge zu entschlüsseln und sie schnell der wissenschaftlichen Gemeinschaft zur Verfügung stellen"*, erklärt Langenberg. Dafür haben wir mit

Kolleg*innen aus München einen interaktiven Webserver entwickelt (www.omicscience.org).

Auch privat gefällt es Claudia Langenberg in Berlin sehr gut: *„Es ist eine tolle Stadt. Mein Vater kam gebürtig aus Berlin und ich finde, es gibt keinen besseren Ort, meiner Familie die Geschichte Deutschlands näherzubringen. Die ‚Berliner Schnauze' ist zwar eher das genaue Gegenteil der ‚feinen englischen Art', aber solche Gegensätze machen das bi-nationale Leben ja gerade so spannend"*, lacht die Mutter von zwei siebenjährigen Zwillingstöchtern. Dies war auch ein Grund für den Wechsel nach Deutschland. *„Ich wollte gern, dass die Mädchen auf eine zweisprachige deutsche Schule gehen, um mehrere Kulturen und Sprachen zu leben. Die internationale Gemeinschaft dort ist wirklich toll, auch wenn der Schulstart im Coronajahr natürlich ziemlich holprig war. Wir hoffen sehr, dass es jetzt wieder regelmäßig Präsenzunterricht geben wird"*, erzählt Claudia Langenberg. Ihr britischer Ehemann ist ebenfalls Wissenschaftler am University College London und arbeitet als Gastwissenschaftler am BIH.

Claudia Langenberg hat noch nicht alle Zelte in England abgebrochen und leitet weiterhin ein Team in Cambridge. *„Einige große, laufende Studien werden erst in ein bis zwei Jahren abgeschlossen sein, und wir sind gerade dabei, neue internationale Kollaborationen zu bilden, die dann von Berlin aus geleitet werden. Ein Gewinn für beide Seiten."*

„Wir konnten die Ergebnisse der Berliner Studie nutzen, um genetische Einflussfaktoren auf COVID-19-relevante Proteine in unseren Studien aus Cambridge zu entschlüsseln und sie schnell der wissenschaftlichen Gemeinschaft zur Verfügung stellen."

Kontakt:

Prof. Dr. Claudia Langenberg
Computational Medicine, Berlin Institute of Health at Charité und MRC Epidemiology Unit, University of Cambridge, UK
Claudia.Langenberg@charite.de

https://www.bihealth.org/de/forschung/arbeitsgruppe/computational-medicine

Profil Claudia Langenberg:

Claudia Langenberg wurde in München geboren, studierte Medizin in Münster und wechselte nach klinischer Ausbildung und Tätigkeit in Deutschland für ihren Master und PhD in Epidemiologie nach England und in die USA. 2016 schloss sie ihre Facharztausbildung in Public Health ab und übernahm 2017 die Programmleitung der Molekularen Epidemiologie an der University of Cambridge, wo sie zuvor genetische Grundlagen und Risikofaktoren metabolischer Erkrankungen entschlüsselt hatte und am Aufbau internationaler Konsortien und an Meta-Analysen beteiligt war. Claudia Langenberg hat über 250 Publikationen, viele in hochrangigen Journalen, und gehört als „Highly Cited Researcher" des Thomson Reuters Highly Cited Researchers Ranking zu den meistzitierten Wissenschaftlerinnen in ihrem Fachgebiet. Sie hat zahlreiche wissenschaftliche Auszeichnungen für ihre Arbeit erhalten und war Chef-Editorin des Reports „Generation Genome" des englischen Chief Medical Officers, der maßgeblich zur Umstrukturierung genomischer Medizin im englischen NHS (National Health Service) beitrug.

Der Digitale FortschrittsHub CAEHR

Optimale & personalisierte Therapieentscheidungen für Herzkreislauf-Patient*innen

von Dagmar Krefting

Fehlende Kommunikation an den Schnittstellen im Gesundheitssystem führt oft durch Informations- und Zeitverlust zu einer nicht optimalen Behandlung. Mit dem Digitalen FortschrittsHub CAEHR – koordiniert von der Universitätsmedizin Göttingen – soll eine forschungskompatible elektronische Patientenakte (ePA) entwickelt und implementiert werden, sodass zukünftig strukturierte Patientendaten an allen Punkten des Versorgungssystems nach einheitlichen Standards erhoben und über die gesamte Versorgungskette genutzt werden können. Damit sind präzisere Prognosen zu Krankheitsverläufen möglich, die wiederum zu optimierten Therapien und neu entwickelten Versorgungsmodellen führen können.

Ob Schlaganfall, Herzinsuffizienz oder koronare Herzkrankheiten – Herz-Kreislauferkrankungen stellen trotz wichtiger Fortschritte in der Behandlung noch immer die häufigste Todesursache in Deutschland dar. Darüber hinaus nehmen diese Erkrankungen oftmals einen chronischen Verlauf. Sowohl die betroffenen Patient*innen als auch die behandelnden Ärzt*innen müssen sich auf eine dauerhafte sowie individuelle und personalisierte Behandlung einstellen.

Was sind die Ziele von CAEHR?

Und genau hier setzt der vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Digitale FortschrittsHub **CAEHR** (**CA**rdiovascular diseases – **E**nhancing **H**ealthcare through cross-sectoral **R**outine data integration) an. Denn werden Herz-Kreislauferkrankungen frühzeitig diagnostiziert, sind diese in vielen Fällen gut behandelbar.

Für genau diese patientenzentrierte Behandlung vereinheitlicht und strukturiert CAEHR die Gesundheitsdaten aus der ambulanten und stationären Versorgung und macht diese in einer forschungskompatiblen ePA allen Akteur*innen entlang des gesamten Behandlungspfades (Sanitäter*innen, Pfleger*innen, Ärzt*innen und Patient*innen) für die individuelle Patientenversorgung zugänglich. Somit geht es CAEHR um eine bessere Gesundheitsversorgung von erkrankten Patient*innen durch die optimierte und zeitnahe Bereitstellung von relevanten, standardisierten und strukturierten Gesundheitsinformationen sowie die Etablierung intelligenter datengetriebener Dienste. In drei Regionen Deutschlands – Hannover/Göttingen, Berlin und Würzburg/Mainfranken – wird CAEHR in den nächsten vier Jahren digitale Lösungen für eine bessere Versorgung der

Die fehlende Kommunikation an den Schnittstellen des Behandlungspfades führt durch Informations- und Zeitverlust zu keiner optimalen Behandlung. Eine Standardisierung und Digitalisierung der Daten ermöglichen die intersektorale Bereitstellung relevanter Informationen am richtigen Ort zur richtigen Zeit – und die Nutzung dieser Daten zur Entwicklung von klinischen Entscheidungsunterstützungssystemen, wodurch eine messbare Verbesserung der Gesundheitsversorgung erreicht werden soll.
(Quelle: © CAEHR-Konsortium)

Quelle: © HiGHmed

Menschen mit Herz-Kreislauferkrankungen erproben und für den späteren bundesweiten Einsatz weiterentwickeln.

Was sind Digitale FortschrittsHubs?

Die vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Digitalen FortschrittsHubs haben das Ziel, die Pionierarbeiten der Medizininformatik-Initiative (MII) zur Digitalisierung in der Medizin aus den Unikliniken – zunächst in Pilotprojekten – in alle Bereiche des Gesundheitssystems einfließen zu lassen: von der ambulanten Versorgung in der Hausarztpraxis über den stationären Aufenthalt im örtlichen Krankenhaus bis zur Versorgung in Rehabilitations- und Pflegeeinrichtungen.

Use Cases

An den verschiedenen Punkten entlang des Behandlungspfades haben die Behandelnden heute oft nur einen unvollständigen Informationszugang zu Patientendaten. CAEHR nimmt daher den Informationsfluss zwischen den verschiedenen Sektoren des Gesundheitssystems anhand dreier Anwendungsfälle (Use Cases) in den Fokus: Der **Use Case „Notfallversorgung“** widmet sich Schlaganfallpatient*innen und der Schnittstelle *akutstationäre und Notfallversorgung*. Der **Use Case „Rehabilitation“** thematisiert bei Patient*innen nach Aortenklappenersatz mittels Kathetertechnik die Schnittstelle *stationäre Versorgung und Rehabilitation* und der **Use Case „Ambulante Betreuung“** fokussiert bei Patient*innen mit chronischer Herzinsuffizienz und koronarer Herzkrankheit die Schnittstelle *stationäre und ambulante Versorgung*.

Der daraus resultierende Mehrwert für die Patient*innen, das beteiligte Fachpersonal, die involvierten Wissenschaftler*innen und das Gesundheitssystem als Ganzes soll in den drei genannten Use Cases nachgewiesen werden.

Use Case 1: Notfallversorgung von Schlaganfällen

Dieser Use Case beschäftigt sich mit einer optimierten Ressourcenallokation. Die bereits im Rettungswagen generierten relevanten Patientendaten sollen direkt in das Ziel-Krankenhaus übertragen und die Notfallmediziner*innen bei zeitkritischen Therapieentscheidungen von KI-getragenen Systemen unterstützt werden.

Use Case 2: Rehabilitation nach einer Herzoperation

In diesem Anwendungsfall werden individuelle Rehabilitationsangebote entwickelt. Dabei soll die Dokumentation digital und nicht mehr papiergebunden erfolgen. Zudem soll die interprofessionelle Planung individueller Therapien gestärkt werden.

Use Case 3: Ambulante Versorgung bei koronarer Herzkrankheit

Dieser Use Case setzt sich eine Verbesserung der Datenschnittstellen durch IT-Lösungen zum Ziel. Tragbare Sensorik soll vermehrt im häuslichen Umfeld Anwendung finden. Darüber hinaus soll eine nahtlose Dokumentation in der Diagnostik und Therapie zur Stärkung der Prävention beitragen.

Projektpartner

Das Projekt CAEHR umfasst insgesamt 28 Verbundpartner und besteht aus den folgenden neun Partnern mit direkter Zuwendung:

- Universitätsmedizin Göttingen und Georg-August-Universität Göttingen
- Medizinische Hochschule Hannover
- Charité - Universitätsmedizin Berlin
- Universitätsklinik Würzburg
- Hochschule Osnabrück
- HiGHmed e.V.
- Vitasystems GmbH
- AOK Niedersachsen
- System Vertrieb Alexander GmbH

Darüber hinaus sind noch weitere zahlreiche Institutionen aus Wissenschaft und Industrie sowie Patient*innenvertretungen aktive Partner im Projekt CAEHR. Eine Übersicht dazu gibt es hier: www.gesundheitsforschung-bmbf.de

Kontakt:

Prof. Dr. Dagmar Krefting
Leiterin des Digitalen FortschrittsHubs CAEHR
Universitätsmedizin Göttingen
dagmar.krefting@med.uni-goettingen.de

Eva König
Presse- und Öffentlichkeitsarbeit HiGHmed
communications@highmed.org

news

Der MTZ®-Award 2022 – jetzt bewerben!

Förderpreis für herausragende Promotionen im Bereich der systemorientierten Gesundheitsforschung

Doktorandinnen und Doktoranden aufgepasst! Ab jetzt können Sie sich für den „MTZ®-Award for Systems Medicine“ bewerben. Die Bewerbungsfrist endet am 18. Februar 2022. Die Preissumme von 10.000 Euro wird unter den besten drei Bewerberinnen und Bewerbern geteilt. Die feierliche Preisverleihung findet während der 8. Internationalen Konferenz „Systems Biology of Mammalian Cells“ vom 16. bis 18. Mai 2022 in Heidelberg statt.

Der „MTZ®-Award for Systems Medicine“ ist ein nationaler Nachwuchs-Förderpreis für herausragende Dissertationen junger Wissenschaftlerinnen und Wissenschaftler auf dem Gebiet der systemorientierten Gesundheitsforschung. Der Preis soll dem vielversprechenden wissenschaftlichen Nachwuchs eine 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.

Der Schwerpunkt der Promotion kann dabei sowohl eine molekulargenetische, klinische, mathematische als auch informatische Herangehensweise verfolgen. Sie soll aber in einem interdisziplinären Umfeld erarbeitet worden sein, um der Arbeitsweise der Systemmedizin gerecht zu werden. Erst durch den Austausch zwischen den Expertisen werden komplexe, miteinander vernetzte biologische Phänomene enträtselt und für die personalisierte Medizin nutzbar gemacht.

Der **MTZ®-Award for Systems Medicine 2022** wird unter diesem Namen zum ersten Mal ausgelobt, setzt aber die erfolgreiche Reihe des MTZ®-Award for Medical Systems Biology fort. Dieser wurde seit 2008 bereits an 21 Preisträgerinnen und Preisträger verliehen.

Weitere Informationen unter:
www.ptj.de/systembiologie

Quelle: Projektträger Jülich

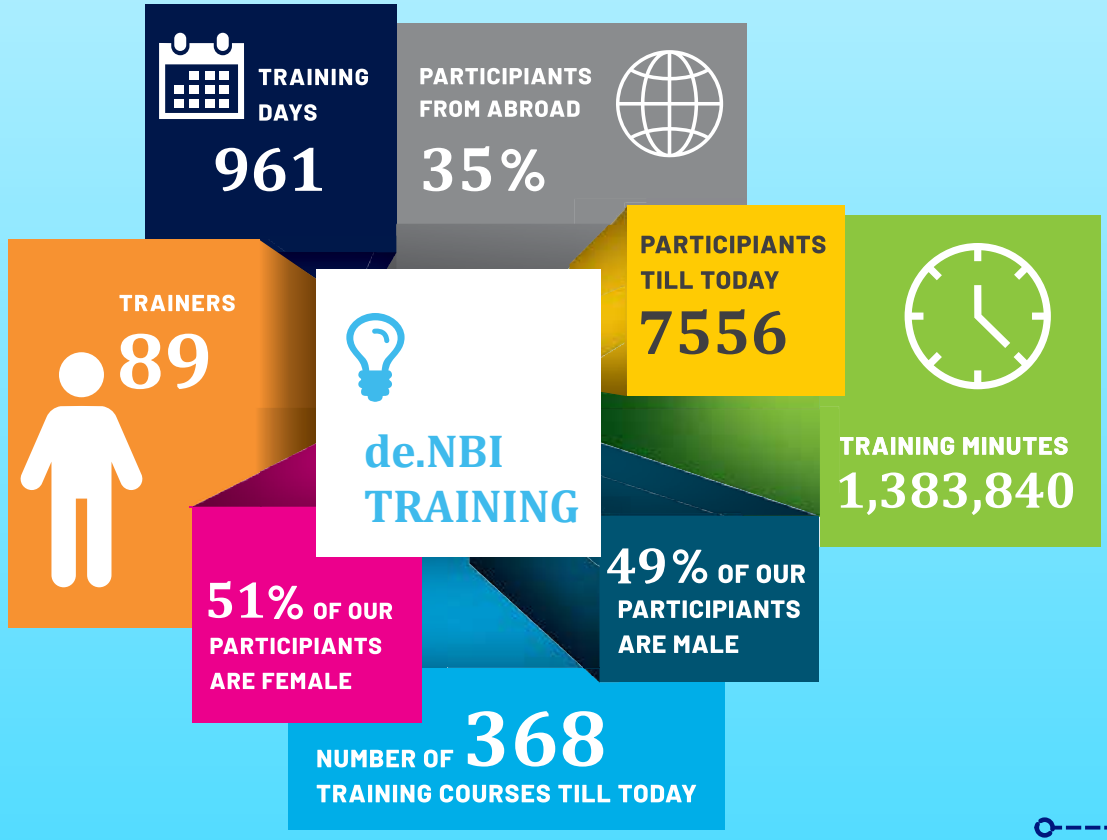
Die Bewerbungsfrist für den „MTZ®-Award for Systems Medicine“ endet am 18. Februar 2022.
(Quelle: shutterstock © Syda Productions).

de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

de.NBI TRAINING

Bioinformatics Training and Competence Building

The German Network for Bioinformatics Infrastructure (de.NBI) organizes training events and provides online training materials to enable bioinformaticians and life scientists to exploit their own and publicly available data more effectively.



Find out more at: www.denbi.de/training

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

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

events

„Genau zur richtigen Zeit“

Die Systems Biology of Human Disease (SBHD) 2021 zum ersten Mal Hybrid

von Franziska Müller und Julius Upmeier zu Belzen
im Namen der SBHD 2021 Organisatoren

Die Gewinnerin des diesjährigen Anne-Heidenthal-Preises Helena Radbruch, links von ihr: Roland Eils (SBHD Chair) und rechts: Georg Draude (Chroma Technology Corp.) (Quelle: Franziska Müller)

Wenn nicht jetzt, wann dann. Unter diesem Motto fand Anfang Juli die 13. jährliche internationale Konferenz zur Systembiologie menschlicher Erkrankungen statt. Vor Ort waren wir nach einer erfolgreichen ersten Konferenz in Berlin 2019 wieder im Kaiserin-Friedrich-Haus. Direkt neben dem berühmten Bettenhochhaus der Charité liegt dieser über 100 Jahre alte Veranstaltungsort.

Trotz einer eher überschaubaren Anzahl an vor Ort Teilnehmenden und einer erheblichen Anzahl hygienischer Vorsichtsmaßnahmen war die Veranstaltung ein voller Erfolg: nach über einem Jahr Abstinenz von jeglichen nicht im virtuellen Raum stattfindenden Treffen, war die Community kaum in Ihrer Euphorie sich von Angesicht zu Angesicht wieder zu sehen, zu bremsen.

„Ich kam mir vor wie auf einem Festival“ war eine von vielen Rückmeldungen, die uns im Nachgang zu Ohren kamen. Schon merkwürdig wie 2 Jahre Pandemie die Wahrnehmung verändern: waren wir doch knapp 60 Leute in einem für bis zu 170 Personen ausgelegten Saal. Aber ja, auch wir können dem Teilnehmer nur zustimmen: wir kamen uns ebenfalls vor wie auf einem Festival, die Stimmung war eben grandios. Ein Stück weit haben das sicher auch die 65 online TeilnehmerInnen wahrgenommen, die, wenn auch nicht vor Ort, gleichermaßen Vorträge hielten, mitdiskutierten und Poster vorstellten.

Aber nicht nur das Rahmenprogramm, welches sichtlich von den TeilnehmerInnen genossen wurde, hat die Veranstaltung zu diesem Erfolg geführt, sondern vielmehr die äußerst spannenden Beiträge und Diskussionen:

Die aktuelle COVID-19 Pandemie war nicht nur in den Hygienemaßnahmen und dem hybriden Konferenzformat präsent, eine Vielzahl der Beiträge behandelte direkt die COVID-19-Forschung. Dabei wurden die verschiedenen Herangehensweisen von Single-cell Transcriptomics über Proteomics und Fluoreszenzmikroskopie bis hin zu klassischer mathematischer Modellierung und Machine Learning, abgedeckt. Ziel dieser Analysen war neben einem verbesserten Verständnis des Virus und seiner molekularen Mechanismen die Identifikation möglicher Wirkstoffe gegen das Virus.

Neben der Erforschung von COVID-19 wurden diese Methoden selbstverständlich weiterhin für die Erforschung von Erkrankungen wie Krebs, Alzheimer, Tuberkulose sowie ALS genutzt. Auch in diesen Feldern gab es daher spannende Ergebnisse zu berichten.

Doch es wurden nicht nur Projekte zu spezifischen Krankheiten präsentiert, sondern auch solche, die wertvolle Erkenntnisse für eine Vielzahl medizinischer Fragen liefern, beispielsweise der Human Lung Cell Atlas. Neben diesen äußerst detaillierten Datensätzen weniger Individuen waren ihr Gegenstück, die Populationskohorten ebenfalls ein Thema. In diesen stehen einige ausgewählte Datenpunkte für eine Vielzahl von Individuen zur Verfügung und erlauben so eine weitere Perspektive auf humane Erkrankungen.

Nicht zuletzt bildet die Methodenentwicklung einen fundamentalen Bestandteil aktueller Fortschritte und wurde entsprechend ausführlich diskutiert. Diese Methoden umfassen verbesserte statistische Verfahren, Differenzialgleichungssysteme, klassisches Machine Learning, sowie Deep Learning. Wiederkehrende Themen bei diesen diversen Herangehensweisen waren die Auswahl und Kritik zugrundeliegender Annahmen, sowie Verbesserungen in der Evaluierung und Untersuchung möglicher „Biases“ dieser Modelle.

Ein Höhepunkt der Konferenz war sicherlich der Vortrag von **Helena Radbruch** vom Institut für Neuropathologie an der Charité über ihre Arbeiten zum Eindringen des SARS-Cov-2 ins Gehirn, für diese sie mit dem Anne-Heidenthal-Preis für Fluoreszenzmikroskopie von Chroma Technology ausgezeichnet wurde.

Zwei weitere Preise, gestiftet vom Journal Molecular Systems Biology by EMBO Press, wurden an diesem Tag für das beste Poster und den besten Short-Talk verliehen. Der Best Poster Award im Wert von 200 € ging an **Jens Hansen** von Mount Sinai, New York für sein Poster zu Pathway-Netzwerkananalysen in Multiomics Daten des Kidney Precision Medicine Projects. Der Best Short Talk Preis ging an **Brian Orcutt-Jahns** von der UCLA für seinen Vortrag über Multivalenz als neue Achse in der Entwicklung von Therapien die regulatorische T-Zellen zum Ziel haben.

Neben diesen ausgezeichneten Vortragenden haben viele weitere exzellente Wissenschaftlerinnen und Wissenschaftler das Programm bereichert, so zum Beispiel **Bree Aldridge** (Tufts School of Medicine), **Bernd Bodenmiller** (Universität Zürich), **Monique van der Wijst** (Universität Groningen), **Trey Ideker** (UC San Diego), und **Claudia Langenberg** (Berlin Insitute of Health at Charité).

Mit Vorträgen und Diskussionen zur Methodenentwicklung, Datensammlung und der spezifischen Erforschung einzelner Krankheiten von ALS über Tuberkulose bis hin zu COVID-19, war die diesjährige SBHD für viele in der Community ein fantastischer Auftakt für die Rückkehr zu Veranstaltungen in Person. Die Wichtigkeit von Präsenzveranstaltungen für das Netzwerken und den wissenschaftlichen Austausch wurde dieses Jahr wieder vielen bewusst.

Wir freuen uns mitsamt der Community auf eine erfolgreiche Fortsetzung der beliebten Konferenzreihe im nächsten Jahr.

Kontakt:

Franziska Müller
BIH at Charité – Zentrum für Digitale Gesundheit
franziska.mueller@charite.de
www.sbhdberlin.org

Roland Eils' Eröffnungsrede zum Konferenzdinner
(Quelle: Claudia Böckermann)

events

Streitsache

Gibt es ein Recht auf medizinische Behandlung mit KI?

von Matthieu-P. Schapranow, Hannelore Loskill und Klemens Budde

KI-basierte Assistenzsysteme können Ärzt*innen künftig bei Diagnose und Behandlung von Krankheiten unterstützen. Müssen sie in Zukunft möglicherweise sogar zu Rate gezogen werden, wenn der Wunsch direkt von den Patient*innen ausgeht? Dieser Frage ging die Plattform Lernende Systeme (PLS) gemeinsam mit der Leibniz Universität Hannover am 22. Sept. 2021 in einem fiktiven Gerichtsverfahren nach. Grundlage war ein realitätsnaher Fall, der am Zentrum für Kunst und Medien (ZKM) in Karlsruhe in einer fiktiven Gerichtsverhandlung diskutiert wurde.

Die Streitsache: Worum geht es?

Der Fall orientiert sich an dem in der PLS entwickelten Anwendungsszenario “Mit Künstlicher Intelligenz gegen Krebs”. Es zeigt anhand eines realitätsnahen medizinischen Fallbeispiels die Einsatzmöglichkeiten lernender Systeme bei der Behandlung von Krebspatient*innen auf. Dabei geht es um den Protagonisten Herrn Anton Merk, der jüngst mit der Diagnose Lungenkrebs konfrontiert wurde.

Nun ist Herr Merk bestrebt die bestmögliche Therapie für sich zu finden. Er hat dazu eine Lungenfachklinik aufgesucht, in der er seine Tumorerkrankung behandeln lassen will.

Seine behandelnde Ärztin möchte sich nicht nur auf ihre Expertise und die der Kolleg*innen im Klinikum verlassen. Daher möchte sie zusätzlich ein Entscheidungsunterstützungssystem einsetzen, das in der Lage ist die Daten vieler ähnlicher Lungenkrebsfälle mit Hilfe von KI-Verfahren zu analysieren und so die Arbeit ihre Arbeit zu unterstützen. Seine Ärztin ist im Umgang mit diesem Assistenzsystem vertraut und hat bereits gute Erfahrungen bei dessen Einsatz gemacht.

Nun fordert Herr Merk also, dass seine Behandlung mit dem KI-Assistenzsystem durchgeführt wird. Die Lungenfachklinik, in der Herr Merk derzeit behandelt wird, lehnt den Einsatz der Software hingegen ab.

Die Klinikleitung fürchtet die unklare Haftungslage im Falle eines Behandlungsfehlers oder Schadens an Herrn Merk. Darüber hinaus werden die Kosten für den Einsatz des KI-Systems bislang nicht von der Krankenkasse übernommen. Die Klinik fühlt sich für den Einsatz in der Versorgung bislang nicht ausreichend gewappnet.

Während einer Operation assistiert ein KI-basiertes Navigationssystem Herrn Merks Ärztin und warnt zuverlässig, wenn sie etwa

Zur Unterstützung des Gerichts in Karlsruhe wurden Sachverständige live zugeschaltet und befragt. Darunter auch Herr Dr. Matthieu-P. Schapranow vom Digital Health Center des Hasso-Plattner-Instituts in Potsdam (Foto: Zentrum für Kunst und Medien Karlsruhe / Elias Siebert).

in die Nähe von Blutgefäßen gerät – besonders nach langen Schichten eine große Hilfe für eventuell erschöpftes Personal. Darüber hinaus möchte Herr Merk, dass das oben beschriebene KI-basierte Entscheidungsunterstützungssystem bei der Auswahl der passenden medikamentösen Behandlung zum Einsatz kommt.

Herr Merk klagt nun vor Gericht, um den Einsatz des KI-Systems zu erzwingen. Die Streitsache lautet: Gibt es ein Recht auf Behandlung mit KI?

Gefragt: Kompetenter Umgang mit Empfehlungen von KI-Assistenzsystemen

Expert*innen der PLS aus den Bereichen Patientenperspektive, Künstliche Intelligenz, Medizin, und Ethik unterstützen das hohe Gericht bei der Entscheidungsfindung als Sachverständige in Ihren jeweiligen Fachdisziplinen.

Hannelore Loskill, Bundesvorsitzende der BAG Selbsthilfe, vertrat die Perspektive der Patient*innen und unterstrich, dass die Wünsche von Herrn Merk im Mittelpunkt der Entscheidungen stehen sollten, damit er nach bestem Stand der Wissenschaft behandelt werden könne.

Dr. Matthieu-P. Schapranow, Dozent und Forscher am Digital Health Center des Hasso-Plattner-Instituts, informierte Zuschauer und das Gericht über die technischen Fähigkeiten von KI-Assistenzsystemen und die Erklärbarkeit von Ergebnissen, die durch KI-Systeme gewonnen werden.

Prof. Dr. Klemens Budde, leitender Oberarzt an der Charité – Universitätsmedizin Berlin, vertrat als dritter Sachverständiger die Position der Mediziner*innen. KI-gestützte Assistenzsystem sind ein Add-on, welches Mediziner*innen mit mehr Wissen versorgt und sogar Zugang zu weltweiten Erfahrungen ermöglicht. Gleichzeitig wies er auch auf Gefahren durch unvollständige oder verzerrte Daten hin, die in weniger zuverlässigen Prognosen resultieren würden. Am Ende werde es aber Herrn Merks Fachärzt*in sein, die die Empfehlungen des Assistenzsystems bewertet und über die Behandlung entscheidet. Budde verwies allerdings auf die Gefahr einer „Kochbuch-Medizin“, in der sich Ärztinnen und Ärzte zu sehr auf das KI-System verlassen. Deshalb sei es wichtig, dass Ärztinnen und Ärzte künftig im Umgang mit Empfehlungen des Systems und ihnen zugrunde liegenden Wahrscheinlichkeiten sowie Daten umzugehen wissen – und im Zweifel auch abweichende Entscheidungen zu treffen wagen.

Mit dem Einsatz eines KI-Assistenzsystems stellen sich aber auch ethische Fragen. Frau **Dr. Jessica Heesen**, Medienethikerin an der Eberhard Karls Universität Tübingen, verwies darauf, dass das KI-System die Kompetenz und Autorität der Ärzt*in infrage stellen könnte. Wem vertrauen die Menschen schlussendlich mehr – der Technik oder ihrer Ärzt*in? Wer entscheidet künftig darüber, ob eine ärztliche Entscheidung richtig oder falsch ist? Diese Fragen können nicht abschließend geklärt werden.

Das Urteil: Ein Kompromiss für alle Beteiligten

Auch wenn es viel Befürwortung für den Einsatz des KI-Systems gab, konnten nicht alle Fragen abschließend geklärt werden. Ob es künftig ein Recht für Patienten auf die Behandlung mit Hilfe eines KI-basierenden Assistenzsystems geben wird, bleibt ebenso abzuwarten. Auf der Bühne des ZKM einigten sich die fiktive Klinik und der Kläger auf einen individuellen Behandlungsvertrag, der die Haftung für die Klinik minimiert und die Kostenübernahme durch Herrn Merk regelte. Auf Grundlage dieses persönlichen Behandlungsvertrags ist der Einsatz des KI-Assistenzsystems nun für beide Seiten realisierbar.

Das in diesem Fall gewählte Format eines fiktiven Gerichtsverfahrens ist eine moderne Form der Wissenschaftskommunikation. In einem exemplarischen Fall werden Fakten, Argumente und bestehende Rechtslage auf lockere und dennoch seriöse Weise in verteilten Rollen präsentiert. Die Veranstaltung richtete sich an

events

interessierte Bürger*innen und setzte weder juristisches noch medizinisches Fachwissen voraus. Ziel der Arbeit der PLS ist es, über Möglichkeiten von KI-Systemen aufzuklären, zugleich ihren Einsatz kritisch zu reflektieren und juristische Voraussetzungen zu klären, die ihren Einsatz ermöglichen.

Das Gerichtsverfahren wurde von Frau **Dr. Jessica Heesen** moderiert, die darauf achtete, dass auch ethische Gesichtspunkte bei der Betrachtung des Gerichts nicht untergingen. Frau **Prof. Dr. Susanne Becker** ordnete die Möglichkeiten eines wie in diesem Fall angestrebten Verfahrens in einen juristischen Kontext ein.

Links:

Weitere Details zur Veranstaltung und zur Arbeit der PLS sowie eine Videoaufzeichnung sind online verfügbar unter:
www.plattform-lernende-systeme.de/streitsache.html

Kontakt:

Dr.-Ing. Matthieu-P. Schapranow
ist Mitglied der Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ der Plattform Lernende Systeme, Leiter der Arbeitsgruppe „In-Memory Computing for Digital Health“ und wissenschaftlicher Leiter Digital Health Innovations am Hasso-Plattner-Institut (HPI).
schapranow@hpi.de
https://hpi.de/digital-health-center/members/working-group-in-memory-computing-for-digital-health/dr-ing-matthieu-p-schapranow.html

Hannelore Loskill ist Bundesvorsitzende der BAG Selbsthilfe und Mitglied der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme.
info@bag-selbsthilfe.de
https://www.bag-selbsthilfe.de/bag-selbsthilfe/ueber-uns/vorstand

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

Nicht nur vor Ort in Karlsruhe sondern auch via Internet-Livestream waren interessierte Bürger*innen eingeladen an der Abendveranstaltung teilzunehmen. Dabei wurde weder juristisches noch medizinisches Fachwissen vorausgesetzt (Foto: Zentrum für Kunst und Medien Karlsruhe / Elias Siebert).

Virtuell die Systemmedizin vernetzen

e:Med Meeting 2021

von Ann-Cathrin Hofer und Silke Argo

Das e:Med Systems Medicine Meeting 2021, 20.-22. September, war willkommener Treffpunkt der deutschen Systemmedizin *Community*. Neu waren insbesondere die *Community Sessions* und viele Varianten für *Networking* und Kommunikation, womit Vorschläge aus der e:Med *Community* umgesetzt wurden. Die erhebliche Bandbreite an Themen ist schon Standard und typisch für e:Med. Die Plattform e:Med@scoocs bot den besucherfreundlichen Konferenzraum für die 230 Teilnehmer.

Kreative Pianoklänge live sind nicht unbedingt das, was man auf einer digitalen Konferenz erwartet, umso begeisterter wurde Pianist **Michael Nickel** (Berlin) wahrgenommen, als Auftakt und Impuls zu der hochkarätigen Forschung.

Feierliche Eröffnungsworte fanden **Dr. Lorna Moll** (DLR PT Bonn) von Seiten des Förderers (BMBF) und **Dr. Matthia Karreman** (DKFZ und Uniklinik Heidelberg), Sprecherin des e:Med Projektkomitees mit motivierenden Schlaglichtern auf die deutsche Systemmedizin. Schon zum Start lockte e:Med 2021 neben den Forschenden der aktuellen e:Med Projekte und der e:Med Vorhaben der ersten Phase internationale Wissenschaftler, um ihre neuste Forschung vorzustellen, miteinander zu diskutieren und sich weiter zu vernetzen.

Professor Philip Rosenstiel (UKSH Kiel) stellte in dem Eröffnungsvortrag seine systemmedizinischen Arbeiten zur Behandlung von chronisch-entzündlichen Erkrankungen vor und erlaubte einen Ausblick auf den künftigen Nutzen und die weitere Entwicklung. Wie genetische Veränderungen in die Beurteilung von koronaren Herzerkrankungen mit einzubeziehen sind, erläuterte **Professor Heribert Schunkert** (DHZ München) der e:Med *Community* in seiner *Keynote Lecture* am

events

zweiten Tag der Konferenz. **Professor Julie George** (Universität zu Köln) zeigte mit translationalem Fokus aktuelle Arbeiten zur molekularen Landschaft von Lungentumoren auf.

In sieben thematischen Strängen von Methoden über Modellierung zur Translation begeisterten die Forschenden in über zwanzig Vorträgen durch aktuellste Ergebnisse, die interessiert diskutiert wurden. Auch Covid-19 war präsent, jedoch nicht als Beschränkung, sondern mit Vorträgen in der Session *„Systems Medicine of Covid-19“*.

In nur zwei Minuten Publikum durch kreativ und anschaulich dargestellte Forschung für ein Poster zu gewinnen: diese Möglichkeit nutzten die Präsentierenden der zwanzig *Flash Talks*. Die digitale Wahl durch die Teilnehmenden lohnte besonders für **Lea Zillich** (ZI Mannheim), sie erhielt den *Creative Award* 2021, einen Gutschein für einen NaWik *Workshop* im Wert von 275 Euro.

Die virtuelle Posterausstellung mit knapp fünfzig Postern bot Einblick in ganz aktuelle Arbeiten aller systemmedizinischen Bereiche in inhaltlicher und grafischer Vielfalt neben klassischen Postern durch Video- und Tonbeiträge. Der e:Med *Poster Award* 2021, dotiert mit je 150 EUR, wurde dieses Jahr an fünf Gewinner vergeben: **Dr. Ulrike Träger** (DKFZ Heidelberg), **Liza Vinhoven** (Uniklinik Göttingen), **Dr. Raik Otto** (HU Berlin), **Robert Müller** (TU Dresden) und **Zhijian Li** (RWTH Aachen).

Zusätzlich zum digitalen Verfolgen der Vorträge und virtuellen Flanieren entlang der Poster ermöglichte die digitale e:Med Plattform es den Forschenden, an *Networking Tables* Platz zu nehmen, um mit anderen Teilnehmenden in Kontakt zu kommen und sich auszutauschen. Anhand der individuellen Profile, spontanen *Video Meetings* und *Chats* gelang dies leicht und gezielt. Erstmalig wurde eine *Community Session* zur Zukunft der Systemmedizin auf Einladung des e:Med Projektkomitees durchgeführt. Das *Brainstorming* mit der gesamten *Community* hatte zum Ziel, die Perspektiven der Systemmedizin zu erörtern. Den Abschluss der Konferenz machte **Professor Rainer Spanagel** (ZI Mannheim), der Sprecher des e:Med Projektkomitees betonte die essentielle Förderung durch das BMBF und die großzügige Unterstützung durch die langjährigen Sponsoren der e:Med Meetings, life & brain und illumina, sowie durch die Partner de.NBI und NaWik.

Das e:Med Meeting 2021 war ein großer Erfolg, wie auch die Umfrage im Nachgang zeigte. Der Wunsch nach persönlichem Treffen bleibt jedoch lebendig. Aus diesem Grund freuen wir uns besonders, die e:Med Community nächstes Jahr vor Ort zu treffen: 28. – 30. November 2022 in Heidelberg.

Kontakt:

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

e:Med Meeting on Systems Medicine 2021, Ansichten von der digitalen Plattform mit Menu, Dashboard, Programm, Poster Session und Video (Sprecherin des e:Med Projektkomitees Dr. Matthia Karreman) (Foto: Quelle: e:Med/scoocs).

impressum

gesundhyte.de

Das Magazin für Digitale Gesundheit in Deutschland – Ausgabe 14, Dezember 2021

gesundhyte.de ist ein jährlich erscheinendes Magazin mit Informationen aus der deutschen Forschung aus den Bereichen Digitale Gesundheit und Systemmedizin. **ISSN 2702-2552**

Herausgeber:

gesundhyte.de wird herausgegeben vom Berlin Institute of Health at Charité und dem Projektträger Jülich.

Redaktion:

Chefredakteur: Prof. Dr. Roland Eils
(BIH at Charité – Universitätsmedizin Berlin)
Redaktionelle Koordination: Franziska Müller
(BIH at Charité – Universitätsmedizin Berlin)
Redaktion:
Dr. Silke Argo (e:Med), Melanie Bergs (PtJ), Dr. Stefanie Gehring (DLR-PT), Prof. Dr. Lars Kaderali (Universität Greifswald), Katharina Kalhoff (PtJ/BIH at Charité), 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 at Charité) und Dr. Gesa Terstiege (PtJ).

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

Aus systembiologie.de wird gesundhyte.de!

apops - Fotolia

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 at 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 at Charité und dem Bundesministerium für Bildung und Forschung (BMBF).

Die Redaktionsmitglieder von gesundhyte.de:

v.l.n.r. erste Reihe: Prof. Dr. Roland Eils (BIH at Charité), Franziska Müller (BIH at Charité), Dr. Silke Argo (e:Med), Melanie Bergs (PtJ), Dr. Stefanie Gehring (DLR-PT), **v.l.n.r. zweite Reihe:** Prof. Dr. Lars Kaderali (Universität Greifswald), Katharina Kalhoff (PtJ/BIH at Charité), 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 at Charité) 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, privat / Bergs

kontakt

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

Kommunikation und Marketing, Pressestelle
Ansprechpartner:
Dr. Stefanie Seltmann und Katharina Kalhoff
Anna-Louisa-Karsch-Str. 2; D-10178 Berlin
E-Mail: stefanie.seltmann@bih-charite.de, katharina.kalhoff@bih-charite.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, Melanie Bergs und Dr. Gesa Terstiege
D-52425 Jülich
E-Mail: y.pfeiffenschneider@fz-juelich.de, m.bergs@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

INTERNATIONAL CONFERENCE ON SYSTEMS BIOLOGY OF HUMAN DISEASE

NASHVILLE • JUNE 20 - 22, 2022

REGISTER @:
www.sbhdNashville.org

THE SBHD 2022 WILL BE HELD IN HYBRID MODE
YOU CAN CHOOSE ONLINE OR ONSITE PARTICIPATION
AT THE REGISTRATION.

LOCATION TBA

ABSTRACT SUBMISSION
DEADLINE TBA

LATE ABSTRACT SUBMISSION (POSTER ONLY)
DEADLINE TBA

ORGANIZED BY:

ROLAND EILS
BERLIN INSTITUTE OF HEALTH (BIH) /
CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN, GERMANY

VITO QUARANTA
VANDERBILT UNIVERSITY, NASHVILLE, USA

IRINA LEHMANN
BERLIN INSTITUTE OF HEALTH (BIH) /
CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN, GERMANY

SUPPORTED BY:

BIH Berlin Institute of Health @Charité

VANDERBILT UNIVERSITY

/vu/qsbc quantitative systems biology center

VANDERBILT UNIVERSITY MEDICAL CENTER

CHROMA®