



„Die Covid-Pandemie hat Anforderungen verschoben

Im Interview: Prof. Dr. med. Jörg Meerpohl, Direktor des Instituts für Evidenz in der Medizin, Uniklinik Freiburg, und von Cochrane Deutschland

Im Titelinterview mit „Monitor Versorgungsforschung“ beschreibt Prof. Dr. med. Jörg Meerpohl, der Direktor von Cochrane Deutschland, wie sich Cochrane ebenso wie die gesamte Methodik-Szene ständig weiterentwickelt. Hierbei spielt unter anderem der GRADE-Ansatz eine maßgebliche Rolle, der vor allem im Kontext von systematischen Übersichtsarbeiten herangezogen wird, um die beste verfügbare Evidenz für einen Endpunkt zu bewerten. Hieran arbeitet Meerpohl mit seinem Team auch im Fall der Covid-19-Pandemie, in der die verfügbare Studienevidenz gesichtet, strukturiert und bewertet wird. Gerade hier will und muss die gesamte Community – und auch Cochrane – schneller werden als bisher, denn – so Meerpohl – „die Covid-Pandemie hat die Anforderungen verschoben, und das ist auch gut so“.

>> Covid-19 hält die Welt und damit auch die Versorgungsforschung in Atem, auch wenn sie medial oft nicht so publikumswirksam in Erscheinung tritt wie die Virologie oder die Epidemiologie. All diese Fachdisziplinen haben das gleiche Ziel: die Schaffung bestmöglicher Evidenz. Gelingt das in diesen turbulenten Tagen?

Bedingt. Auf der einen Seite gibt es unheimlich viele wissenschaftliche Aktivitäten und entsprechend viele wissenschaftliche Publikationen zu Covid-19. Auf der anderen Seite sind diese Aktivitäten nicht gut aufeinander abgestimmt. Es gibt nur wenige große randomisierte Studien, dafür aber viele kleine Studien, häufig ohne Kontrollgruppe. Insofern ist die Evidenz, die wir für Antworten auf viele wichtige Fragen benötigen, noch nicht sehr gut.

Hat Covid-19 zu Anpassungen bei der Bewertung von Evidenz geführt?

Im Bereich der EbM bzw. EbHC hat sich das methodische Vorgehen durch die Corona-Pandemie nicht grundsätzlich verändert. Das heißt, wir haben an der methodischen Herangehensweise festgehalten, die sich über viele Jahrzehnte im Bereich der evidenzbasierten Medizin (EbM) und der evidenzbasierten Gesundheitsversorgung (EbHC) etabliert haben. Unser wichtigstes Instrument sind dabei systematische Übersichtsarbeiten, die für eine konkrete Fragestellung systematisch alle verfügbaren Studien identifizieren, deren Qualität bewerten, und diese dann entweder narrativ oder statistisch im Rahmen einer Metaanalyse zusammenfassen und eine sorgfältige Interpretation der Ergebnisse liefern.

Und wie geht man dabei mit Ergebnissen von Studien zu Covid-19 um, die nun gleich zu Dutzenden auf Pre-Print-Servern veröffentlicht werden?

Tatsächlich sehen wir in der Pandemie vermehrt Vorabveröffentlichungen von Studienergebnissen. Dies ist dem Bedürfnis geschuldet, der Wissenschaft und auch der Öffentlichkeit Ergebnisse möglichst rasch zur Verfügung zu stellen. Es ist allerdings zu beachten, dass diese Ergebnisse noch nicht den wissenschaftlichen Qualitätssicherungsprozess eines Peer Review durchlaufen haben; sie sollten also mit Vorsicht und immer im Kontext anderer Studien interpretiert werden. Ergebnisse von Pre-Print-Publikationen können durchaus auch Eingang in systematische Übersichtsarbeiten finden. Hier gilt aber, dass die wissenschaftliche Qualität sehr sorgfältig geprüft werden muss, und dass nach Publikation der Studienergebnisse in einer Fachzeitschrift mit Peer Review diese finalen Ergebnisse die Pre-Print-Ergebnisse in einem Update der systematischen Übersichtsarbeit ersetzen sollten. In der Pandemie arbeitet die Wissenschaft unter Hochdruck und muss nach Wegen suchen, den wissenschaftlichen Erkenntnisprozess zu beschleunigen.

Die Hauptherausforderung ist demnach die Zeit?

Natürlich braucht man in einer akuten Notlage wie jetzt Antworten so schnell wie irgend möglich. Die größte Herausforderung sehe ich aber darin, trotz des Zeitdruckes gute, d.h. richtige Antworten auf der Basis von hochwertigen Studien zu finden. Wir müssen aufpassen, dass das Tempo nicht zu sehr zu Lasten der Qualität geht.

Es ist niemandem geholfen, wenn wir voreilig falsche Schlussfolgerungen ziehen und unser Handeln daran ausrichten. Diese Abwägung von Geschwindigkeit und methodischer Sorgfalt betrifft übrigens nicht nur Primärstudien, sondern auch Evidenzsynthesen. Auch hier versuchen wir ohne Abstriche bei der Qualität aufs Tempo zu drücken. Bisher

benötigten klassische systematische Übersichtsarbeiten wie die von Cochrane ein, zwei oder manchmal gar drei Jahre bis zur Veröffentlichung. Dieses sehr sorgfältige und somit eben auch oft langwierige Vorgehen hat normalerweise durchaus seine Berechtigung, unter den Bedingungen einer Pandemie ist es aber nicht akzeptabel und für Gesellschaft und Politik kaum vermittelbar. Um die Erstellung von Evidenzsynthesen zu drängenden Fragen zu beschleunigen, erarbeitet eine eigene Methoden-Arbeitsgruppe von Cochrane schon seit 2015 vertretbare Anpassungen des methodischen Vorgehens für sogenannte Cochrane Rapid Reviews. Dabei gilt die Prämisse, möglichst wenig methodische Abstriche zu machen und dafür mit größerer Manpower und durch einen höheren Grad von Kooperation und Koordination der Teilaufgaben schneller zu Ergebnissen zu kommen. Es gibt darüber hinaus noch eine zweite, wichtige methodische Innovation, die aufgrund der enormen Forschungsdynamik in den Vordergrund gerückt ist: den Ansatz der „lebenden“ Evidenzsynthese. Solche Living-Systematic-Reviews werden in kurzen Abständen regelmäßig um die neueste Evidenz ergänzt und bleiben dadurch aktuell.

Oft wird aber auch, so zeigt ein Blick in einige Evidenzsynthesen zu Covid-19, der Blick geweitet.

Stimmt. Die klassische Herangehensweise einer systematischen Übersichtsarbeit beschränkte sich in der Vergangenheit häufig auf den paarweisen Vergleich von zwei Interventionen, also zum Beispiel: Ist A besser als B? In der Realität und gerade auch jetzt bei Covid-19 gibt es aber oft noch Option C, D, E und F. Dafür eignen sich die paarweisen Vergleiche in klassischen Systematischen Reviews nicht so gut. Um mehrere Interventionen gleichzeitig in einer systematischen Übersichtsarbeit miteinander vergleichen zu können, gibt es den methodischen Ansatz der Netzwerk-Metaanalyse. Allerdings ist der Aufwand hierfür auch größer.

Reicht dafür die vorhandene Manpower aus?

Nicht, wenn wir nur unsere Arbeitsgruppe in Freiburg am Institut für Evidenz in der Medizin und der Cochrane Deutschland Stiftung zur Verfügung hätten. Wir sind als Arbeitsgruppe aber Teil des internationalen Cochrane-Netzwerks und darüber hinaus auch in der methodischen Community der EbM und der Evidence-based Healthcare, kurz EbHC, vernetzt. Die Pandemie hat uns allen einen starken Push gegeben, noch mehr als früher in breit aufgestellten Kollaborationen zu denken und zu arbeiten. So stehen wir im regelmäßigen Austausch mit Partnern aus aller Welt, um gemeinsam Arbeiten durchzuführen und die nötigen Arbeitsschritte möglichst sinnvoll zu koordinieren und zu verteilen.

„Die größte Herausforderung sehe ich darin, trotz des Zeitdruckes gute, d.h. richtige Antworten auf der Basis von hochwertigen Studien zu finden.“

die
n“

Klappt das denn schon alles so ganz perfekt?

Wir sind auf einem guten Weg, d.h. in einzelnen Projekten funktioniert die Kollaboration gut.

Zum Beispiel?

Die WHO hat beispielsweise das „Evidence Collaborative on COVID-19 Network“ ins Leben gerufen, in dem inzwischen mehr als 80 Partner weltweit mitarbeiten. Dieses Netzwerk trifft sich einmal pro Monat virtuell, um sich auszutauschen, zum Beispiel welche Gruppe an welchem Thema arbeitet, und wie sich die Gruppen gegenseitig unterstützen können. Ein weiteres Beispiel ist das COVID-NMA-Projekt¹. Hier erarbeitet ein internationales Team, an dem auch wir beteiligt sind, lebende Evidenzsynthesen zu wichtigen pharmakologischen Fragestellungen.

Das wird sicher eine positive Langzeitfolge von Covid-19 sein. Denn diese Art der Zusammenarbeit wird hoffentlich mit der Pandemie nicht enden, sondern zu einem guten Teil bleiben.

Dafür gibt es zumindest gute Chancen. Doch auch schon vor Corona gab es – besonders im methodischen Bereich – eine große Notwendigkeit und Bereitschaft zu kooperieren. Viele der großen methodischen Projekte fanden auch schon vorher in Netzwerken statt, da die Fördermöglichkeiten für solche Projekte begrenzt sind. Trotzdem hege ich die Hoffnung, dass sich die in der jetzigen Situation etablierenden Kollaborationen verstetigen lassen.

Vielleicht sehen auch die Förderer immer mehr ein, wie wichtig methodische Forschung und Wissensmanagement sind.

Da tut sich schon einiges. So wurden im Rahmen des vom BMBF initiierten und durch die Charité koordinierten „Netzwerks Universitätsmedizin“ Fördermittel bereitgestellt. Das Institut für Evidenz in der Medizin (und Cochrane Deutschland) hat gemeinsam mit 20 anderen Universitätskliniken und externen internationalen Kollaborationspartnern ein Projekt CEOsys² beantragt, das seit Anfang September 2020 gefördert wird. In diesem Projekt sichten, strukturieren und bewerten wir umfänglich die verfügbare Studienevidenz zu Covid-19. Im Fokus stehen u.a. die Behandlung von Covid-19, aber auch Fragestellungen aus dem Public-Health-Bereich.

Da wird sicher schon gefragt: Wann seid ihr damit fertig?

Klar. Das gesamte Netzwerk ist großzügig mit Fördermitteln ausgestattet worden, und entsprechend hoch sind die Erwartungen, dass alles in kürzester Zeit aufgebaut und umgesetzt wird. Das ist eine große Herausforderung, weil es in den meisten Kliniken und Instituten neben der Krankenversorgung ja auch noch andere wissenschaftliche Aktivitäten gibt, die man nicht einfach auf Eis legen kann. Hinzu kommt, dass einige der klinischen Partner in unserem Projekt derzeit stark in die Versorgung von Covid-19-Patienten eingebunden sind, und somit wenig Zeit für wissenschaftliche Projekte bleibt.

ist es für Medien und Öffentlichkeit, aber auch für Wissenschaftler sehr schwierig, hier den Überblick zu behalten. Immer wieder werden einzelne, neue Studienergebnisse in Medien und in der Öffentlichkeit ohne eine angemessene Berücksichtigung der bereits bestehenden Studienerkenntnisse diskutiert. In anderen Situationen werden erste Ergebnisse einer kleinen Studie interpretiert, als ob hier bereits eine verlässliche und abschließende Antwort gefunden worden wäre. Es ist aus meiner Sicht zentral wichtig, dass einzelne Studien immer im Kontext anderer, ähnlicher Studien interpretiert werden.

Wie sollte denn mit dieser großen Masse an vielfältigen Forschungsdaten umgegangen werden?

Wir sehen in dieser Pandemie eine immense Dynamik im Bereich der Primärforschung und demzufolge auch auf der Ebene von systematischen Übersichtsarbeiten und Evidenzsynthesen. Was sich hier aufhäuft, ist kaum zu überblicken; vieles ist zudem redundant, nicht koordiniert und leider auch oft von nicht ausreichend guter Qualität. In der großen Masse finden sich so auch meist Studien oder gar Evidenzsynthesen, die die erhofften „Ergebnisse“ zeigen. Tatsächlich sieht man zur Zeit oft, wie wissenschaftliche Studien und Daten selektiv so ausgewählt werden, dass sie die eigene Argumentationslinie unterstützen. Findet diese Rosinenpickerei in den Medien oder

der Politik statt, ist das besonders problematisch. Genau diese Lücke wollen wir mit dem Projekt CEOsys – dem Covid-Evidenz-Ökosystem – im Rahmen des Netzwerks Universitätsmedizin schließen. Wir wollen damit einerseits Licht in den oft undurchdringlichen Dschungel aus Studienergebnissen bringen. Zum anderen bereiten wir in Kooperation mit der AWMF das dann

sortierte und bewertete Wissen in Form von evidenzbasierten Handlungsempfehlungen mit hohem Praxiswert auf. Diese sollen dann auch als Basis für auf unterschiedliche Zielgruppen angepasste Informationsmaterialien dienen – zum Beispiel für Politik und breite Öffentlichkeit. Das ist sicher ein sehr ambitioniertes Ziel, das wir zudem in ausgesprochen kurzer Zeit realisieren wollen.

Wie bewertet Cochrane bessere oder schlechtere Evidenz?

Zunächst ist es wichtig zu unterscheiden, um welche Art von Fragestellung es sich handelt. So ist der klassische RCT in der Evidenzhierarchie für therapeutische Fragen nach wie vor ganz oben angesiedelt. Das gilt so aber nicht für prognostische Fragestellungen.

Die Frage bestimmt nun einmal den notwendigen Studientyp.

Ja, von der Fragestellung hängt das optimale Studiendesign ab. Wenn wir im Bereich von Intervention, Prävention oder im Gesundheitsmanagement bleiben, gilt vereinfacht die bekannte Evidenzhierarchie – beginnend bei RCTs über prospektive Kohortenstudien bis hin zu Einzelfallberichten. Diese Evidenzhierarchie ist im Prinzip sinnvoll, weil wie wir alle wissen, gewisse Verzerrungsrisiken – Biases – in einem randomisierten Studiendesign besser in den Griff zu bekommen sind als mit anderen, nicht-randomisierten Studiendesigns. Es ist aber auch klar, dass es nicht nur große und gut durchgeführte RCTs gibt, sondern auch solche mit sehr kleinen Fallzahlen oder mangelhaftem Design. Zudem werden noch immer viele RCTs gar nicht oder nicht vollständig berichtet, sprich: Wir haben für Entscheidungen

„Ich hege die Hoffnung, dass sich die in der jetzigen Situation etablierenden Kollaborationen verstetigen lassen.“

Links

- 1: www.COVID-nma.com
- 2: www.COVID-evidenz.de
- 3: <https://de.gradeworkinggroup.org>

Wie ist derzeit Ihr Eindruck der tatsächlichen Nutzung der Empirie zu Covid-19 in den Medien und in der Politik?

Aufgrund der Masse und der Dynamik

unter Umständen nur Zugriff auf eine positiv selektierte Auswahl von Studienergebnissen. Das heißt: Simplizistisches Schwarz-Weiß-Denken ist hier nicht hilfreich. Darum verwendet Cochrane, wenn es um die Bewertung von Vertrauenswürdigkeit von Evidenz geht, den GRADE-Ansatz, der über viele Jahre international entwickelt wurde. Damit beschäftigen wir uns auch in Deutschland sehr intensiv und haben bereits im Jahr 2013 das Freiburger GRADE-Zentrum³ etabliert.

Was macht die Methodik des „Grading of Recommendations, Assessment, Development and Evaluation“-Ansatzes, kurz GRADE, aus?

Dieser methodische Ansatz zielt darauf ab, die Vertrauenswürdigkeit von Evidenz auf der Basis einer systematischen Übersichtsarbeit zu bewerten, also eben nicht mehr nur auf der Ebene einzelner Studien. Die Vertrauenswürdigkeit von Evidenz wird vielmehr aus einer Gesamtschau auf alle relevanten Studien abgeleitet. Hierbei wird nicht nur das Verzerrungsrisiko (Risiko für Bias) innerhalb der Studien berücksichtigt, sondern auch Aspekte wie die Inkonsistenz der Studienergebnisse oder auch das Problem des Publikationsbias. Wichtig ist auch, dass die Vertrauenswürdigkeit der Evidenz für jeden wichtigen Endpunkt separat bewertet wird. Ein Pool von Studien, wir nennen das dann Evidenzkörper, wird dadurch nicht per se als gut oder schlecht bewertet, vielmehr kann er Evidenz von hoher Vertrauenswürdigkeit für den einen, aber nur Evidenz von niedriger Vertrauenswürdigkeit für einen anderen Endpunkt liefern.

Kann man bei diesem methodischen Vorgehen auch die Ergebnisse einzelner Studien nutzen, obwohl diese vielleicht für sich gesehen gar nicht so gut gemacht sind?

Ja. GRADE bewertet die Evidenz – wie gesagt – auf der Basis einer systematischen Übersichtsarbeit, und zielt darauf ab, die beste, verfügbare Evidenz für einen Endpunkt zu nutzen. Insbesondere wenn es keine guten Studien gibt, werden dann eben die „mittelguten“ Studien berücksichtigt. Dieses Vorgehen bringt mit sich, dass manchmal für einige Endpunkte die Daten aus RCTs genutzt werden, und für andere, beispielsweise langfristige oder seltene unerwünschte Ereignisse, die Daten aus großen Kohortenstudien.

Wo bleibt dabei die Berücksichtigung von Biases und die Bewertung der internen Validität?

Diese wird natürlich berücksichtigt. Der erste Aspekt, der in der GRADE-Methodik bewertet wird, ist das Verzerrungspotenzial der Studien, also das Risiko für Bias. Doch auch Publikations- oder Disseminationsbias – ein übergeordnetes Phänomen des Studienpools –, sowie die Inkonsistenz und statistische Präzision der Ergebnisse werden bewertet. Wichtig ist es auch, einzuschätzen, wie direkt bzw. indirekt die Ergebnisse sind. Es ist entscheidend zu wissen, wie gut oder schlecht sich das, was man in der Literatur findet, auf die konkrete Fragestellung anwenden lässt. Dies ist der Bereich der externen Validität, also der Übertragbarkeit von Studienergebnissen. All diese Aspekte sind im GRADE-Ansatz zusammengeführt, der es so erlaubt, transparent und systematisch zwischen vertrauenswürdiger und weniger vertrauenswürdiger Evidenz zu unterscheiden.

Das ist eine Fortentwicklung der methodischen Standards zur Bewertung von Evidenz?

Das kann man so sagen. GRADE hat sich mittlerweile weltweit durchgesetzt, sei es für die Bewertung von Evidenz als Basis von Leitlinien oder in Cochrane Reviews. Mittlerweile ist GRADE das Herzstück eines jeden Cochrane Reviews. GRADE-Bewertungen erscheinen dort

besonders prominent in der „Summary of Findings“-Tabelle, welche die wichtigsten Ergebnisse zusammen mit deren Grad der Vertrauenswürdigkeit darstellt.

Wie sehen Sie mit diesem Bewertungsprinzip Studien zu Covid-19, in dem Wissen, dass deren Qualität sicher zum Teil unter dem Drang zur Schnelligkeit gelitten hat?

Studien haben schon immer Limitationen, nicht erst seit Corona. Nicht alle Studien werden optimal geplant und durchgeführt, so dass sich an vielen Stellen Risiken für Bias ergeben. Häufig beginnt für uns das Problem schon damit, dass viele Studien nicht ausreichend detailliert und präzise berichtet werden, d.h. es ist oft gar nicht klar, was genau wie gemacht wurde. Zum Beispiel geht aus Publikationen manchmal nicht eindeutig hervor, welche Endpunkte an wie vielen Probanden in welcher Art und Weise und zu welchem Zeitpunkt gemessen worden sind. Dazu kommt, dass z.B. auch oft unklar bleibt, ob der Outcome-Assessor – also derjenige, der den Endpunkt erhebt – verblindet wurde oder nicht. Die Herausforderungen bestehen also in Limitationen der Studienqualität, aber auch Limitationen des Berichtens von Studien.

Trotzdem spricht man von verlässlichen Informationen, dass A besser ist als B.

Zumindest wenn nach sorgfältiger Bewertung der Vertrauenswürdigkeit der Evidenz für den Vergleich A versus B sich dies so darstellt.

Wie stellt sich das nun bei Covid-19 dar?

Es gibt zu Covid-19 viele kleinere Studien, vor allem monozentrische Studien. Hinzu kommt, dass viele Arbeitsgruppen oft sehr gute Ideen für klinische Forschungsansätze hatten, in ihren Studien dann aber oft ihre Rekrutierungsziele nicht erreichen konnten – besonders über den Sommer hinweg, als die Fallzahlen weltweit zurückgingen. Derartige heterogen durchgeführte und heterogen geplante Studien, die in Teilen dann aufgrund von Rekrutierungsproblemen abgebrochen werden mussten, helfen natürlich nur bedingt. Überspitzt formuliert haben wir teilweise mehr „Studienkonkurrenz“ als Kooperation gesehen. Dazu kommt, dass viele Studien ähnliche Fragen adressieren und andere, womöglich wichtigere Fragen nicht untersucht wurden. Hier würde man sich eine – am besten weltweit – sehr viel koordiniertere Forschungsanstrengung wünschen. Auf diese Weise könnte man die wichtigsten Fragestellungen in wenigen, aber dafür sehr gut durchgeführten großen, gegebenenfalls multinationalen Studien besser und schneller beantworten.

Diesen Vorwurf wird sich die wissenschaftliche Community – auch die deutsche – anhören müssen.

Stimmt. Leider ist das bis auf wenige Positivbeispiele wie die Solidarity- oder Recovery-Trials noch nicht ausreichend gut gelungen.

In der Versorgungsforschung spielt – bei Covid-19 derzeit erst beginnend – die Auswertung von Routinedaten eine große Rolle. Wo würden Sie diese in Ihrer Rangfolge einordnen?

In der Diskussion um die Nutzung von Routine- und Registerdaten muss man zunächst einmal aufpassen, welche Begrifflichkeiten verwendet werden und was man genau unter Routine-, Register- oder Real-World-Data versteht. Die Qualität dieser Daten ist immens variabel.

Irgendwo in diesen vielen Daten und Datenquellen wird es doch gute geben.

Sicher gibt es die. Wichtig ist, dass Daten standardisiert und vollständig erhoben werden. Und wenn wir solche Daten nutzen wollen, um Aussagen bzgl. Therapieverfahren zu machen, brauchen wir auch entsprechend gute Erhebungen und Dokumentation zu möglichen Confoundern.

So wie es eben gute und schlechte RCT gibt, gibt es auch gute und schlechte Register.

Absolut. Die vereinfachende Dichotomisierung von „RCTs sind immer gut“ versus „Alles andere ist immer schlecht“ ist nicht sinnvoll. Stattdessen brauchen wir die Einordnung aller relevanten Ergebnisse nach unterschiedlichen Graden von Vertrauenswürdigkeit im Kontext einer gegebenen Fragestellung.

Kommen wir zurück zu Corona. Helfen in der Zeit der Pandemie die vielen aktuellen Studien oder verwirren sie eher?

Die vielen einzelnen Studien, die zum Teil dann ja auch in den Medien und der breiten Öffentlichkeit diskutiert werden, helfen vermutlich nicht viel. Vielmehr sehe ich hier die Gefahr der Verwirrung. Genau deswegen bemühen wir uns nicht nur um die Zusammenstellung der verfügbaren Evidenz, sondern auch um deren Bewertung und Interpretation. Eine transparente und nachvollziehbare Bewertung und Nutzung der vorhandenen Evidenz ist meines Erachtens nach zentral wichtig, um der Verunsicherung der Bevölkerung und auch der Politik entgegenzuwirken. Dazu gehört auch, Unsicherheit und Wissenslücken darzustellen. Das allein reicht aber noch nicht. Man braucht zusätzlich eine adäquate Kommunikations-Strategie, die auf der Basis der existierenden Evidenz nachvollziehbare Handlungsanweisungen und verständliche Informationsangebote anbietet. So kann man hoffentlich der Politik-Müdigkeit entgegenwirken, die man gerade in Zeiten der Pandemie gar nicht brauchen kann.

Natürlich wäre es für die Bürger schöner, wenn es eine klare, einheitliche und dauerhafte Linie gäbe. Doch funktioniert das nicht so einfach, wenn sich der Stand der wissenschaftlichen Ergebnisse dauernd ändert.

Jeder Wissenschaftsprozess ist nun einmal mit einer gewissen Unsicherheit und auch Dynamik verbunden, begleitet durch unterschiedliche Interpretationen und natürlich Anpassungen durch neu hinzukommende Erkenntnisse. Das sieht man zum Beispiel beim Thema der Mund-Nase-Masken. Zu Beginn der Corona-Krise wurde angenommen, dass lediglich FFP2-Masken schützen und Alltagsmasken keinerlei Wirkung zeigen würden. Heute gibt es vielfältige Hinweise, dass auch Alltagsmasken eine gewisse Schutzwirkung haben, ohne dass sich dieser Effekt im Einzelnen ganz genau quantifizieren lässt.

Und sicher zu Beginn der Corona-Pandemie auch einfach die Verfüg-



Prof. Dr. med. Jörg Meerpohl

ist seit 2018 Direktor des Instituts für Evidenz in der Medizin, Universitätsklinikum Freiburg, Direktor von Cochrane Deutschland, Wissenschaftlicher Vorstand der Cochrane Deutschland Stiftung und bereits seit 2013 Direktor des Freiburger GRADE Centers. Er ist ebenso Mitglied der STIKO (Ständige Impfkommision Deutschland). Er habilitierte sich im Bereich der experimentellen Medizin.

barkeit noch nicht gegeben war.

Das kam noch hinzu. Es ist aber auch ein Beispiel dafür, wie nach und nach mehr Evidenz zusammenkam, die inzwischen in der Gesamtschau dafür spricht, dass selbst einfache Mund-Nase-Masken nützlich sind.

Auch hier der Faktor Zeit: Es braucht eben seine Zeit, bis Ergebnisse vorliegen und diese verdichtet werden. Ist die Wissenschaft einfach für eine solche Situation zu langsam?

Das ist eine Kritik, die ich durchaus teile. In der Vergangenheit haben viele Cochrane Reviews zu lange gebraucht: Von der Definition einer relevanten Fragestellung bis zum fertigen Review dauerte es häufig ein bis zwei Jahre – und neueste Studien der letzten Monate sind dann oft noch nicht einmal berücksichtigt, weil die Literatursuche ja am Anfang des Prozesses steht. Keine Frage: Da muss die gesamte Community – und auch Cochrane – besser werden. Die Covid-Pandemie hat hier die Anforderungen verschoben, und das ist auch gut so.

Eine andere Kritik ist die, dass Cochrane lange teilweise nur RCTs berücksichtigt hat.

Cochrane entwickelt sich wie auch die gesamte Methodik-Szene stetig weiter. Im Rahmen des GRADE-Ansatzes wird für Fragen z. B. im Kontext der Leitlinienerstellung, die beste verfügbare Evidenz herangezogen. Wenn keine RCTs vorliegen, sind das dann meist (nicht-randomisierte) Beobachtungsstudien. Die Daten dieser Studien können wir in GRADE mit der gleichen Methodik bewerten. Natürlich hat Cochrane nach wie vor eine aus meiner Sicht auch berechnete klare Präferenz für RCT, doch wird dies nicht mehr ganz so schwarz-weiß gesehen wie früher. Hier hat sich Cochrane deutlich geöffnet – beschleunigt zuletzt auch durch die Pandemie. Im Bereich von Covid-19 wird nun sogar zum Teil mit Modellierungsstudien gearbeitet, ein

Novum innerhalb von Cochrane.

Gibt es eine explizit formulierte Forschungs-Strategie bei Cochrane?

Innerhalb von Cochrane wurden in den letzten Jahren Priorisierungs-Mechanismen etabliert. Man muss aber auch wissen, dass Cochrane vor mehr als 20 Jahren als eine Art wissenschaftliche Grassroots-Bewegung entstanden ist. Damit einher ging die Art und Weise, wie Themen ausgewählt, und dann auch in der Gemeinschaft abgearbeitet werden – nämlich meist forschungs- und interessengetrieben. Nach und nach hat sich mit der Weiterentwicklung von Cochrane eine engere Abstimmung mit Stakeholdern – auch auf übergeordneter Ebene mit der WHO – durchgesetzt. Dabei wird mehr und mehr darauf geachtet, dass Cochrane die drängendsten Fragen abarbeitet. Ich hoffe, dass das mittelfristig dazu führt, dass Cochrane Reviews noch relevanter für die Nutzer werden.

Worin sehen Sie den Nutzen der unter dem Stichwort Big Data zusammengefassten Methoden? Und wo die größten Gefahren für Fehlschlüsse und falsche Aussagen? Ihr Vorgänger im Amt, Professor Antes, war an der Stelle sehr kritisch.

Seine Hauptkritikpunkte halte ich für berechtigt: Wenn ich große Datenmengen habe und darin lange genug suche, werde ich jede Menge falsche Assoziationen – sogenannte „Spurious Correlations“ – finden. Da gibt es ganz herrlich abstruse Beispiele, etwa die Korrelation zwischen der Menge des Speiseeisverkaufs und der Anzahl von Menschen, die in Gewässern ertrinken. Oder der Kurve des Pro-Kopf-Käsekonsums in den USA, die über Jahre parallel zur Zahl der Todesfälle durch Verheddern im eigenen Bettlaken verläuft.

Weil man in großen Datenmengen auch viel Unsinniges finden kann.

Deshalb muss man gerade bei Big Data sehr sorgfältig in der Analyse und der Interpretation sein. Dennoch sehe ich in Big Data, also großen bis sehr großen Datensätzen, durchaus Potenzial für einen Nutzen. Schließlich sind mir auch bei Studien große Studien mit 10.000 Patienten lieber als kleine Studien mit 100. Generell gilt aber: Wenn die Datenqualität schlecht ist, haben auch sehr viele „schlechte“ Daten nur begrenztes Potenzial. Bei der Analyse sind dann besondere Sorgfalt und Kompetenz erforderlich. Dies trifft auch auf Big Data und Real-World-Data zu.

Deswegen heißt bezüglich der Datenqualität groß eben nicht gleich gut.

Stimmt. Primär ist die Qualität entscheidend. Entscheidend wichtig ist jedoch – wie gesagt – die sorgfältige Analyse dieser Daten, gepaart mit Fachwissen, um eben nicht fälschlicherweise irgendwelchen „Spurious Correlations“ auf den Leim zu gehen und daraus auf kausale Effekte zu schließen. <<

Danke für das Gespräch.

Das Interview führte MVF-Herausgeber Prof. Dr Reinhold Roski, Bearbeitung durch MVF-Chefredakteur Peter Stegmaier.

Zitationshinweis

Meerpohl, J.: Roski, R., Stegmaier, P.: „Covid-19 hat die Evidenz-Anforderungen verschoben“, in „Monitor Versorgungsforschung“ (01/21), S. 6-11; doi: <http://doi.org/10.24945/MVF.01.21.1866-0533.2269>

SNOMED CT: Eine gemeinsame Sprache für Gesundheitsdaten

>> Die vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Medizinformatik-Initiative (MII), an der alle Standorte der deutschen Hochschulmedizin und viele weitere Partner aus Forschung und Gesundheitswirtschaft beteiligt sind, hat 2020 die internationale Terminologie SNOMED CT in Deutschland eingeführt und deren Nutzung in ihrem Netzwerk ausgerollt. Gemäß Patientendaten-Schutz-Gesetz (PDSG) wird diese erfolgreiche Pilotierung nunmehr vorzeitig in die nationale E-Health-Strategie überführt: Seit dem 1. Januar 2021 ist Deutschland Mitglied bei SNOMED International, und auch Institutionen außerhalb des MII-Netzwerks können die erforderlichen Lizenzen innerhalb Deutschlands kostenlos nutzen. In diesem Zuge geht die Zuständigkeit für das die Lizenzen verwaltende SNOMED CT National Release Center ab Jahresbeginn von der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. auf das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) über.

„Wir freuen uns, dass die Nachfrage nach SNOMED CT auch außerhalb der medizinischen Forschung so groß geworden ist und wir damit bei der Förderung von Interoperabilität deutlich schneller in die flächendeckende Nutzung kommen. Die medizinische Forschung ist hier ihrer Aufgabe als Innovator gerecht geworden“, sagt Christian Luft, Staatssekretär im Bundesministerium für Bildung und Forschung (BMBF). „Wie wichtig die frühzeitige Investition in Dateninfrastrukturen für Forschung und Versorgung und insbesondere das Fördern von Standardisierung wie beispielsweise durch die Einführung von SNOMED CT ist, hat sich in besonderem Maße in der Corona-Krise 2020 gezeigt: Die notwendigen Festlegungen von standardisierten Datensätzen zur Covid-19-Forschung konnten bereits auf SNOMED CT und auf die weiteren Standardisierungsergebnisse der in der Medizinformatik-Initiative zusammengeschlossenen Expertinnen und Experten zurückgreifen.“ Das BMBF fördert die Initiative in der laufenden Aufbauphase von 2018 bis 2022 mit rund 160 Millionen Euro.

Die Terminologie SNOMED CT verwandelt medizinische Sachverhalte und Fachbegriffe, für die es weltweit unterschiedliche Bezeichnungen und Maßstäbe gibt, in sprachunabhängig unmissverständliche und international eindeutige Zahlencodes, also in eine „einheitliche Sprache“, die von Computern gelesen und verarbeitet werden kann. Dies

schafft die Voraussetzungen für intelligente Datenanalysen in Gesundheitsversorgung, medizinischer Forschung und Gesundheitsstatistik. Direkte Vorteile für die Gesundheitsversorgung ergeben sich zum Beispiel, wenn klinische Parameter oder Medikationsdaten standort- und sogar länderübergreifend einheitlich ausgetauscht werden. Zudem können die mit SNOMED CT einheitlich beschriebenen medizinischen Inhalte auch für die Entwicklung von medizinischen Entscheidungshilfen verwendet werden, die auf Künstlicher Intelligenz basieren.

Das Nationale Steuerungsgremium (NSG) der MII hat hierzu bereits während der Konzeptphase zum MII-Aufbau im Dezember 2016 die Festlegung getroffen, dass für eine auch international anschlussfähige terminologische Normierung der Routinedaten für Forschungszwecke die Nutzung des international anerkannten Terminologiestandards SNOMED CT unerlässlich sei. SNOMED CT wird herausgegeben von der Non-Profit-Organisation SNOMED International; Nutzungs- und Mitwirkungsrechte erwirbt man durch nationale Mitgliedschaften in dieser Organisation. Bis dato hatte Deutschland diese Mitgliedschaft nicht vollzogen. Das BMBF hat daher mit Unterstützung der MII-Koordinationsstelle mit SNOMED International eine Sondervereinbarung für eine bis zu dreijährige Pilotphase ausgehandelt. Auf dieser Basis ist seit Mitte März 2020 SNOMED CT über die MII erstmals in Deutschland verfügbar. Als Teilnehmer der MII haben nahezu alle deutschen Universitätskliniken und weitere Partner der Initiative (Universitäten, Hochschulen, Forschungseinrichtungen und Industrieunternehmen) eine SNOMED CT-Lizenz erhalten. Die Lizenzgebühren trug in dieser Pilotphase das BMBF; die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Sitz der MII-Koordinationsstelle, fungierte als National Release Center.

Mit dem durch das Patientendaten-Schutz-Gesetz vom 14.10.2020 gesetzlich geregelten Vollzug der nationalen Mitgliedschaft Deutschlands bei SNOMED International ist diese vorbereitende Pilotphase vorzeitig zum 01.01.2021 beendet worden. Die Funktion des National Release Centers übernimmt das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Bundesinstitut im Geschäftsbereich des Bundesministeriums für Gesundheit (BMG), welches nunmehr die Sublizenzen für deutsche SNOMED CT-Nutzer herausgeben wird. <<