

Interview mit Prof. Dr. Gerd Antes: Digitalisierung und Evidenz – Hat die klassische Methodik in der Wissenschaft noch eine Zukunft?

Das Interview führte Dr. Matthias Fellhauer

Die *lege artis* durchgeführte klinische Studie gilt nach wie vor als Goldstandard bei der Generierung von Evidenz in der Medizin. Mit der zunehmenden Digitalisierung und der Verfügbarkeit riesiger Datenmengen stellt sich inzwischen aber die Frage, ob die Methoden der klassischen Wissenschaft nicht ebenso gut oder sogar noch besser von einer intelligenten maschinellen Datenanalyse abgelöst werden können. Der Freiburger Mathematiker, Methodenwissenschaftler und langjährige Direktor der Cochrane Deutschland Stiftung, Professor Gerd Antes, vermisst den kritischen Diskurs über diese Thematik.

Herr Professor Antes, in einem Ihrer Vorträge stellten Sie fest, dass die Evidenz beim Beweis des Nutzens in der Medizin zunehmend durch die permanente Wiederholung des Unbewiesenen abgelöst werde. Sehen Sie diese Gefahr auch beim Nachweis der Wirksamkeit von Arzneimitteln? Gibt es dafür schon konkrete Anhaltspunkte?

Diese Aussage bezog sich auf die gebetsmühlenartig in den Raum gestellten Aussichten auf eine bessere Welt durch die Digitalisierung durch die Analyse von riesigen Datenmengen und die damit untrennbar verbundene künstliche Intelligenz. Diese auf allen Kanälen laufende Dauerberiesung erfolgt tatsächlich nur durch ständige Wiederholung, ohne begründete und nachvollziehbare Belege dafür zu liefern, dass die versprochenen Auswirkungen eintreten werden.

Der Wirksamkeitsnachweis von Arzneimitteln ist nur ein kleiner Ausschnitt unserer Welt. Da wir jedoch mitten in der „digitalen Transformation“ unserer Gesellschaft stecken, ist selbstverständlich und geradezu zwangsläufig auch die Arzneimittelbewertung davon betroffen.

Die US-amerikanische und europäische Arzneimittelbehörde (FDA und EMA) haben Initiativen gestartet, um ihre Infrastruktur hinsichtlich der Gewinnung von Erkenntnissen aus Big Data zu erweitern. Als konkrete Anhaltspunkte lassen sich Artikel, regulatorische Anforderungen und Pläne zur Nutzung von Big Data nennen, zur Orientierung vor allem von der US-amerikanischen und europäischen Arzneimittelbehörde (FDA und EMA). Am deut-

lichsten wird das durch den 48 Seiten langen Bericht der HMA-EMA Joint Big Data Taskforce vom Februar 2019. Dort wird klar festgestellt, dass Big Data ein weithin benutzter Begriff ist, es jedoch keine allseitig akzeptierte Definition gibt. Der üblichen Dateneuphorie wird in dem Bericht eine nüchterne Betrachtung gegenübergestellt, die das mögliche Potential benennt, aber sehr differenziert die besondere Sorgfalt im regulatorischen Prozess betont und fordert. Die Unverzichtbarkeit von Validierung und Qualität in den Prozessen steht in deutlichem Gegensatz zum allgegenwärtigen Hype, der diese Notwendigkeit negiert und vermittelt, dass die Masse an Daten dank künstlicher Intelligenz keiner besonderen Anstrengungen mehr bedarf.

Auch bei der FDA sind aus der zweiten Hälfte von 2018 und aus diesem Jahr Papiere sichtbar, die die intensive Beschäftigung mit dem Thema zeigen. Auch hier scheint jedoch keine pure Begeisterung zu herrschen. So beklagt das Center for Drug Evaluation and Research (CDER) der FDA das Bombardement durch Big Data, das die

Prof. Dr. Gerd Antes, Institut für Didaktik und Ausbildungsforschung in der Medizin (Senior Fellow), Klinikum der Universität München, LMU München, E-Mail: antes@cochrane.de

Dr. Matthias Fellhauer, Apotheke der Schwarzwald Baar-Klinikum Villingen-Schwenningen GmbH, Klinikstraße 11, 78052 Villingen-Schwenningen, E-Mail: herausgeber-kph@adka.de

Entscheidungsfindung behindert und neue Strukturen erfordert.

Strukturell und methodisch bedeutet das die Erweiterung der Forderung nach Evidenz jenseits von RCTs durch Real World Data (RWA); durch Big Data bewegen wir uns nun – bildlich gesprochen – in RWA XXL.

Einer Ihrer wesentlichen Kritikpunkte an Big Data ist die unkritische Gleichsetzung des kausalen Zusammenhangs mit der statistischen Korrelation. Wo sollten Apothekerinnen und Apotheker besonders auf der Hut sein und weiterhin kritisch differenzieren?

Einzelne Apothekerinnen und Apotheker sind – genauso wie alle Gesundheitsberufe sowie Patienten und Angehörige – sehr gefordert und größtenteils wohl auch überfordert, individuell kritisch zu differenzieren. Diese kritische Differenzierung muss institutionell geleistet werden und den Gesundheitsprofessionen und Patienten als Service geliefert werden. Genau diese Unterstützung wird ja vom Arzneimittelzulassungsprozess durch den Nachweis der kausalen Ursache für maximale Wirksamkeit bei minimalem Risiko und Schaden geliefert.

Grundsätzlich sollte von allen Beteiligten verstanden werden, dass Korrelationen, also das gleichzeitige Auftreten von zwei Beobachtungen, nicht bedeuten, dass das eine Ereignis das andere ursächlich ausgelöst hat. Ein sehr einfaches Beispiel, das seit Generationen den Medizinstudentinnen und -studenten nahegebracht wird: Die Abnahme der Storchenpopulation und der Rückgang der Geburtenzahlen liefern nicht der Beweis dafür, dass der Klapperstorch die Babys bringt. Dieses Beispiel mag trivial erscheinen, es gibt jedoch eine Fülle von ähnlichen Zusammenhängen, die nicht so leicht zu durchschauen sind. Der falsche Eindruck wird erzeugt durch andere Einflüsse im Hintergrund, sogenannte Störvariablen bzw. Confounder, die beide Variablen gleichzeitig beeinflussen. In diesem Beispiel wäre es die zeitliche Dimension der Entwicklung der Zivilisation. Diese Einflüsse können beliebig komplex sein und sind die große Herausforderung bei der Datenanalyse, wenn es darum geht, die Variablen zu identifizieren, die als Stellschrauben für Veränderungen dienen können.

Dass eine weitgehende technische (und damit zwangsläufig datenbasierte) Unterstützung ein Qualitätsgewinn sein kann, ist unbestritten. So ist z. B. die Unit-Dose-Technologie ein wesentlicher Bestandteil des von der ADKA propagierten Closed-Loop-Prozesses bei der stationären Arzneimittelversorgung. Im Gegensatz dazu werden die aus der statistischen Analyse großer Datenmengen gewonnenen Erkenntnisse von ihren Kritikern auch als „künstliche

Dummheit“ bezeichnet. Wie kann Ihrer Meinung nach die Gratwanderung zwischen technischer Unterstützung und künstlicher Dummheit gelingen?

Ganz einfach, durch die Rückkehr zur Bewertung des Patientennutzens und durch die Rückbesinnung auf den Qualitätsbegriff als alles überragenden Parameter für die Strukturen und Prozesse, die heute mit marketinggetriebenem Hype und Versprechungen überschwemmt werden. Wir erleben eine von Schlagworten dominierte, völlig überzogene und systematisch übertriebene Nutzenerwartung bei gleichzeitiger Ignoranz von Risiken und möglichen Schäden sowie der Kosten. Damit wird von der seit den 1960er Jahren wissenschaftlich entwickelten und fortlaufend validierten Technikfolgenabschätzung (allgemein TA, speziell für die Medizin HTA = Health Technology Assessment) von den drei Dimensionen Nutzen – Risiko/Schaden – Kosten nur die erste Dimension genutzt und gleichzeitig werden die letzten beiden vorsätzlich und konsequent ignoriert.

Die Gratwanderung zwischen technischer Unterstützung und künstlicher Dummheit wird erschwert oder sogar verhindert durch eine verwirrte und verwirrende Diskussion, in der weder Begriffe eindeutig definiert werden noch Ziele und die Kriterien für deren Erreichung benannt werden.

Das hier genannte Beispiel der Closed-Loop-Arzneimittelversorgung ist eine lokale Lösung, die hohe Ansprüche an die Sicherheit stellt. Sie braucht jedoch weder eine Big-Data-Analyse über Institutionen, Bundesländer- und nationale Grenzen hinweg noch den Zugang zu persönlichen genetischen Daten für Forschungszwecke. Diese und weitere Ziele werden jedoch undifferenziert in einem Atemzug genannt und implizit mit der digitalen Transformation begründet, in der unsere Gesellschaft sich ja angeblich befindet.

Angenommen, man ist sich der Problematik des kausalen Zusammenhangs versus der statistischen Korrelation bewusst: Haben wir in Zeiten der globalen Digitalisierung denn überhaupt noch Chancen, in unserer fachlichen Arbeit zu differenzieren? Müssen wir z. B. bei der Bewertung von Studien für eine Therapieleitlinie jetzt auch die Qualität der Durchführung kritisch hinterfragen?

Ja, allerdings, unbedingt. Das galt jedoch immer schon. Die Lagerbildung auf der einen Seite „RCT = Evidenzbasierte Medizin“ vs. „Beobachtungsstudien = Versorgungsforschung“ auf der anderen Seite ist eine fundamentale, speziell deutsche Fehlentwicklung. Ein chronischer Fehler darin ist die Beschränkung auf die Sichtweise, die RCT nur als Studiendesign zu sehen und für die Qualitätsbewertung diverse Aspekte der Durchführung zu ignorieren. Als Beispiel kann hier die Beschreibung der Rando-

misierung genannt werden. Es gibt eine stattliche Menge an Literatur aus den letzten Jahrzehnten, die die methodischen Schlampereien wie auch die vorsätzlichen Täuschungsmanöver bezüglich der Behandlungsallokation ausführlich beschreibt. Diese Arbeiten waren Motivation für die Entwicklung des Equator Networks, das mit seinen Reporting Guidelines for main study types eine unschätzbare Arbeit geleistet hat.

Generell ist für die Nutzung der Studienergebnisse aus der Literatur notwendig, einen detaillierten Methodenteil zu formulieren, der die Qualität der Studie und damit das Vertrauen in die Ergebnisse einschätzen lässt. Speziell für die Randomisierung heißt das, dass nicht nur festgehalten werden muss, dass die Behandlungszuteilung randomisiert erfolgte, sondern wie. Bei hochwertigen Studien heißt das heute üblicherweise Telefonrandomisierung. Diese Forderung ist einer der zentralen Qualitätsparameter und trotzdem wird sie auch heute noch erstaunlich oft nicht angemessen befolgt.

Kausale Ursachen aus Beobachtungsstudien zu ermitteln, ist mehr als tückisch, da Störvariablen bzw. Confounder zu massiven Täuschungen führen können und insbesondere nicht messbare oder nicht gemessene Einflussgrößen ein Eigenleben mit sich regelmäßig widersprechenden Studienergebnissen führen. Genau diese schädlichen Auswirkungen werden durch die Randomisierung verhindert.

Auf die Spitze getrieben wird die Nutzung von Real-World-Data nun durch die Analyse von Big Data durch Deep Learning, neuronale Netze und Machine Learning. Sie scheinen gegen diese Gefahren und chronischen Fehler immun zu machen. Unbegrenzte Datenmengen scheinen Theorie und intensive methodische Anstrengungen in der Empirie überflüssig zu machen. Das ist natürlich grob falsch, jedoch Grundelement in den grassierenden Versprechungen.

Um hier auch über verschiedene Studientypen Anhalt für das Vertrauen in die Ergebnisse zu erhalten, hat sich die Bewertung des Risk of Bias zu einem unverzichtbaren Instrument entwickelt, das in allen modernen Anleitungen eine unverzichtbare Rolle spielt. Das Biasrisiko kann nicht null werden, die Minimierung sollte jedoch zentrales Ziel sein. Nur mit methodisch sauberer Arbeit ist maximale Aussagesicherheit vor allem auch für die Frage der Kausalität zu erreichen.

Die Digitalisierung ist inzwischen ein Überbegriff für die unterschiedlichsten IT-Technologien von der elektronischen Patientenakte über die Gesundheits-App bis zur Datenbank in einem Gerät der Medizintechnik. Selten wird der Begriff aber im Zusammenhang mit Qualitätsstandards gebraucht. Täuscht dieser Eindruck, oder läuft der

Patient tatsächlich Gefahr, im Zuge der Digitalisierung einen Qualitätsverlust bei seiner Behandlung zu erleiden?

Diese Gefahr ist sehr groß und wird in der Gesundheitsforschung und -versorgung nicht angemessen erkannt und thematisiert. Der Qualitätsbegriff, die Fehlerkultur, um damit adäquat umzugehen sowie das Grundprinzip in der Medizin, Entscheidungen unter Unsicherheit zu fällen, sind weitgehend aus der Diskussion verschwunden. Die in den letzten Jahrzehnten entwickelten und implementierten Elemente von Good Scientific Practice, die sich alle um den Qualitätsbegriff ranken, spielen keine sichtbare Rolle in der neuen digitalisierten Welt. Das Verschwinden ist deutlich zu beobachten, ebenso wie die fehlende Diskussion um die Qualität der Methoden.

Der marketinggetriebene Hype dient nicht nur dem Vertrieb, sondern ist auch in die relevanten Bereiche der Wissenschaft eingedrungen und ist unvereinbar mit den Grundprinzipien der Evidenzbasierten Medizin (EbM), die ja gerade die Integration hochwertiger Methodik in den klinischen Alltag als zentrales Ziel hat und dazu in den letzten Jahrzehnten prägend gewirkt hat. Auf den Punkt gebracht könnte man sagen, dass wir gegenwärtig die Ablösung einer altmodischen Welt, die an Prinzipien mit Qualität als Focus klebt, durch eine neue, datenorientierte Welt beobachten, in der die Rolle von Qualität durch Menge, Geschwindigkeit und Verzicht auf Theorie übernommen wird. Es ist sicherlich nicht überraschend, um nicht zu sagen vorhersehbar, wenn daraus ein Qualitätsverlust bei der Behandlung folgt. Besondere Bedeutung kommt dem 2008 deklarierten Verzicht auf Theorie zu. Der daraus abgeleitete Paradigmenwechsel der Wissenschaftlichkeit ist äußerst kritisch zu sehen, sind die großen Durchbrüche und wissenschaftlichen Entwicklungen doch aus dem funktionierenden Wechselspiel von Theorie und Empirie entstanden. Es gibt keine belegbaren und überzeugenden Gründe, dass der Theorieverzicht die Qualität verbessert oder auch nur erhält. Jüngste, tiefgründig theoretische Arbeiten zeigen sogar, dass selbst mehr richtige Daten nicht zwangsläufig mehr Wissen bedeuten und sie sogar den Erkenntnisprozess massiv beschädigt können.

Sehen Sie Bereiche, wo die Evidenzbasierte Medizin von Big Data profitieren könnte?

Praktisch überall, da die Evidenzbasierte Medizin ja die systematische Integration von Studienergebnissen, also Wissen aus Daten, als Leitziel verfolgt und damit Daten Teil des Fundaments sind. Entscheidend ist, ob es gelingt, die methodischen Ansprüche der EbM mit den Sichtweisen innerhalb der Datenwissenschaften zu harmonisieren, das Profitieren also wechselseitig erfolgen sollte. Ich bin da einigermaßen skeptisch. Die Entwicklung und Kommuni-

kation in Silos und Echokammern scheinen diese Welten durch Barrieren zu trennen, die gut funktionieren.

Nur aus der EbM-Perspektive betrachtet, sind die Bedingungen einfach zu formulieren: Rigide Anwendung der methodischen Qualitätswerkzeuge für Studien. Wie oben teilweise beschrieben, Risk of Bias als wesentliche Orientierung für die Qualitätsbewertung von Studienergebnissen, Transparenz und Integrität in der Planung, Durchführung und Publikation von Studien, vollständige und informative Methodikbeschreibung, Datenverfügbarkeit für systematische Übersichtsarbeiten und Reanalysen und weitere in der Welt der Studiendurchführung etablierte Qualitätsparameter. Insgesamt bedeutet das ein eher konservatives Vorgehen, im Gegensatz zur Data Science Welt: Neu ist nicht immer besser, schneller ist nicht immer besser, neues Wissen immer im Kontext des vorhandenen Wissens. Diese Prinzipien stehen in fundamentalem Widerspruch zu den überstrapazierten Hype-Wörtern Innovation, Disruption oder sogar Sprunginnovation. Das Bundesministerium für Bildung und Forschung (BMBF) leistet für den Blick in die Zukunft Nachhilfe: Es schafft die Grundlagen für revolutionäre Innovationen wie das Smartphone. Solche „disruptiven“ Innovationen fördert das BMBF mit der „Agentur für Sprunginnovationen“.

Es ist schwierig, zu sehen, wie diese Welten zueinander kommen können.

Sie fordern einen breiten gesellschaftlichen und kritischen Diskurs vor einer mehr und mehr ausufernden Nutzung von Big Data im Gesundheitswesen – vor allem auch, um ein gesundes Gleichgewicht zwischen Ethik und Technik zu bewahren. Welchen Rat geben Sie den Krankenhausapothekern, um diesen Diskurs in ihrem fachlichen Umfeld zu führen?

Ethische Aspekte scheinen oft wie ein fünftes Rad am Wagen neben Entwicklungen herzulaufen und oft auch nur als Feigenblatt zu dienen. Das ist auch gegenwärtig zu beobachten. Gleichzeitig sehen wir eine bedeutende Zunahme an Diskussionen um die Ethik der Digitalisierung

und unter anderem neugeschaffene Professuren für dieses Themengebiet. Das ist extrem wichtig, wird doch das Menschenbild massiv beeinflusst und verändert. Die Schaffung eines neuen Menschenbildes und die Automatisierung von Entscheidungen im therapeutischen und sozialen Bereich schaffen Abhängigkeiten, deren Ausmaß kaum einzuschätzen ist. Speziell für Krankenhausapotheker scheinen mir zwei Aspekte besonders relevant zu sein.

Einmal wird die Datensicherheit eins der am schnellsten wachsenden Themen sein. Die Diskussion wird gegenwärtig fast ausschließlich um Datenschutz geführt, Schutz vor Diebstahl und damit möglicher individueller Erpressung oder auch Verkauf der Daten. Diskreditiert wird dieser Schutz durch sinnleere Sprüche wie „Datenschutz ist für Gesunde“. Oft völlig übersehen – und viel bedrohlicher – ist jedoch die mangelnde Datensicherheit. Gemeint ist der Schutz vor Zerstörung – oder noch perfider speziell für Patientendaten – vor Veränderung von Daten, die selbst automatisiert durch Robots möglich ist. Wenn auch nur ein kleiner Teil dessen, was als Internet der Dinge für private Haushalt angepriesen wird, in die Krankenhäuser hereingelassen wird, schafft man damit ein Hacker-Eldorado, das unvermeidlich auch die Arzneimittelversorgung in den Krankenhäusern betrifft. Hier wachsam zu sein, ist sicherlich für Krankenhausapotheker extrem relevant.

Der andere wesentliche Aspekt ist die mit der Digitalisierung verbundene Entmenschlichung. Empathie-Apps und lächelnde Computermännchen sind keine geheimen Gedanken mehr. Welche Schäden die Abschaffung des Handauflegens, des menschlichen Blicks in die Augen und des Anlächelns für Patienten bedeutet, ist nicht einmal ansatzweise abzuschätzen. Auch hier ist der Krankenhausapotheker massiv vom Gleichgewicht zwischen Ethik und Technik betroffen und sollte eine sehr aktive Rolle bei der Gestaltung des Versorgungsalltags spielen. Die Ethik spielt eine entscheidende Rolle, ob die Technik Knecht oder Herrscher sein wird. Diese Entwicklung betrifft jeden, sodass Beteiligung und Einflussnahme selbstverständlich sein sollte, auch für Krankenhausapotheker.