

Künstliche Intelligenz mit der Radiologie als Vorreiter für Super-Diagnostics: Ein Essay

Artificial Intelligence with Radiology as a Trailblazer for Super-Diagnostics: An Essay



Prof. Dr. med. Michael Forsting

Korrespondenzadresse

Prof. Dr. med. Michael Forsting

Direktor des Instituts für diagnostische und interventionelle Radiologie und Neuroradiologie, Medizinischer Direktor der Zentralen IT der Universitätsmedizin Essen
Universitätsklinikum Essen, Hufelandstraße 55, 45147 Essen
michael.forsting@uk-essen.de

Bibliografie

DOI <https://doi.org/10.1055/a-0808-7772>

Online-Publikation: 2020

Fortschr Röntgenstr 2019; 191: 73–78

© Georg Thieme Verlag KG, Stuttgart · New York

ISSN 1438-9029

Gucken wir auf die Medizin, sehen wir im Moment 2 große Trends. Der eine Trend heißt Digitalisierung: In der Radiologie und in der klassischen Labormedizin ist der Alltag schon lange digital. Die Patientenakte, die Pathologie, die Mikrobiologie und die Virologie haben unterschiedliche Reifegrade in der Digitalisierung, aber es ist absehbar, dass die komplette Digitalisierung nicht mehr lange auf sich warten lässt.

Der zweite große Trend in der Medizin ist die „Personalisierung“: Gemeint ist damit bis jetzt vor allem eine maßgeschneiderte Pharmakotherapie, bei der die individuelle physiologische Konstitution und auch molekularbiologische Konstellationen berücksichtigt werden.

Guckt man außerhalb der Medizin, gibt es einen weiteren Megatrend: Künstliche Intelligenz (KI). Aus Gründen, die später erörtert werden, ist dieser Trend in der Medizin noch nicht richtig angekommen, aber in der industriellen Fertigung und bei der Verarbeitung großer Daten nicht mehr wegzudenken. Facebook, Amazon und Google würden ohne KI nicht funktionieren.

Die Herausforderung für die Medizin ist jetzt, die Trends Digitalisierung und KI so einzusetzen, dass personalisierte Medizin Alltag wird. Ohne KI wird es keine personalisierte Medizin geben: Die Datenmengen sind so immens, dass es mit „Kopfrechnen“ – dem persönlichen Wissen einiger Experten – nicht mehr funktionieren wird.

Warum ein Essay?

Eigentlich erwarten Sie eine wissenschaftliche Übersichtsarbeit, die die Möglichkeiten von KI in der Radiologie und in der Medizin beschreibt. Das von mir gewählte Essay-Format hat damit zu tun, dass man schwarze Schwäne naturwissenschaftlich nicht überzeugend darstellen kann. Schwarze Schwäne sind das Symbol für

Dinge, mit denen wir nicht rechnen (wenn Sie nach diesem Essay tiefer in das schwarze-Schwan-Problem einsteigen wollen, lesen Sie das gleichnamige Buch von TALEB). Nehmen wir als Beispiel für einen schwarzen Schwan die Terrorattacken des 11. September 2001: Jeder Sicherheitsexperte hätte dieses Terrorszenario in einem wissenschaftlichen Aufsatz vorhersagen können. Die Möglichkeit einer Flugzeugentführung war gegeben, es war möglich, dass ein Entführer eine Pilotenlizenz haben könnte und technisch war es möglich, die Transponderfunktion der Flugzeuge abzuschalten, sodass Flughöhe, Kennung und Geschwindigkeit nicht mehr erfasst wurden. Selbst wenn ein Sicherheitsexperte dieses Szenario vorausgesehen hätte, sein Aufsatz wäre sehr wahrscheinlich nicht zur Publikation angenommen worden. Undenkbares Denken bringt das Weltbild zu sehr durcheinander.

Das Auto ist auch ein schwarzer Schwan gewesen. Die Strecke von Mannheim nach Pforzheim wäre mit dem Pferd viel schneller bewältigt worden und man hätte sich das Theater mit dem Kauf von Benzin in einer Apotheke sparen können. Ein Aufsatz zum potenziellen Nutzen des Autos wäre von den Reviewern als unwahrscheinliche Perspektive abgelehnt worden. Die essayistische Tat von Berta Benz war die wichtigere Publikation.

Andere schwarze Schwäne sind der Niedergang der Telefonsparte von Siemens, das Verschwinden der Schreibmaschinenhersteller vom Markt, der Ausstieg aus der Atomenergie getrieben von der CDU, der Aufstieg von Google und der Niedergang vom Blackberry. Was können wir daraus lernen? Es gibt schwarze Schwäne – gar nicht so selten –, sie werden von der klassischen Wissenschaft immer wieder unterschätzt und sie können die Welt gehörig durcheinanderbringen.

Gut, aber wir sind ja in der Medizin. Schwarze Schwäne sind kein Tabu. Naja, nehmen wir das Magen- und Duodenal-

Geschwür. Bis Mitte der 1980er-Jahre war dies eine psychosomatische Erkrankung. Man hatte eine Hypothese entwickelt und so lange nach Bestätigung gesucht, bis alle es glaubten. Bis 2 australische Kollegen ihre Theorie zur bakteriellen Genese des Ulkus in einem Selbstversuch bewiesen, völlig unmedizinisch, essayistisch gewissermaßen. Oder nehmen wir die Anorexia nervosa, eine sehr monomorph verlaufende Erkrankung, deren Ursache nach heutigem Verständnis ein Zusammenwirken von psychischen und sozialen Faktoren sein soll. Irgendwann kommen wir mithilfe hypothesenfreier Forschung und KI auf die wirkliche Ursache.

Am Ende dieses Abschnitts sollten Sie mental bereit sein, die Existenz von schwarzen Schwänen anzuerkennen und für denkbar halten, dass schwarze Schwäne auch die Medizin gehörig verändern können. Und Sie sollten bereit sein, einen Essay in einer wissenschaftlichen Zeitschrift zu akzeptieren.

Was ist der Unterschied zwischen CAD und KI?

Die Radiologie hat schon mit der Digitalisierung des Faches begonnen, Algorithmen zu entwickeln, die die Diagnostik einfacher machen sollten, Computer-aided diagnosis (CAD) ist das Stichwort. Und viele Skeptiker der KI referieren auf den fehlenden Durchbruch der CAD-Systeme als Beleg für ihre berechtigte Skepsis.

CAD-Systeme brauchen grundsätzlich eine Hypothese und danach muss diese Hypothese mathematisch modelliert werden. Nehmen wir als Beispiel die CAD-Systeme für die MR-Mammografie. Die Hypothese ist, dass die Kontrastmittelanflutung und das Auswaschen des Kontrastmittels die entscheidenden Kenngrößen für ein Mammakarzinom sind. Also hat man Techniken entwickelt, um diese Kontrastmittelkinetik im MR robust darzustellen, Normkollektive untersucht und mathematisch die Unterschiede zwischen Normal und Karzinom modelliert. Wenn die Hypothese richtig ist, kann man mit solchen CAD-Systemen gute Ergebnisse erzielen. Ein KI-System lernt hypothesenfrei, d. h. man präsentiert dem Algorithmus nicht nur die KM-Kinetik, sondern auch viele andere Bilddaten (T2, SWI, Diffusion, alles, was das System hergibt), kann ihm aber auch klinische Daten, Laborwerte und genetische Informationen zum Lernen geben. Und der Algorithmus findet in dieser Datenfülle die Parameter, die mit der höchsten Wahrscheinlichkeit für oder gegen ein Mammakarzinom sprechen. Am Ende wird sicher die KM-Kinetik eine Rolle bei der Entscheidung finden, aber der Algorithmus wird mit unterschiedlicher Wichtigkeit auch andere Parameter prognostisch wichtig finden. Ohne Hypothese. Ich will hier nicht auf die verschiedenen Formen der KI mit Schlagworten wie „machine learning“, „deep learning“ oder „supervised“ und „unsupervised“ eingehen. Sie sollen sich nur merken, dass KI hypothesenfrei lernt. Und – darauf komme ich später – die Güte von KI-Anwendungen davon abhängt, wie gut die Daten sind, mit denen das System trainiert wurde.

Alltagsanwendungen von KI in der Radiologie

Ein Vorteil des Computers ist, dass er nicht müde wird und Routinarbeiten sehr zuverlässig durchführen kann. Und jetzt können wir nach solchen Tätigkeiten in der Radiologie suchen. Gibt es nicht? Also, ich finde das Zählen von MS-Plaques, die Größenbestimmung der Ventrikel-Weite im Verlauf nach einer SAB, das Ausmessen der Lebermetastasen während der Therapie, die Knochenalter-Bestimmung des Handskeletts und das Vermessen von Winkeln bei der Beurteilung der Skoliose nicht wirklich spannend. Und hier wird KI ganz schnell Routinen übernehmen. Alles, was wir dazu brauchen, sind Algorithmen, die darauf trainiert wurden, das Zielorgan und die Läsion zuverlässig zu erkennen. Und dann müssen die Algorithmen noch quantifizieren können. Diese Anwendungen gibt es bereits, überwiegend noch im Anfangsstadium. Nicht perfekt, aber wir erinnern uns, die erste Autofahrt war auch langsam. Und statt der blumigen Beschreibung von MS-Plaques an unterschiedlichen Lokalisationen im Gehirn mit oder ohne zarte Aufnahme von Kontrastmittel wird der Befund am Ende nur „plus 19“ (dann läuft die Therapie nicht gut) oder „minus 4“ (dann ist es prima) lauten. Der Radiologe kann sich dann darauf konzentrieren, einen der häufigsten Fehler in der Radiologie – Satisfaction of search – zu reduzieren: Vor lauter Zählerei der Plaques übersieht er nämlich gerne das Aneurysma am Ramus communicans anterior oder das Hypophysen-Adenom. Eine andere Anwendung, die gegenwärtig an vielen Stellen in der Welt verfolgt wird, ist die Befundung von konventionellen Thorax-Aufnahmen. Der zu tief liegende Tubus, die Schrittmacherelektroden, der ZVK, der Pneumothorax, die Tuberkulose, die Pneumonie im Mittellappen und 40 andere Pathologien werden schon sehr zuverlässig erkannt. Es gibt Algorithmen, die 70 oder gar 90 dieser Pathologien erkennen können und automatisch befunden. Ein Problem dieser Anwendungen ist, dass sie bisher schlecht in den Workflow der Radiologie oder des Radiologen integriert sind. Das kann mit unserer Hilfe schnell besser werden. Letztlich wurde uns von einer kleinen Firma ein Produkt präsentiert, das 40 typische Pathologien im konventionellen Röntgenbild des Thorax diagnostizieren konnte. Bei den einfachen Dingen (ZVK, Tubus usw.) war der Algorithmus perfekt, bei Lungenfibrosen katastrophal. Warum? Die Trainingsdaten kamen aus einem drittklassigen Krankenhaus. Ein anderes und viel schwerwiegenderes Problem bei der Etablierung von KI in der Radiologie sind also die Trainingsdaten. Darauf gehe ich weiter unten ein.

Zum Schluss dieses Abschnitts eine kurze Betrachtung des Screenings: Stellen Sie sich vor, China würde ein Mammografie-Screening-Programm starten: Dieses würde niemals mit Radiologen funktionieren. Es gibt viel zu wenig Radiologen und viel zu viele Frauen, die ein Screening bekommen sollten. Natürlich würde dieses Screening in China von intelligent trainierten Algorithmen durchgeführt. Und dann stellen Sie sich vor, das Lungen-screening wird in den USA und Deutschland eingeführt. Welcher Radiologe hätte Spaß daran, den ganzen Tag nur Screening-CTs der Lunge anzugucken? Auf dem ECR 2018 wurde in einer Sitzung, in der KI-Anwendungen gezeigt wurden, in der Diskussion gefragt, ob wir statt des „4-Augen-Prinzips“ mit 2 Radiologen

beim Mammografie-Screening nicht besser das „KI plus 1 Radiologe“-Prinzip einsetzen sollten? Der schwarze Schwan wurde diskutiert. Aber klar, wir können uns auch widersetzen. Nur, ob das hilft?

Diagnostischer Mehrwert durch KI: Radiomics

Möglicherweise stecken in den Bilddaten ja mehr Informationen als bisher gedacht, also ein schwarzer Schwan. Die Datenfülle ist aber so groß, dass der Radiologe diese Mehr-Informationen nicht sieht. Bei Radiomics-Anwendungen wird ein Algorithmus wieder mit bekannten Daten trainiert. Nehmen wir das Beispiel eines Glioblastoms. Der Algorithmus bekommt alle MR-Bilddaten (von T1 nativ, über T2, SWI, DWI, KM-Bilder) und molekulargenetische Informationen. Am Ende kann der Algorithmus in unbekanntem MR-Datensätzen nicht nur die molekulare Differenzierung des Tumors erkennen, sondern auch ohne Gabe von Kontrastmittel das Ausmaß der Blut-Hirn-Schrankenstörung. Keine Utopie, auch schon da. Nicht überall, aber in einigen Abteilungen.

Ich gebe Ihnen ein paar weitere Beispiele: Wir haben einen Algorithmus darauf trainiert, auf PET-MR-Aufnahmen beim Zervix-Karzinom nur bei Betrachtung des Tumors mit über 95%iger Wahrscheinlichkeit zu erkennen, ob die Patientin Organmetastasen hat oder nicht. Mit entsprechenden Trainingsdaten kann KI die molekularbiologische Typisierung von Tumoren ohne Biopsie erkennen, vorhersagen, ob ein Tumor auf Strahlentherapie anspricht, und die Ansprechwahrscheinlichkeit auf Chemotherapien vorhersagen. Alles Utopien? Hier darf ich die essayistische Variante kurz verlassen und verweise auf das Literaturverzeichnis: Alphabetisch sortiert zeigt es unterschiedliche Anwendungsbeispiele von KI in der Radiologie [1–39].

Warum ist KI noch nicht so richtig in der Medizin angekommen?

KI ist in vielen Bereichen der Industrie und der Robotik schon Standard. Aber in der Medizin reden wir erst seit kurzer Zeit davon. Warum hängt die Medizin hinterher? Vielleicht, weil Medizin (auch Radiologie) viel zu kompliziert ist? Sicher nicht. Mustererkennung – und nicht nur Radiologie ist „pattern recognition“ (dazu später noch ausführlicher) – ist eine Stärke von KI. Militärische Anwendungen von KI beruhen zum großen Teil auf dem Erkennen von Mustern. Die Grundvoraussetzung für das Training von intelligenten Algorithmen sind valide Trainingsdatensätze. Und gucken wir uns die Player in der KI-Szene an (Google, Microsoft, Amazon...), dann wird klar, dass diese eigentlich keinen Zugriff auf valide medizinische Datensätze haben. Google weiß wahrscheinlich eher als jeder Amtsarzt, wo demnächst die nächste Grippeepidemie losgeht – weil die Menschen vermehrt nach irgendwelchen Symptomen über die Suchmaschine suchen –, aber valide radiologische Datensätze sind nicht einfach zu bekommen. Aber Siemens, GE oder Philips! Was ist mit denen? Ok, sie könnten theoretisch auf viele CT- und MR-Datensätze gucken (wir lassen bei dieser Betrachtung mal das Datenschutzproblem außer

Acht), aber – vereinfacht – die Diagnosen müssen ja nicht richtig sein. „Old“ und „New Economy“ können also nicht so ohne weiteres KI in die Medizin (und die Radiologie) bringen.

Also fragen wir weiter: Wer hat das nötige Wissen und das Material? Lehrbuchverlage wäre eine Antwort. Schon seit vielen Jahren reicht es den Verlagen nicht mehr, wenn die Autoren eine Erkrankung nur mit einem Bild illustrieren. Der Autor soll das Epidural-Hämatom mit einem 3D-Datensatz illustrieren. Verlage haben also Lehrbuchwissen, in der IT-Sprache ist das „ground truth“ – also die echte Wahrheit. Jetzt braucht man zum Trainieren von intelligenten Algorithmen nicht die vielzitierten „big data“, sondern nur valide Daten und gut annotierte Bilder, aber $N > 1$ braucht man dann doch. Also wahrscheinlich haben die Verlage einfach nicht genug Daten, um einen Algorithmus so zu trainieren, dass er das Knochenalter von Kindern bestimmen kann. Und jetzt sind wir auch bei der Lösung: Große Krankenhäuser und namentlich Universitätskliniken haben solche Daten. Leider strubbelig abgelegt und unsortiert, manchmal vielleicht auch falsch. Wir brauchen also in den großen Krankenhäusern so etwas wie „Data-Validation“-Abteilungen. Da werden die Daten (Radiologie, Labor, Pathologie, elektronische Patientenakte...) auf Richtigkeit überprüft und in einen Daten-See gegeben, der eine Struktur hat, die für KI zugänglich ist. Geht nicht? Naja. Zu Zeiten, als die Krankenhäuser noch nach der Zahl der belegten Betten von den Kostenträgern honoriert wurden, brauchten wir auch keine Kodierungsabteilungen.

Und jetzt? DRG hat es möglich gemacht. Ok, da war ja der ökonomische Druck. Stimmt. Aber KI macht ökonomischen Sog! Stellen Sie sich vor, Sie haben ganz viele Daten zum Prostata-Karzinom und trainieren einen Algorithmus mit Bilddaten, Laborbefunden, molekularbiologischen, genetischen und pathologischen Daten. Und dann haben Sie noch den Verlauf von Hunderten von Patienten. Wenn es ein Muster für die gutartige und die bösartige Variante des Prostata-Karzinoms gibt, dann wird KI es in dem Daten-See finden. Und dann haben Sie einen Algorithmus – in dem übrigens keine Patientendaten mehr sind, sondern nur noch das Ergebnis des Trainings –, und den können Sie gerne mit einem strategischen Partner als Produkt verkaufen. Super Idee! Und wenn das dann alle machen? Also alle können schon mal nicht alles machen! Wir könnten „Leber“, aber nicht „Rheuma“, und so hat jeder Schwerpunkte. Und außerdem gibt es ja auch mehr als eine Firma, die Autos oder MR-Geräte herstellt. Am Ende ist das eine Produkt vielleicht besser in den Workflow integriert, das andere schneller, das dritte mit mehr oder weniger Parametern trainiert, die Daten sind besser oder schlechter validiert. Es gibt also genügend Merkmale, die die einzelnen Anwendungen voneinander unterscheiden. Und unabhängig vom ökonomischen Sog: KI wird ein riesiges Forschungsgebiet in der Medizin und der Radiologie werden. Für die Universitätskliniken ein nicht zu verachtender Nebeneffekt.

Ganz am Ende dieses Abschnitts noch ein weiterer Aspekt: Oben habe ich gesagt, dass Google und Amazon ein Problem haben, an valide medizinische Datensätze zu kommen. Das stimmt bis jetzt auch. Aber beide könnten in der Zukunft selbst Krankenhäuser betreiben – vielleicht erst nur für Mitarbeiter und auch nur den ambulanten Bereich. Dann hätten sie Daten und zwar in einer Struktur, die sie haben wollen. Und der Gesundheits-

markt ist so gigantisch groß, dass die Googles dieser Welt dort einsteigen werden. Sie glauben, dass niemand in ein Amazon-Krankenhaus geht, weil der Datenschutz ein Problem sein könnte? Wie viele Menschen kaufen täglich bei Amazon ein?

Super-Diagnostics

Sie können sich jetzt vorstellen, was KI in der Radiologie alles können wird? Langweilige Tätigkeiten eigenständig übernehmen und Radiomics von der Detektion zur Prädiktion eine neue Ära einläuten. Springen wir kurz in andere diagnostische Fächer: KI wird lernen, in die Petrischale des Mikrobiologen zu gucken und feststellen, dass es Streptokokken sind, die sich dort tummeln. Und KI wird lernen, bei der HIV-Medikation die Resistenzen vorherzusagen. Die digitale Pathologie ist am Ende so etwas wie Radiologie mit Farbe, auch Mustererkennung. Routinen wird KI schnell lernen, Pathomics wird es als Begriff für tiefere Analysen geben. Genetische Analysen werden mit KI in eine andere Dimension geführt. Und Krankheiten aus bestimmten Konstellationen von Laborwerten diagnostizieren kann in Zukunft auch nicht mehr nur der weise Professor.

Kurz und gut: Wir werden KI ziemlich schnell in allen diagnostischen Disziplinen haben und die richtige – personalisierte (siehe oben) – Therapie-Entscheidung wird zunehmend aus diesen Super-Diagnostic-Zentren kommen. Jetzt wird Ihnen schwindelig? Data-Validation-Abteilung und Super-Diagnostic-Zentrum.

Alles zu viel? Aber wenn Sie schon eine Data-Validation-Abteilung haben, dann ist der Schritt zum SD-Zentrum doch nur noch sehr klein. Es braucht eine Struktur, aber irgendwie ist es wie ein Tumorboard. Am Ende nur schlanker, weil ein Teil der Diagnosen und Therapieempfehlungen automatisch erfolgt, also für ein Glioblastom mit einer bestimmten Mutante immer die dafür vorgesehene Standardtherapie. Und natürlich nicht nur für Tumoren, sondern für eine Vielzahl von Erkrankungen. Wer kann solche Zentren für Super-Diagnostics aufbauen? Eigentlich nur die Universitätskliniken, weil in der Regel nur sie alle diagnostischen Disziplinen vorhalten und damit auch in der Lage sind, die Daten strukturiert zusammenzuführen. Vielleicht wird Super-Diagnostics ein wesentliches Merkmal der Universitätsmedizin werden.

Ethische und juristische Dimension von KI

Wenn ich – gelegentlich – einen Vortrag über KI halte, ist gar nicht so selten eine Spezies im Publikum, die Medizin und Jura parallel oder nacheinander studiert hat, und immer ist jemand im Publikum, der warnend den Zeigefinger erhebt und das Wort „Ethik“ in den Raum wirft.

Fangen wir mit der Ethik an. Wir sind uns einig, dass Medizin immer eine ethische Dimension hat. Es gilt also, in der Diskussion herauszufinden, welche spezifische ethische Dimension KI dabei hat. Wäre das Thema des Vortrags: „Mit mehr Intelligenz die Medizin besser machen“, gäbe es keine Ethik-Anmerkung. Und nur weil es „künstliche Intelligenz“ ist, macht da jemand Theater. Ersetzen wir doch „künstliche Intelligenz“ durch „Computerintelligenz“. Hat jemand „Ethik“ gesagt, als die Computertomografie die Palpation der Leber ersetzt hat? Ich erinnere mich nicht.

Eigentlich ist es eher verwunderlich, dass kein Ethiker bisher aufgestanden ist und gesagt hat, dass es wahrscheinlich unethisch ist, eine KI-Technik nicht einzusetzen, wenn man damit schneller und sicherer Diagnosen stellen kann. Und schon sind wir bei der Juristik. Letztlich erhob sich nach einem Vortrag ein Richter aus dem Publikum und prophezeite mir – er meinte nicht mich persönlich, sondern bestimmt alle, die so ketzerische Gedanken haben – großen juristischen Ärger, weil ein Richter nicht nachvollziehen könnte, wie KI auf eine Diagnose kommt. Und überhaupt, was denn wäre, wenn KI falsch liegen würde. Die große Keule.

Aber fangen wir mit der Deeskalation am Ende an. Google plant schon länger, dass man sich seine Mammografie im Netz befunden lassen kann. Für 1 Dollar. Ist noch nicht Wirklichkeit, vielleicht werden es auch 5 Dollar, aber es kommt. Und nun ist das Szenario, dass 2 deutsche Radiologen die Mammografie als unauffällig befundet haben; die Frau hat ein halbes Jahr später Lebermetastasen eines eben nicht vorhandenen Mamma-Karzinoms. Der Ehemann der Patientin lädt die Mammografien bei Google hoch und der Algorithmus sagt, das Mamma-Karzinom war schon da. Die beiden Radiologen bleiben bei ihrer Einschätzung und 2 Gutachter auch. Und das passiert jetzt 100-mal. Geschenkt, das wird so nicht passieren, weil wir Radiologen schon vorher in zig Studien rausgefunden haben, ob das so ist (und sein kann) und durch die Studienergebnisse schon den Algorithmus als Unterstützung für die eigene Befundung genommen haben. Und dann hätte genau der Radiologe ein „großes juristisches Problem“, der den Algorithmus nicht verwendet hat. Ob und wieviel der Richter von dem Algorithmus versteht, wird die Sache nicht entscheiden. Bleibt das Problem mit der Fehldiagnose des Algorithmus. Also erstmal wird es weniger Fehldiagnosen geben, weil der Algorithmus ja immer mit Expertenwissen trainiert wurde. Nicht alle Ärzte sind Experten für alles und auch nicht alle Radiologen auf Experten-Niveau. Und dann gucken wir doch einfach in einen anderen Bereich der Medizin, in dem schon lange Ergebnisse vollautomatisch ermittelt werden – in die Labormedizin. Wie merkt man, dass die Maschine außer Rand und Band ist? Qualitätskontrollen. Und genau so wird die gesamte KI-Diagnostik streng qualitätskontrolliert. Es werden immer Bilddaten oder mikrobiologische oder pathologische Proben durch das System gehen, deren Ergebnis bekannt ist und das die Maschine richtig diagnostizieren sollte.

Zusammengefasst: KI kann eine ethische Dimension bekommen, aber nur, wenn sie nicht angewendet wird. Und die juristischen Lehrbücher müssen nicht umgeschrieben werden.

Werden nur die „technischen Fächer“ durch KI verändert?

Anfang des Jahres habe ich einen Vortrag in einer großen Stadt gehalten und der Moderator – ein Neurologe und Psychiater – führte auf meinen Vortrag hin mit ein paar überschwänglichen Bemerkungen zum Wert der sprechenden Medizin und zum baldigen Ende des Radiologen durch KI. Was soll ich sagen, ich war gut vorbereitet. Warum reden wir gerade so viel von Radiologie, wenn es um KI geht? Weil wir schon lange digital sind und Daten haben, die KI lesen kann. Wenn man mal fast philosophisch an KI und

Medizin rangeht, muss man sich eigentlich die Frage stellen, wo wird KI den meisten Mehrwert bringen? Da, wo die meisten Fehler gemacht werden. Und in der Radiologie werden relativ wenig Fehler gemacht! Stimmt wirklich. Wenn wir ein CT oder MR der Leber machen und finden dort eine Läsion von 3 cm Größe, dann wird da schon eine Läsion mit ungefähr dieser Größe sein. Gut, wir können je nach Qualifikation die Artdiagnose falsch stellen, aber wir wissen, es gibt ein Problem in der Leber. Die meisten Fehler werden in der sprechenden Medizin gemacht. Üblicherweise gibt es an dieser Stelle in dem Vortrag vor Neurologen und Psychiatern Theater im Saal. Aber gucken wir in die Realität: Eine Mitte 50-jährige Frau hat Rückenschmerzen und geht nacheinander zum Allgemeinmediziner, zum Orthopäden und Physiotherapeuten. Sie hat fast immer mehr als ein MR der LWS im Laufe von 2 Jahren – die LWS zeigt die alterstypischen Veränderungen, die wir – leider – auch immer ausführlich beschreiben –, sie hat zig Fangopackungen und Faszien-Therapien bekommen und vielleicht auch eine Operation der LWS mit Versteifung von 4 Segmenten. Die Rückenschmerzen sind immer noch da. Wir alle kennen diese Fälle. Und nach 2 Jahren – im Durchschnitt – kommt jemand auf die Diagnose „Depression“. Wenn Sie die Fotos, die diese Frau auf Instagram hochgeladen hat, analysieren oder – noch besser – ihre „daily activity“ auf Facebook, können Sie die Diagnose sofort stellen. Oder Sie bedienen sich des Algorithmus, der darauf trainiert wurde, die Sprachmelodie und Stimmlage bei depressiven Patienten zu erkennen. Dann haben sie es nach 2 Minuten raus. Auch als Radiologe. Und sobald digitale Daten aus der sprechenden Medizin vorliegen – Google arbeitet wie verrückt daran – wird die sprechende Medizin völlig verändert werden.

Aber sonst bleibt Radiologie unverändert?

Nein, natürlich nicht. Sie haben akzeptiert, dass Radiologie durch KI in die Radiomics-Dimension vorstößt und Zentren für Super-Diagnostics die Zukunft sind. Aber auch das Geschäftsmodell der Radiologie wird sich ändern. Als ich 1987 anfang mit der Radiologie war alles unter einem Dach. Der Arzt saß am Gerät und guckte am Monitor die Bilder an. Und es gab einen Film, der entwickelt werden musste und der zur Befundung verwendet wurde. Die Befundung fand in den Räumen der Radiologie statt. Heute arbeitet schon ein Teil meiner Mitarbeiter im Home-Office. Und heute werden die MR-Geräte auch schon „remote“ bedient. In einem Raum, der weit weg ist von allen MR-Geräten, sitzen hochqualifizierte MTRAs und fahren 2 oder 3 MR-Untersuchungen parallel. Am Gerät und in der Wartezone übernehmen Nicht-MTRAs das Patientenmanagement inklusive der Lagerung. Mit neuen, deckenartig konzipierten Oberflächenspulen ist die Lagerung weniger kompliziert und das Gerät erkennt – dank KI – die zu untersuchende Region zunehmend selbständig.

Und dann gibt es noch das Geschäftsmodell von Amazon – die Plattform. Man könnte auch Uber oder AirBnB nehmen. Die Bilder werden auf eine Maklerplattform gestellt und jeder Radiologie, der Zeit und Lust hat, kann sie – qualitätsgesichert – befunden. Welch eine Riesenchance für Krankenhäuser, die keinen Radiologen mehr finden und welche Chance für große Praxisverbände und Krankenhausabteilungen, die wachsen können ohne zu inves-

tieren. Und vielleicht wird die Indikation zum LWS-MR bei Rückenschmerzen auch nicht mehr ausschließlich vom Orthopäden gestellt, sondern von einer KI, die die Anamnese des Patienten erhebt, und der Arzt gibt noch den körperlichen Untersuchungsbefund ein. Vielleicht ist das aber auch ein Physician Assistant. Und dann geht der Patient zum dem MR-Gerät, das bei Aldi um die Ecke steht und remote bedient wird. Und der Befund wird auf der Plattform gemacht.

Ich gebe Ihnen Recht. Sehr utopisch und vielleicht nicht morgen Realität. Und vielleicht kommt es auch ein bisschen anders. Aber denken Sie an den schwarzen Schwan aus der Einführung. Denken Sie nicht darüber nach, warum es nicht gehen wird (juristisch, ethisch, Datenschutz, Gewerkschaft, Ärztekammer, Stromausfall). Ich weiß, dass das Suchen und Finden von Problemen in Deutschland gerne als hohe intellektuelle Leistung angesehen wird. Als wir 2001 RIS-PACS eingeführt haben, gab es innerhalb und außerhalb der Radiologie fast nur Zweifler und Problem-Erkenner. Und die Übergangsphase – Voraufnahmen analog im Archiv und die aktuellen Aufnahmen im PACS – war nicht lustig. Und so wird es auch mit KI werden. Die Technologie wird kommen und wir können sie aktiv mitgestalten. Also denken Sie an Lösungen.

Michael Forsting

Literatur

- [1] Abajian A et al. Predicting Treatment Response to Intra-arterial Therapies for Hepatocellular Carcinoma with the Use of Supervised Machine Learning-An Artificial Intelligence Concept. *J Vasc Interv Radiol* 2018; 29: 850–857 e851
- [2] Abdollahi H et al. Cochlea CT radiomics predicts chemoradiotherapy induced sensorineural hearing loss in head and neck cancer patients: A machine learning and multi-variable modelling study. *Phys Med* 2018; 45: 192–197
- [3] Al'Aref SJ et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 2018. doi:10.1093/eurheartj/ehy404
- [4] Al-Masni MA et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput Methods Programs Biomed* 2018; 157: 85–94
- [5] Al Ajmi E et al. Spectral multi-energy CT texture analysis with machine learning for tissue classification: an investigation using classification of benign parotid tumours as a testing paradigm. *Eur Radiol* 2018; 28: 2604–2611
- [6] AlBadawy EA et al. Deep learning for segmentation of brain tumors: Impact of cross-institutional training and testing. *Med Phys* 2018; 45: 1150–1158
- [7] Arbabshirani MR et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *Npj Digital Medicine* 2018; doi.org/10.1038/s41746-017-0015-z
- [8] Baessler B et al. Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy-Preliminary results. *Eur J Radiol* 2018; 102: 61–67
- [9] Bektas CT et al. Clear Cell Renal Cell Carcinoma: Machine Learning-Based Quantitative Computed Tomography Texture Analysis for Prediction of Fuhrman Nuclear Grade. *Eur Radiol* 2018. doi:10.1007/s00330-018-5698-2
- [10] Bluemke DA. Radiology in 2018: Are You Working with AI or Being Replaced by AI? *Radiology* 2018; 287: 365–366

- [11] Caballo M et al. An unsupervised automatic segmentation algorithm for breast tissue classification of dedicated breast computed tomography images. *Med Phys* 2018; 45: 2542–2559
- [12] Chilamkurthy S et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet* 2018. doi:10.1007/s00330-018-5698-2
- [13] Choi W et al. Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer. *Med Phys* 2018; 45: 1537–1549
- [14] Choy G et al. Current Applications and Future Impact of Machine Learning in Radiology. *Radiology* 2018; 288: 318–328
- [15] Chung SW et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthop* 2018; 89: 468–473
- [16] Couture HD et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer* 2018; 4: 30
- [17] Del Gaizo J et al. Using machine learning to classify temporal lobe epilepsy based on diffusion MRI. *Brain and Behavior* 2017; 7: doi:10.1002/brb3.801
- [18] Dreyer KJ, Geis JR. When Machines Think: Radiology's Next Frontier. *Radiology* 2017; 285: 713–718
- [19] Hosny A et al. Artificial intelligence in radiology. *Nat Rev Cancer* 2018; 18: 500–510
- [20] Katzen J, Dodelzon K. A review of computer aided detection in mammography. *Clin Imaging* 2018; 52: 305–309
- [21] Krupinski EA. Deep Learning of Radiology Reports for Pulmonary Embolus: Is a Computer Reading My Report? *Radiology* 2018; 286: 853–855
- [22] Kumar V et al. Automated and real-time segmentation of suspicious breast masses using convolutional neural network. *PLoS One* 2018; 13: e0195816
- [23] Lee H et al. Deep feature classification of angiomyolipoma without visible fat and renal cell carcinoma in abdominal contrast-enhanced CT images with texture image patches and hand-crafted feature concatenation. *Med Phys* 2018; 45: 1550–1561
- [24] Liu F et al. Deep Learning MR Imaging-based Attenuation Correction for PET/MR Imaging. *Radiology* 2018; 286: 676–684
- [25] Montoya JC et al. 3D Deep Learning Angiography (3D-DLA) from C-arm Conebeam CT. *Am J Neuroradiol* 2018; 39: 916–922
- [26] Nichols JA et al. Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophys Rev* 2018. doi:10.1007/s12551-018-0449-9
- [27] Nishio M et al. Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. *PLoS One* 2018; 13: e0195875
- [28] Orooji M et al. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. *J Med Imaging (Bellingham)* 2018; 5: 024501
- [29] Park SH. Regulatory Approval versus Clinical Validation of Artificial Intelligence Diagnostic Tools. *Radiology* 2018; 288: 910–911
- [30] Pesapane F et al. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights Imaging* 2018; 9: 745–753
- [31] Salem M et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin* 2018; 17: 607–615
- [32] Savadjiev P et al. Demystification of AI-driven medical image interpretation: past, present and future. *Eur Radiol* 2018. doi:10.1007/s00330-018-5674-x
- [33] Tajmir SH et al. Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability. *Skeletal Radiol* 2018. doi:10.1007/s00256-018-3033-2
- [34] Ting DSW et al. Clinical Applicability of Deep Learning System in Detecting Tuberculosis with Chest Radiography. *Radiology* 2018; 286: 729–731
- [35] Titano JJ et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med* 2018; 24: 1337–1341
- [36] van Rosendael AR et al. Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry. *J Cardiovasc Comput Tomogr* 2018; 12: 204–209
- [37] Yasaka K et al. Deep learning for staging liver fibrosis on CT: a pilot study. *Eur Radiol* 2018. doi:10.1007/s00330-018-5499-7
- [38] Yepes-Calderon F et al. Automatically measuring brain ventricular volume within PACS using artificial intelligence. *PLoS One* 2018; 13: e0193152
- [39] Zhao X et al. Agile convolutional neural network for pulmonary nodule classification using CT images. *Int J Comput Assist Radiol Surg* 2018; 3: 585–595