

Da hilft nur noch Kollege Computer

Moderne bildgebende Verfahren liefern den Ärzten immer mehr Informationen. Mit der Auswertung sind sie zunehmend überfordert. Künstliche Intelligenz könnte ein Ausweg sein. Ersetzt sie bald den ausgebildeten Fachmediziner?

Von Michael Brendler

Nicht einmal der beste Doktor ist unfehlbar. Das führt zwangsläufig dazu, dass immer wieder Krankheitsbilder übersehen werden. Zum Beispiel der schwarze Hautkrebs. Jedes Jahr kostet er dreitausend Deutsche das Leben, weil er zu spät erkannt wurde. Selbst geübten Dermatologen passiert es regelmäßig, dass ihnen erste Anzeichen des Melanoms entgehen.

Dem Computerforscher Sebastian Thrun vom Labor für Künstliche Intelligenz lief das keine Ruhe. Vor einigen Jahren begann der Gründer der Google-Entwicklungsschmiede Google X deshalb damit, seine Maschinen mit fast einhundertdreitausend Aufnahmen verschiedener Hautveränderungen zu füttern. Die eingesetzte KI erwischt sich als gefähiger Schüler. Beim Test an rund zweltausend Fotografien übersah sie in der Regel nicht nur weniger Melanome als zu erfahrene Dermatologen, denen dieselben Aufnahmen vorgelegt wurden. Sie suchte auch seltener eine harmlose Verbindung als Tumor ein. „Die Technik hat das Potential, die Medizin substantiell zu verändern“, folgerten Thrun und seine Ko-Autoren im vergangenen Jahr in ihrer Veröffentlichung in *Nature*.

Vor allem Röntgenärzte verstehen das als Kampfansage. Und das zu Recht. Schließlich haben sich Computer schon bei ganz anderen Herausforderungen dem Menschen überlegen gezeigt. Bereits 1997 wies eine Maschine den Schachweltmeister Gari Kasparow in die Schranken, 2016 mussten sich auch die klügsten Köpfe beim Brettspiel Go geschlagen geben. Thruns Arbeit auf dem Gebiet der Medizindiagnostik ist nur ein besonders beeindruckendes Beispiel von vielen, die gezeigt haben: Wenn es um das Aufspüren von Krankheitszeichen auf Bildern geht, ist eine selbstlernende Maschine womöglich der aufmerksamere Beobachter. Vielleicht gelingt es ihr in Zukunft einfach besser, die schwachen Schattierungen, winzigen Formveränderungen und leicht zu übersehenden Flecken zu interpretieren, die auf Röntgen- und Hautbildern den Unterschied zwischen gesund und krank ausmachen.

Gerade die Röntgenärzte haben im Vergleich zu den anderen Mediziner besonders viel zu verlieren. Für sie ist die Suche nach verräterischen Strukturen nicht irgendein ungeliebter Nebenjob. Für sie geht es beim Wertlauf mit dem Computer um ihr Hauptgeschäft und damit um ihre Existenzgrundlage. Selbst auf ihrem ureigenen Spezialgebiet, bei der Auswertung von Mammographie-Aufnahmen, konnte die Künstliche Intelligenz kürzlich zeigen, dass sie der echten mindestens ebenbürtig ist.

Im vergangenen Jahr sah das auch die strenge amerikanische Arzneimittelbehörde FDA ein. Sie akzeptierte eine Software, die das Herzvolumen von Patienten nicht nur schneller messen kann als ausgebildete Radiologen, sie soll auf Basis der ermittelten Daten bald auch eigene Diagnosen stellen. „Man sollte aufhören, weiter Röntgenärzte auszubilden“, forderte provokant der Informatiker Geoffrey Hinton und KI-Pionier von der Universität Toronto 2017 in der Zeitschrift *New Yorker*. „Es ist völlig offensichtlich, dass Künstliche-Intelligenz-Systeme in fünf Jahren besser sind.“

Andreas Boss, leitender Arzt am Institut für diagnostische und interventionelle Radiologie des Universitätsklinikums Zürich, hat den neuen Konkurrenten auf die Probe gestellt. Moderne Grafikkarten und eine Art frei verfügbares Google-Betriebssystem, so berichtet er, machen es auch ihm ohne große Umstände möglich, auf die neue Technik umzusatteln. Es geht dabei um das sogenannte Deep Learning, das seit fünf, sechs Jahren vor allem in der Sprach- und Gesichtserkennung Furore macht. Bei dieser Anwendung imitieren künstliche neuronale Netzwerke die Arbeit der Nervenzellen des Gehirns. Signale werden in Schichten verarbeitet, jedes Neuron und jede Lage widmet sich einem anderen Aspekt und leitet die Information an die nächste Schicht weiter. Die Maschine lernt, genau wie der Mensch, indem sie punktuell die Effizienz verändert, mit der die Daten übertragen werden.

Letztlich erwerbe der Computer dadurch sein Wissen nicht anders als ein Kind, sagt die Fachleute. Indem er aus den eigenen Fehlern lernt. Wer einen kleinen Jungen nur oft genug Hundelein und Wölfe zeigt, kann darauf vertrauen, dass dieser bald beide von allein zu unterscheiden lernt. Zunächst wird man ihn noch korrigieren müssen, irgendwann wird er sich die maßgeblichen Merkmale selbst herauspicken. Auf ähnliche Weise legte Andreas Boss einer Vorläuferversion seines neuronalen Netzwerks ausgewertete Mammographien von gesunden und krebstranken Frauen vor. „Die Methode hat beängstigend gut funktioniert“, berichtet der Radiologe, der auch als Physiker ausgebildet ist. Der Algorithmus entdeckte auf den Aufnahmen eine genügend große Anzahl von Tumorkennzeichen, um Brustkrebs genauso zuverlässig aufzuspüren wie erfahrene Röntgenärzte. „Und das nach dem Studium von nur zweihundertachtzig Bildern“, staunt der Arzt. „Da wusste ich, dies wird unseren Beruf revolutionieren.“

Die Firma IBM hofft sogar, mit Hilfe der Künstlichen Intelligenz und ihrem Watson-Supercomputersystem einmal ganze Krankengeschichten analysieren zu können. Doch bei der Analyse von CT-, Kernspiri- und Ultraschallbildern könnte die Technik noch viel früher ihre Stärken ausspielen, sagt Michael Forsting, Direktor des Instituts für Diagnostische und Interventionelle Radiologie der Universitätsklinik Essen. Auf die Erkennung von Mustern und kleinen Details sind die Algorithmen besonders spezialisiert. Zu-

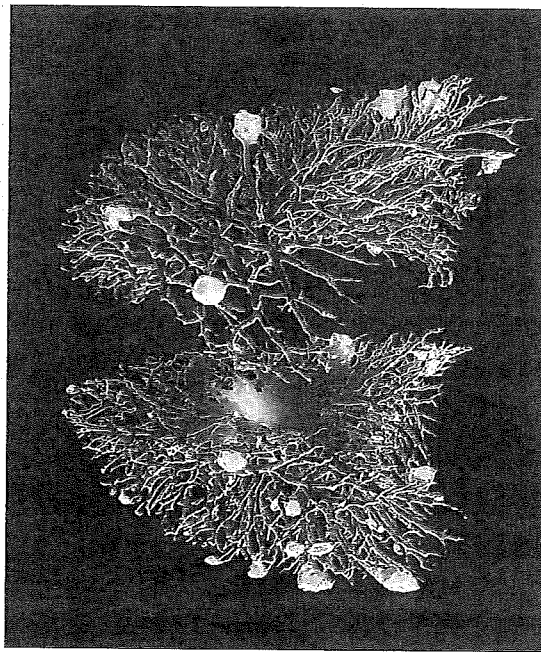
dem gebe es gerade in seinem Fachgebiet ausreichend. Schulungsmaterial für die Geräte, schließlich sei die Radiologie mittlerweile komplett digitalisiert.

Womöglich bleibt den Röntgenmediziner in Zukunft auch gar keine andere Wahl, als auf externe Hilfe zurückzugreifen. Anders werden sie der Datenflut wohl kaum noch Herr werden. Wer früher per Computertomograph den ganzen Körper untersuchte, erinnert sich Boss, der musste ein paar hundert Schichtaufnahmen durchscannen. „Heute produzieren wir für denselben Zweck in neunzig Sekunden dreitausend Querschnitte, die muss sich dann irgendein armer Kerl angucken.“

Weil auch andere Geräte immer genauer werden, fällt eine menschliche Schwäche bei der Arbeit zunehmend ins Gewicht: Selbst der aufgeweckteste Radiologe wird, wenn er eine fast identische CT- oder Mammographie-Aufnahme nach der anderen studiert, irgendwann müde und unkonzentriert. Im telemedizinischen Nachdienst, wenn ein einzelner Arzt die Aufnahmen aus einem ganzen Netz von Kliniken überwachen muss, steigt deshalb die Fehlerrate mit der Dauer der Schicht extrem an. Zudem urteilen auch Mediziner nicht immer objektiv: Zu den bekanntesten Fehlern bei der Röntgenbefundung zählt die sogenannte „satisfaction of search“, die Befriedigung des Findens. Wer am Wochenende das Opfer eines Motorradunfalls untersucht, gibt sich möglicherweise mit dem Entdecken einer Wirbelfraktur zufrieden und übersieht dabei eventuell den zufällig mit aufgenommenen Lungenknötchen.

Dem Computer sind solche Gefühle fremd, seine Leistungskurve bleibt selbst bei den langweiligsten Routineaufgaben noch stabil. Ein Radiologe, so konnte Michael Forsting in Essen zeigen, kann deshalb beim augenblicklichen Stand der Technik zwar bei der Beurteilung von Lungen-Röntgenaufnahmen gerade noch mit der Maschine mithalten. Beide können etwa gleich gut Erkrankungen des Organs erkennen. „Aber am Ende wird der Computer besser sein“, sagt der Experte, „weil er weniger Schwankungen kennt.“

In seiner Klinik nutzt Forsting die Deep-Learning-Systeme auch dazu, den Bilddaten ganz neue Informationen zu entlocken. „Mit dem Computer gehen wir tief in die Biologie des Tumors hinein“, sagt er. Allein durch die Analyse von Form und Struktur von Gebärmutterhals-Krebsgeschwüren konnte sein Algorithmus beispielsweise Metastasen an anderen Körperstellen voraussagen oder das Überleben der Patienten prognostizieren. Der Mensch war mit solchen Analysen bisher überfordert. Weil der Algorithmus völlig unabhängig lernt, sucht er



Metastasen im Gefäßnetz einer Lunge. Auch hier hat der Computer oft das bessere Auge. Foto: Maurizio, Bildagentur-online, US-National Cancer Institute, Fraunhofer MEVIS

sich beim Studieren der Schulungsaufnahmen eigene Erkennungszeichen, an denen er gesundes und krankes Gewebe erkennt. Nicht selten sind das welche, die einem echten Arzt gar nicht auffallen würden.

Allerdings kann genau das auch gewaltig in die Irre führen, wie man am Fraunhofer-Institut für Bildgesteuerte Medizin (MEVIS) in Bremen erfahren musste. Externe Kollegen hätten in der Vergangenheit versucht, ein Deep-Learning-System auf das Erkennen von Pleuratergüssen zu trainieren, berichtet Horst Hahn, der Di-

rektor des Instituts. Dabei handelt es sich um eine Ansammlung von Wasser in der Brusthöhle, beispielsweise bei einer Lungenerkrankung oder Herzschwäche. Tatsächlich erkannte der Computer das Problem sehr zuverlässig, nur leider an den falschen Kennzeichen. Ihm waren während des Trainings vor allem die Herzschrittmarker ins Auge gefallen, die ebenfalls die Röntgenbilder von gesunden und kranken Probanden unterscheiden. „Ein Mensch versteht sofort, dass das eine mit dem anderen nichts zu tun hat“, sagt Hahn. Nicht so der Computer. Er ist

nicht in der Lage, sich Fragen nach der Kausalität zu stellen wie: Sind Schrittmacher und Erguss wirklich Folge der gleichen Krankheit? Oder ist der Gesundheitszustand der Pleura-Patienten nicht einfach insgesamt schlechter, und sie haben deshalb nur zufälligerweise häufiger Herzrhythmusstörungen? Der Algorithmus versucht dagegen nur, die gestellte Aufgabe nur anhand der ihm vorgelegten Daten zu lösen.

Auch aus diesem Grund tat sich die Deep-Learning-Software anfangs schwer, als ihr Hahn und seine Kollegen beizuhelfen.

„Mit Maschinen kann man nicht diskutieren“

Ein Interview mit dem Mediziner Eckhard Nagel über die Grenzen der Künstlichen Intelligenz in der Medizin

Nun könnte man allerdings auch Maschinen mit den Symptomen und der Krankengeschichte füttern...

Das stimmt, und sie würden - entsprechend programmiert - sicherlich auch Vorschläge äußern, welche Diagnose sie für die wahrscheinlichste halten. Nur: Was soll der Betroffene dann anfangen? Zudem ignoriert dieses Vorgehen den wesentlichen Aspekt der Arzt-Patienten-Beziehung, nämlich die Wahrnehmung des Kranken in seiner Situation und die Einordnung von Gefühlen und Informationen in den Gesamtkontext der Befunde. Das macht den Arzt selbst heute, wo die Menge der Messdaten zunehmend unüberschaubar wird, immer wichtiger. Denn irgendetwas muss all diese Dinge zusammenführen und integrieren.

Unter welchen Umständen und in welchem Rahmen wäre es Ihrer Meinung nach in der Medizin vorstellbar, menschliche Intelligenz durch eine künstliche zu ersetzen?

Ich glaube, dass die Technik uns Ärzten helfen kann, an anderer Stelle nicht die Orientierung zu verlieren. Ich kann zum Beispiel bei der Verschreibung eines Medikaments nicht immer alle Neben- und Wechselwirkungen im Kopf haben. Wenn mich ein Computer rechtzeitig davor warnt, es falsch einzusetzen, dann halte ich das für eine hilfreiche und absolut positive Entwicklung. Aber die Maschine kann mir nicht die Entscheidung abnehmen, wie ich bei einem kritischen Patienten auf der Intensivstation weiter vorgehen soll. Hier reicht es nicht, einen Algorithmus zu bemühen.

Worum nicht?

Weil die Maschine dem Arzt nicht die Verantwortung abnimmt. Die Vorstellung, keine Radiologen mehr auszubilden, wäre deutlich kompetenter.

bringen versuchen, anhand von CT-Bildern eine menschliche Leber zu identifizieren und zu analysieren. An den Organen von Asiaten scheiterte es zunächst regelmäßig, ebenso an den Organen, die zu operiert worden waren. Die Leber europäischer Standardmaße definiert. Deep-Learning-Systeme bleiben wissenschaftlich kapierbar, sagt Hahn, wenn man sie weiter mit entsprechendem Anschauungsmaterial füttert, sehen sie schnell ein, dass die Art des Gewebes und der Umgebung viel zuverlässiger anzeigt, um welches Organ es sich handelt. Die Informanten haben so-

lösbar. Die Merkmale, an denen ein

Computer zum Beispiel einen Pleuraerguss zu erkennen glaubt, kann der Mensch nachkontrollieren und -korrigieren. Auch der MEVIS-Algorithmus habe inzwischen kapiert, sagt Hahn, dass sich eine Leber nicht nur durch europäische Standardmaße definiert. Deep-Learning-Systeme bleiben wissenschaftlich kapierbar, sagt Hahn, wenn man sie weiter mit entsprechendem Anschauungsmaterial füttert, sehen sie schnell ein, dass die Art des Gewebes und der Umgebung viel zuverlässiger anzeigt, um welches Organ es sich handelt. Die Informanten haben so-

gar Pläne in der Schublade, wie man solche Probleme präventiv per Umprogrammierung vermeiden könnte.

Aber das ist einweilen noch Zukunftsmusik und zeigt: Ohne menschliche Aufsicht kommt auch ein Deep-Learning-System bis auf weiteres nicht aus. Gerade in der Medizin, wo kein Körper wie der andere ist und auch normale, symptomlose Abnutzungerscheinungen oder das Zusammenwirken von mehreren Krankheiten gleichzeitig den Computern verwirren können. Mensch und Maschine seien deshalb nicht als Konkurrenz zu sehen, meint der MEVIS-Chef.

„Künstliche neuronale Netze sind viel erdümmer und systematischer. Die Menschen können in die Zusammenarbeit ihre überlegene Vorstellungskraft und ihren vielseitigen Erfahrungsschatz einbringen.“

Andreas Boss kennt noch ein anderes gewichtiges Argument, das gegen einen Verdrängungswettbewerb in der Radiologie spricht: Jemandem muss am Ende für die Diagnosen geradestehen.

„Auch der Computer wird Fehler machen, das steht fest“, sagt er. Eine hundertenprozentige Sicherheit strecke in den Bildern einzeln nicht drin, auch die beste Kernspin-Aufnahme kann die Realität nicht eins zu eins wiedergeben. Am Ende heißt das, muss immer ein Arzt dem Urteil zustimmen, selbst wenn er nicht unbedingt der Klügere ist. Der Zürcher Mediziner hat seine Computer deshalb so programmiert, dass sie nur noch als eine Art besonders erfahrener Berater agieren. Gewebeprobe empfohlen, Überwachung wahrscheinlich ausstehend, weiter lehnt sich sein Algorithmus nicht aus dem Fenster. Nun möchte er dem Gerät noch beibringen, diese Entscheidungen auch zu begründen.

Also nicht nur anhand selbstgewählter Merkmale für den Arzt nicht nachvollziehbare Entscheidungen zu fällen, sondern Schritt für Schritt darzustellen, welche Auffälligkeiten für die Diagnose sprechen.

„Der Röntgenarzt wird noch ganz lange auf dem Fahrersitz sitzen“, glaubt auch der Essener Michael Forstung. Nur werde ihm der Computer in Zukunft weitgehend die radiologischen Pflichten abnehmen, während er sich um die bandenärztlichen vom Hals halten, bei denen der Mensch seine Schwächen hat. Das Zählen der stets gleichen Entzündungsherde im Gehirn von Multiple Sklerose-Patienten zum Beispiel. Oder die gleichförmigen Verlaufsberechnungen von Tumorkranken. Auch bei der schnellen Suche nach Knochenfrakturen oder im Mammographie-Screening könnte ein solches Werkzeug wahrscheinlich die Arbeit erleichtern. „Ich glaube nicht, dass die Künstliche Intelligenz uns irgendwann abschaffen wird“, sagt Michael Forstung. „Aber Sie wird uns mit Sicherheit besser machen.“

die am Ende eine autonome Entscheidung treffen soll. Eine Künstliche Intelligenz kann aber mit dem Menschen auf dieser Ebene überhaupt nicht kommunizieren. Sie kann nur sagen, ich habe hier ein Röntgenbild, das bedeutet operieren. Debattieren kann man mit einem Computer nicht, er wird nicht antworten. Insofern droht das in Zukunft noch mehr an Brisanz zu gewinnen.

Haben wir die Entwicklung in einer immer mehr auf Effizienz getrimmten Medizin und in den zunehmend auf wirtschaftliche Ziele ausgerichteten Institutionen überhaupt noch in der Hand? Wird der maschinelle Arzt nicht früher oder später ohnehin kommen, weil er stetiger, geduldiger und womöglich auch billiger ist?

Ich glaube nicht. Bisher ging zumindest in der Medizin der Plan nur sehr selten auf, durch ein Mehr an Technik und Automatisierung Effizienzsteigerungen zu erzielen. In den Krankenhäusern und Praxen hat der technische Fortschritt eher dazu geführt, dass immer mehr gemacht wird. Wir behandeln heute Krankheiten, von denen wir vor zwanzig Jahren noch gar nicht wussten, dass es sie gibt. Die Künstliche Intelligenz wird den Arzt nicht ersetzen, sie wird wenn wir verantwortlich mit ihr umgehen - dazu führen, dass die Medizin noch besser wird.

Die Fragen stellte Michael Brenner

Eckhard Nagel ist geschäftsführender Direktor des Instituts für Management- und Gesundheitswissenschaften der Universität Bayreuth.



Foto: Anshu

Manche Radiologen würden gerne so weit gehen, dass sie beispielsweise Mammographie-Screening-Aufnahmen von KI-Systemen vorsortieren lassen, um selbst nur noch die verdächtigen Bilder zu bearbeiten. Eine gute Idee?

Das ist ja heute schon Realität. Kein Arzt würde sich noch jedes EKG ansehen, die Auswertung der allermeisten dieser Herzuntersuchungen übernimmt längst der Computer. Dagegen ist zunächst einmal nichts einzuwenden. Trotzdem bleibt es weiterhin Aufgabe des Mediziners, zu kontrollieren, ob diese Ergebnisse mit der Realität übereinstimmen. Deshalb muss er eben doch noch selbst den Puls tasten oder die Röntgenbilder sichten, benannt gegenchecken. Auch eine Maschine macht Fehler und hat eine Irrtumswahrscheinlichkeit, und die gilt es so weit wie möglich zu minimieren.

Schon jetzt beklagen viele Patienten das Gefühl, sich in einer Maschinerie zu befinden, auf die sie kaum einwirken können. Droht dieses Problem nicht durch den Einzug der Künstlichen Intelligenz in die Klinik noch größer zu werden?

Die Fülle von Gen-, Labor- und anderen Daten, die wir im Rahmen einer personalisierten Medizin inzwischen über jeden einzelnen Patienten gewinnen, stellt uns tatsächlich schon heute vor die Frage: Ist dieser Ansatz nicht zum Teil gleichbedeutend mit der Auflösung der menschlichen Identität? Sehen wir den Menschen nicht zunehmend als eine Variationsbreite von biologischen Mustern und vergessen zu oft, dass das Individuum auch noch etwas anderes ist: eine eigenständige Persönlichkeit mit spirituellen und psychologischen Aspekten, die ebenfalls entscheidende Auswirkungen auf die Gesundheit haben? Und

den, wäre deshalb gleichbedeutend mit der Idee, in Zukunft statt Piloten nur noch Computer ins Cockpit zu setzen. Die können ja auch fliegen. Kein Mensch würde sich in ein Krankenhaus begeben, in dem er sein Schicksal in die Hände eines Automaten legt - er will sich einen Partner anvertrauen, der als Arzt Verantwortung für und mit ihm übernimmt.

Manche Radiologen würden gerne einen maschinellen Bildwahrer als Berater hinzuziehen oder denken an einen KI-Sicherheitsassistenten, der ihnen bei der Arbeit über die Schulter schaut. Was halten Sie als Mediziner von dieser Idee?

Was wir auf vielen Aufnahmen sehen, sind ja heute schon die Ergebnisse eines Rechners. Und wir vertrauen ihm und gehen davon aus, dass die Bilder, die er aus den Signalen rekonstruiert, der Realität entsprechen. All das kann ausgesprochen hilfreich sein, zum Beispiel bei Operationen. Es gibt ein System, das ihnen vor einem Lebereingriff detailliert anzeigt, wie die Arterien und Venen im Organ verlaufen. Dadurch erhält der Chirurg nicht nur eine Auskunft darüber, wo ein Tumor liegt, sondern auch, welche Blutgefäße er bei seiner Entfernung kapfen muss. Sind es zu viele, kann man den Eingriff oft nicht durchführen, weil ihm das Organ nicht übersehen würde. Aber selbst wenn beim Computer die Anzeige „inoperabel“ blinkt, heißt das ja noch lange nicht, dass der Patient unheilbar ist.

Dann kann man zwar vielleicht den ursprünglich geplanten Eingriff nicht durchführen. Aber man hat immer noch die Möglichkeit, einen anderen therapeutischen Weg zu gehen und zum Beispiel den Tumor anders zu zerstören.

Man könnte man allerdings auch Maschinen mit den Symptomen und der Krankengeschichte füttern...

Das stimmt, und sie würden - entsprechend programmiert - sicherlich auch Vorschläge äußern, welche Diagnose sie für die wahrscheinlichste halten. Nur: Was soll der Betroffene dann anfangen? Zudem ignoriert dieses Vorgehen den wesentlichen Aspekt der Arzt-Patienten-Beziehung, nämlich die Wahrnehmung des Kranken in seiner Situation und die Einordnung von Gefühlen und Informationen in den Gesamtkontext der Befunde. Das macht den Arzt selbst heute, wo die Menge der Messdaten zunehmend unüberschaubar wird, immer wichtiger. Denn irgendetwas muss all diese Dinge zusammenführen und integrieren.

Unter welchen Umständen und in welchem Rahmen wäre es Ihrer Meinung nach in der Medizin vorstellbar, menschliche Intelligenz durch eine künstliche zu ersetzen?

Ich glaube, dass die Technik uns Ärzten helfen kann, an anderer Stelle nicht die Orientierung zu verlieren. Ich kann zum Beispiel bei der Verschreibung eines Medikaments nicht immer alle Neben- und Wechselwirkungen im Kopf haben. Wenn mich ein Computer rechtzeitig davor warnt, es falsch einzusetzen, dann halte ich das für eine hilfreiche und absolut positive Entwicklung. Aber die Maschine kann mir nicht die Entscheidung abnehmen, wie ich bei einem kritischen Patienten auf der Intensivstation weiter vorgehen soll. Hier reicht es nicht, einen Algorithmus zu bemühen.

Worum nicht?

Weil die Maschine dem Arzt nicht die Verantwortung abnimmt. Die Vorstellung, keine Radiologen mehr auszubilden, wäre deutlich kompetenter.