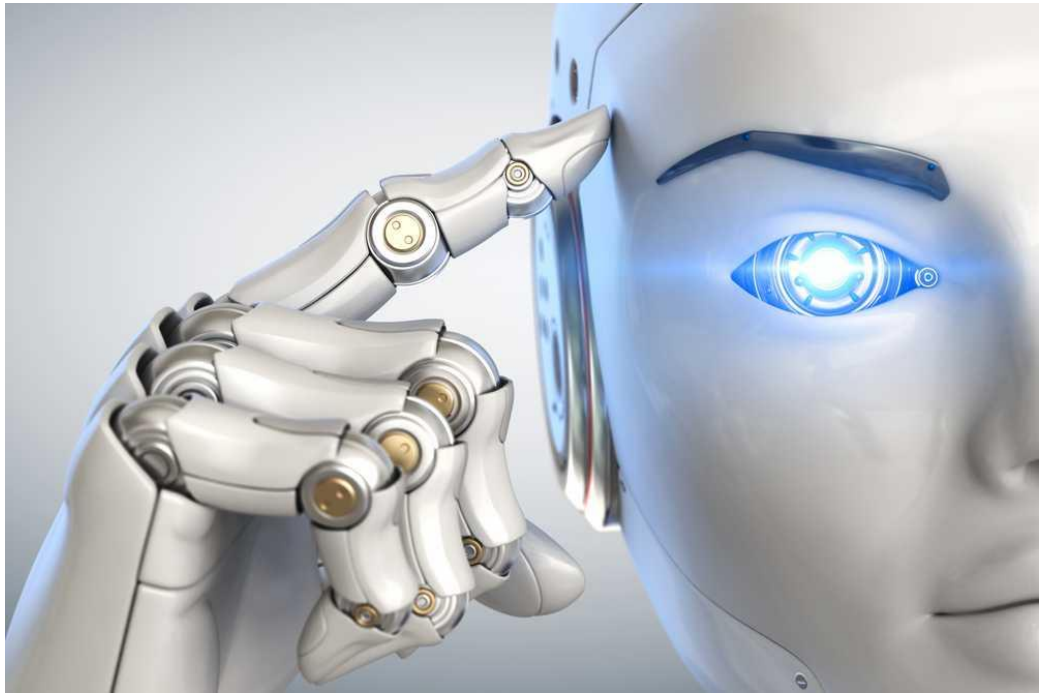
Interaktion Mensch und Maschine im Gesundheitswesen

Michael Forsting

michael.forsting@uk-essen.de

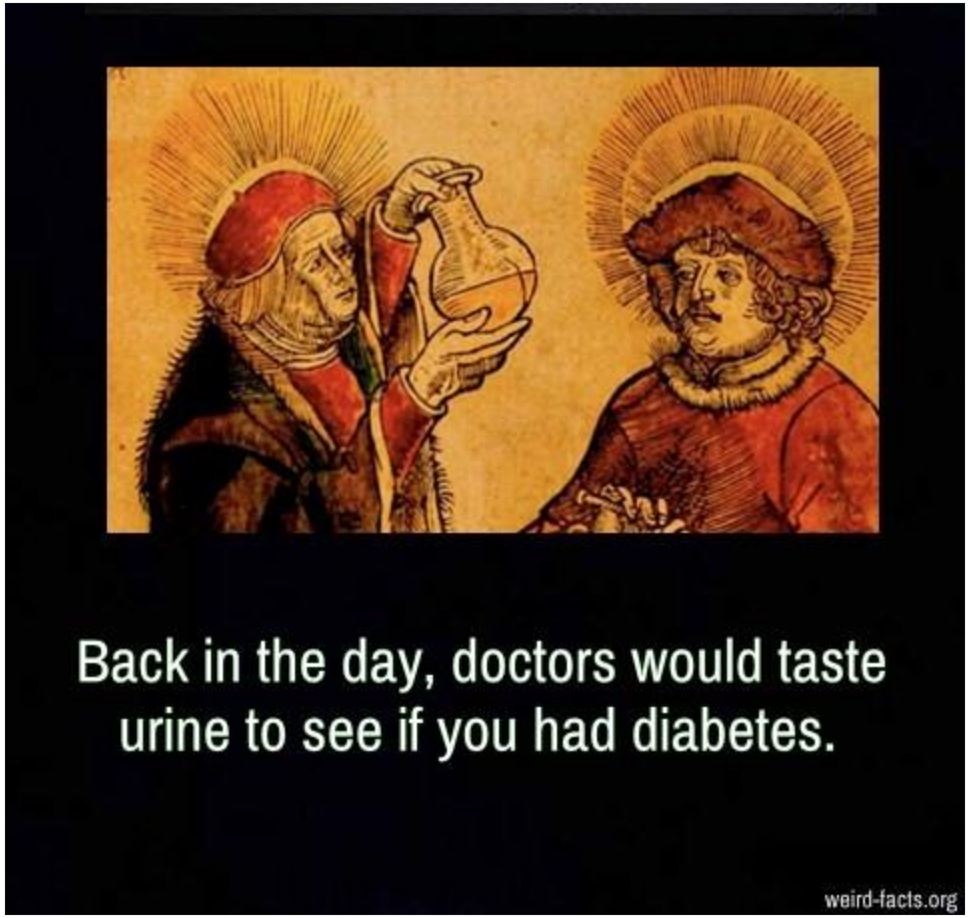
Kurz und knapp



- Unmet needs definieren
- Ground truth ist entscheidend
- Anwendungen in die klinische Routine integrieren
- Produkte auf den Markt bringen

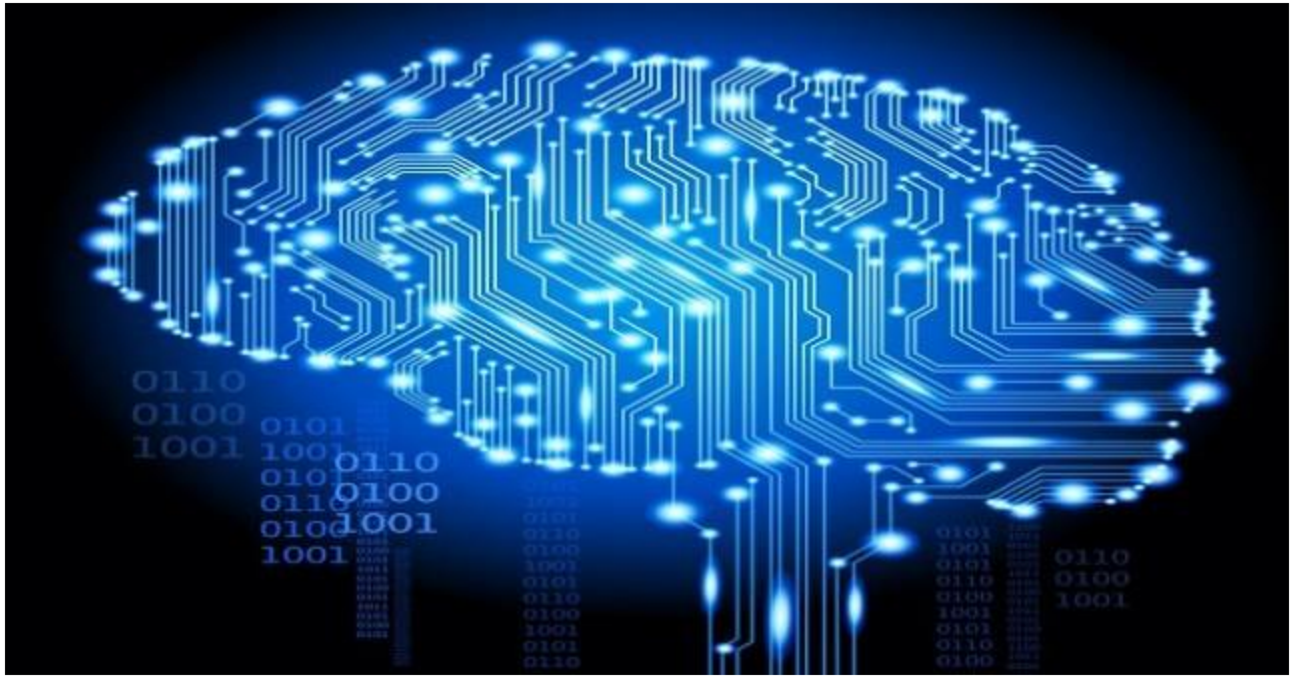
Wollen Sie, dass ein Computer Ihre Krankheit diagnostiziert?

A view into the past

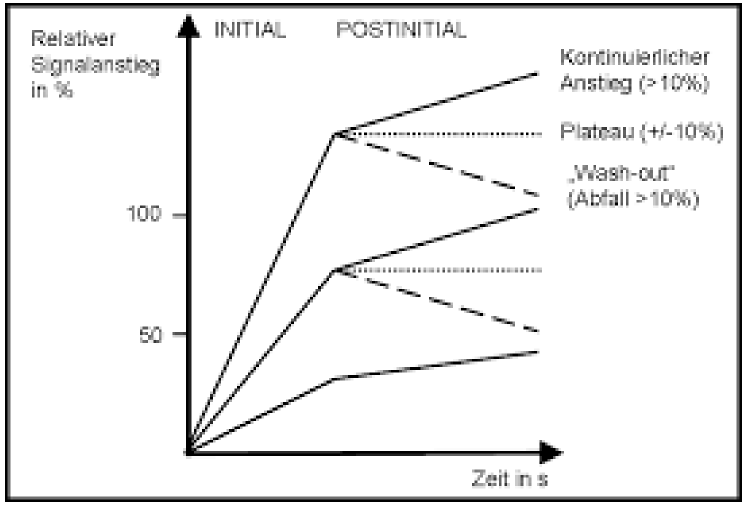


… and sometimes they were wrong

DEEP LEARNING



Was ist der Unterschied zwischen Modellieren und KI?



- Ich brauche eine Hypothese
- Die Hypothese muss mathematisch modelliert werden
- Wenn die Hypothese falsch ist, hilft auch das Modell nicht

Und auch im richtigen Leben funktioniert KI

Bevor wir über Radiologie reden

Wo wird KI die grössten Veränderungen bringen?

communi
care

Longitudinales Curriculum
Sprechende Medizin

- Technische Disziplinen sind fehlerfrei
- „Sprechende Medizin“ ist hypothesengetrieben und macht viele Fehler

Kahneman-Zitat

(Nobelpreisträger für Behavioural Economics)

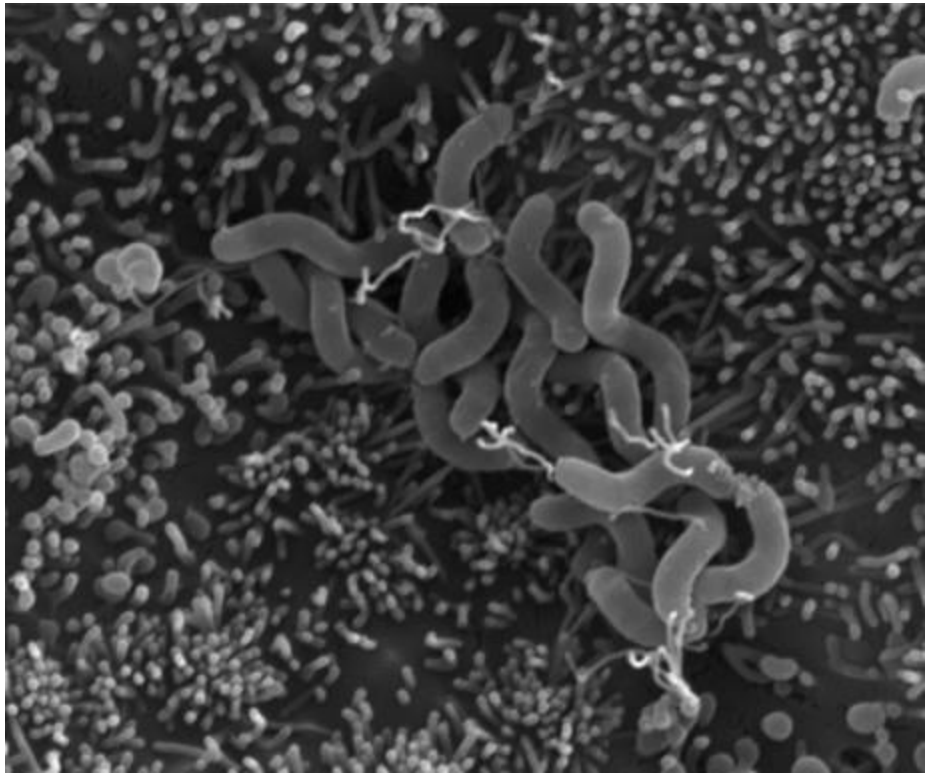
Der Mensch ist nicht in der Lage, statistisch-quantitative Entscheidungen permanent gut genug zu treffen und lässt
sich häufig täuschen (meist durch Optimismus).

Lange Liste von typischen Denkfehlern:
http://en.wikipedia.org/wiki/List_of_cognitive_biases

Die optimistische Annahme: „Wird schon in mein Fachgebiet gehören“

- Kardiologie
- Neurochirurgie
- Orthopädie
- Angiologie
- ….

Sprechende Medizin (Hypothesen-getriebene Medizin) ist extrem fehleranfällig



Das Magengeschwür und Helicobacter pylori

Bis 1984…

- Managerkrankheit
- Zu viel Stress
- Neurotische Persönlichkeit
- Konflikte
- Einstellung

- Proximale selektive Vagotomie
- Beruhigungsmittel
- Rollkuren
- Alternative Medizin:
  - Akupunktur
  - Yoga
  - Entspannung
  - Selbstfindung
  - Reflexzonentherapie

The Lancet · Saturday 16 June 1984

UNIDENTIFIED CURVED BACILLI IN THE STOMACH OF PATIENTS WITH GASTRITIS AND PEPTIC ULCERATION*

BARRY J. MARSHALL J. ROBIN WARREN

Departments of Gastroenterology and Pathology, Royal Perth Hospital, Perth, Western Australia

Summary Biopsy specimens were taken from intact areas of antral mucosa in 100 consecutive consenting patients presenting for gastroscopy. Spiral or curved bacilli were demonstrated in specimens from 58 patients. Bacilli cultured from 11 of these biopsies were gram-negative, flagellate, and microaerophilic and appeared to be a new species related to the genus *Campylobacter.* The bacteria were present in almost all patients with active chronic gastritis, duodenal ulcer, or gastric ulcer and thus may be an important factor in the aetiology of these diseases.

Introduction

GASTRIC spiral bacteria have been repeatedly observed, reported, and then forgotten for at least 45 years.[1-3] In 1940 Freedburg and Barron stated that "spirochaetes" could be found in up to 37% of gastrectomy specimens,[4]- but examination of gastric suction biopsy material failed to confirm these findings.[5] The advent of fibreoptic biopsy techniques permitted biopsy of the antrum, and in 1975 Steer and Colin-Jones observed gram-negative bacilli in 80% of patients with gastric ulcer.[6] The curved bacilli they illustrated were said to be *Pseudomonas,* possibly a contaminant, and the bacteria were once more forgotten. The repeated demonstration of these bacteria in inflamed gastric antral mucosa[7] prompted us to do a pilot study in twenty patients. Typical curved bacilli were present in over half the biopsy specimens and the number of bacteria was closely related to the severity of the gastritis. The present study was designed to confirm the association between antral gastritis and the bacteria, to discover associated gastrointestinal diseases, to culture and identify the bacteria, and to find factors predisposing to infection.

*Based on paper read at Second International Workshop on Campylobacter Infections (Brussels, 1983).

Patients and Methods

Patients

All patients referred for gastroscopy on clinical grounds were eligible for the study which continued until there were 100 participants who gave informed consent and in whom biopsy was considered to be safe. The study was approved by our hospital's human rights committee.

Questionnaire

Where possible patients completed a clinical questionnaire designed to detect a source of infection or show any relationship with "known" causes of gastritis or *Campylobacter* infection, rather than give a detailed account of each patient's history. The emphasis was on animal contact, travel, diet, dental hygiene, and drugs, rather than symptoms.

Endoscopy

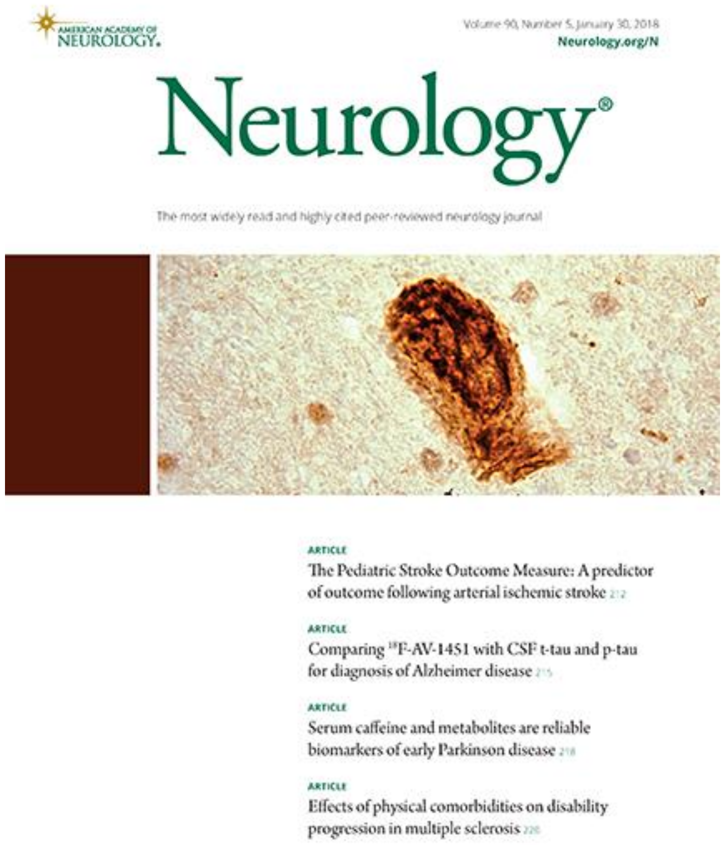
The gastroscopies were done by colleagues at the Royal Perth Hospital. Participants fasted for at least 4 h before endoscopy. An Olympus GIF-K fibreoptic gastroduodenoscope was used. Routine biopsies were done when indicated. For the study two extra specimens were taken from an area of intact antral mucosa, at a distance from any focal lesion such as an antral ulcer. When the mucosa appeared inflamed the specimens were taken from a red area, otherwise any part of the antrum was used. One biopsy was immediately fixed in phosphate-buffered formalin for histological examination, the other was placed in chilled anaerobic transport medium and taken to the microbiology laboratory within 1 h. In a few cases an extra specimen was taken for ultrastructural examination.

The gastroenterologist dictated his report soon after the endoscopy. We had not planned to analyse these reports so a standard terminology was not used and no special attention was paid to minor endoscopic lesions. Findings of doubtful clinical significance, such as mild endoscopic gastritis or duodenogastric bile reflux, may thus have been under-reported. (Hereafter the term "gastritis" refers to a histological grade of chonic gastritis unless stated otherwise.) Before we analysed the data, the endoscopy reports were coded for the major diagnoses.

Histopathology

Sections were stained with haematoxylin and eosin (H & E) and graded for gastritis (by J. R. W.) as 0 (normal), inflammatory cells rarely seen; 1 (normal), lymphoid cells present but within normal limits and with no other evidence of inflammation (see below); 2 (chronic), chronic gastritis; or 3 (active), active chronic gastritis.

AI will change other disciplines even more than radiology



REGULAR ARTICLE | OPEN ACCESS

Instagram photos reveal predictive markers of depression

Andrew G Reece and Christopher M Danforth

EPJ Data Science 2017 **6**:15 | https://doi.org/10.1140/epjds/s13688-017-0110-z
Received: 28 March 2017 | **Accepted:** 22 June 2017 | **Published:** 8 August 2017

The Erratum to this article has been published in *EPJ Data Science* 2017 **6**:21

Abstract

Using Instagram data from 166 individuals, we applied machine learning tools to successfully identify markers of depression. Statistical features were computationally extracted from 43,950 participant Instagram photos, using color analysis, metadata components, and algorithmic face detection. Resulting models outperformed general practitioners' average unassisted diagnostic success rate for depression. These results held even when the analysis was restricted to posts made before depressed individuals were first diagnosed. Human ratings of photo attributes (happy, sad, etc.) were weaker predictors of depression, and were uncorrelated with computationally-generated features. These results suggest new avenues for early screening and detection of mental illness.

Predicting Depression via Social Media

Munmun De Choudhury **Michael Gamon** **Scott Counts** **Eric Horvitz**

Microsoft Research, Redmond WA 98052
{munmund, mgamon, counts, horvitz}@microsoft.com

Abstract

Major depression constitutes a serious challenge in personal and public health. Tens of millions of people each year suffer from depression and only a fraction receives adequate treatment. We explore the potential to use social media to detect and diagnose major depressive disorder in individuals. We first employ crowdsourcing to compile a set of Twitter users who report being diagnosed with clinical depression, based on a standard psychometric instrument. Through their social media postings over a year preceding the onset of depression, we measure behavioral attributes relating to social engagement, emotion, language and linguistic styles, ego network, and mentions of antidepressant medications. We leverage these behavioral cues, to build a statistical classifier that provides estimates of the risk of depression, *before* the reported onset. We find that social media contains useful signals for characterizing the onset of depression in individuals, as measured through decrease in social activity, raised negative affect, highly clustered

laboratory test for diagnosing most forms of mental illness; typically, the diagnosis is based on the patient's self-reported experiences, behaviors reported by relatives or friends, and a mental status examination.

In the context of all of these challenges, we examine the potential of social media as a tool in detecting and predicting affective disorders in individuals. We focus on a common mental illness: Major Depressive Disorder or MDD[1]. MDD is characterized by episodes of all-encompassing low mood accompanied by low self-esteem, and loss of interest or pleasure in normally enjoyable activities. It is also well-established that people suffering from MDD tend to focus their attention on unhappy and unflattering information, to interpret ambiguous information negatively, and to harbor pervasively pessimistic beliefs (Kessler et al., 2003; Rude et al., 2004).

People are increasingly using social media platforms,

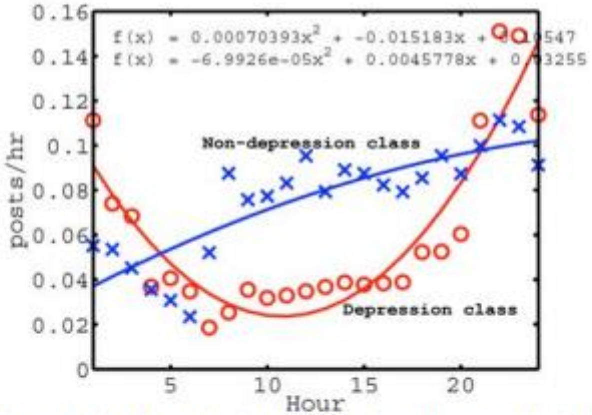


Figure 2: Diurnal trends (i.e. mean number of posts made hourly throughout a day) for the two classes. The line plots correspond to least squares fit of the trends.

FACE PROFILE BY MAREK POLAKOVIC, SK | NOUN PROJECT

Depression detector

Analyzing speech patterns can predict if a subject is depressed.

by Rob Matheson October 23, 2018

To diagnose depression, clinicians interview patients, asking specific questions—about, say, past mental illnesses, lifestyle, and mood.

Machine learning that can detect words and intonations associated with depression could help with diagnostics. But such models tend to predict depression from the person's specific answers to very specific questions.

A new neural-network model developed at MIT can be unleashed on raw text and audio data from interviews to discover speech patterns indicative of depression. Given a new subject, it can accurately predict whether the individual is depressed without needing any other information about the questions and answers.

"The model sees sequences of words or speaking style, and determines that these patterns are more likely to be seen in people who are depressed or not depressed," says EECS graduate student and CSAIL

日本語要約

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

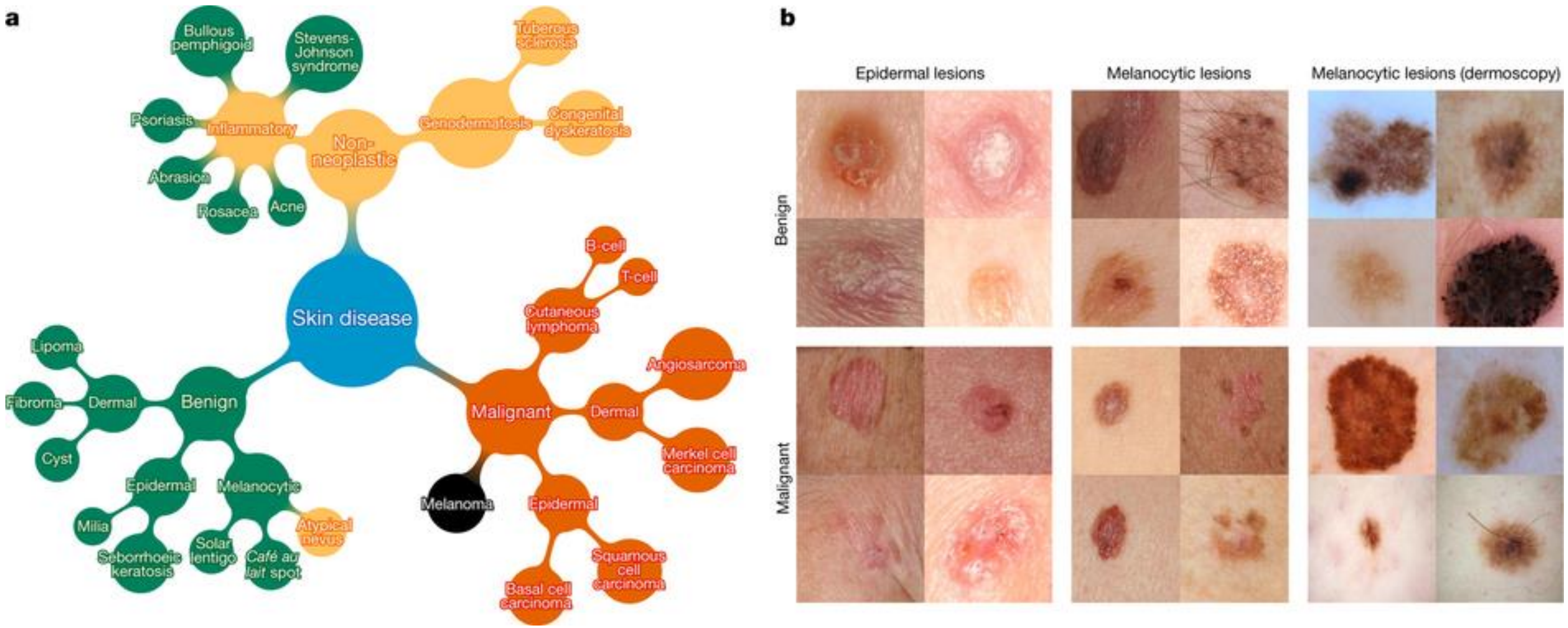
Affiliations | Contributions | Corresponding authors

Nature **542**, 115–118 (02 February 2017) | doi:10.1038/nature21056
Received 28 June 2016 | Accepted 14 December 2016 | Published online 25 January 2017

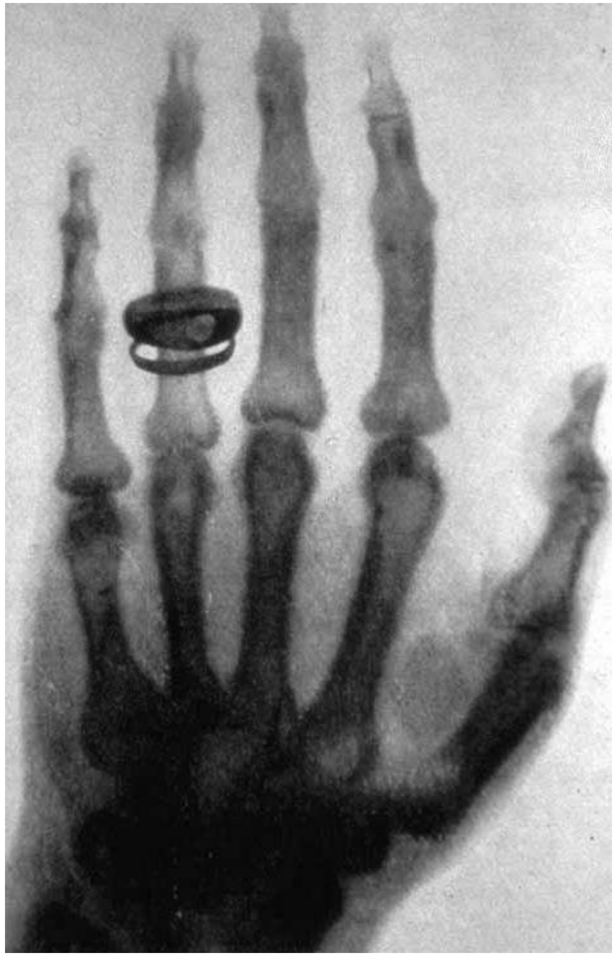
PDF | Citation | Reprints | Rights & permissions | Article metrics

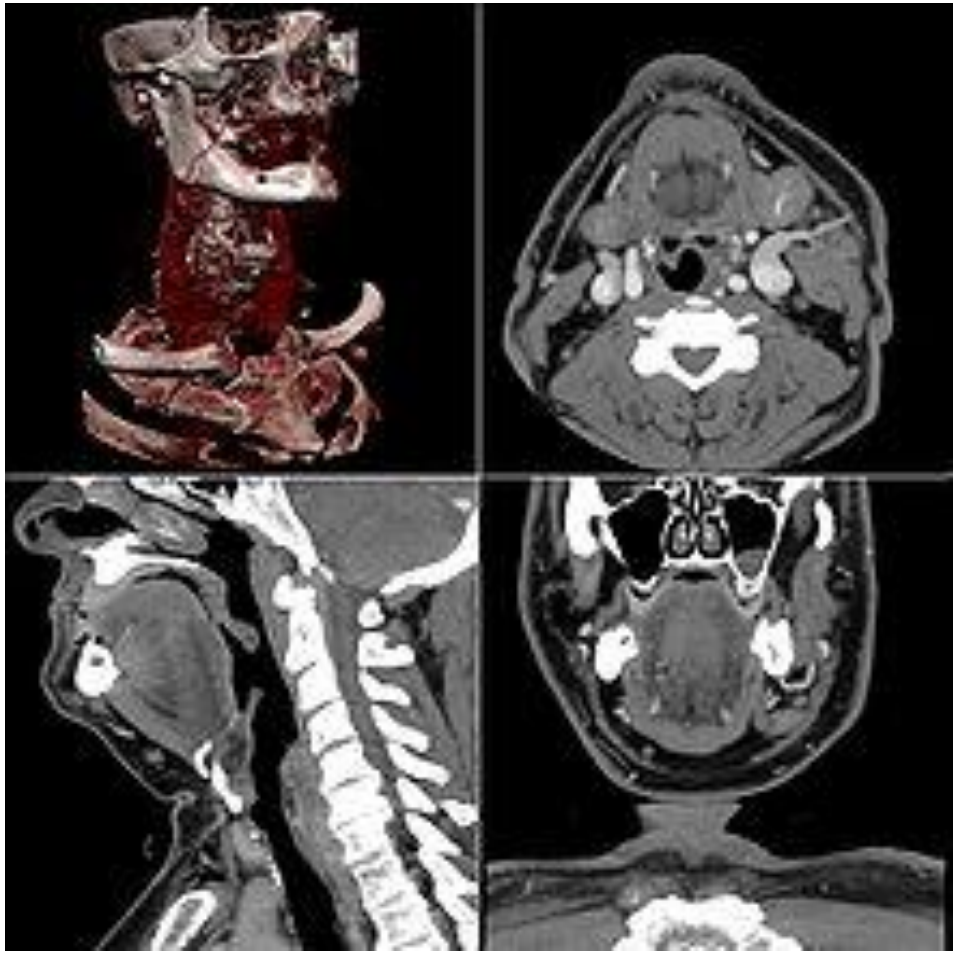
Skin cancer, the most common human malignancy[1,2,3], is primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis, a biopsy and histopathological examination. Automated classification of skin lesions using images is a challenging task owing to the fine-grained variability in the appearance of skin lesions. Deep convolutional neural networks (CNNs)[4,5] show potential for general and highly variable tasks across many fine-grained object categories[6,7,8,9,10,11]. Here we demonstrate classification of skin lesions using a single CNN, trained end-to-end from images directly, using only pixels and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets[12]—consisting of 2,032 different diseases. We test its performance against 21 board-certified dermatologists on biopsy-proven clinical images with two critical binary classification use cases: keratinocyte carcinomas versus benign seborrheic keratoses; and malignant melanomas versus benign nevi. The first case represents the identification of the most common cancers, the second represents the identification of the deadliest skin cancer. The CNN achieves performance on par with all tested experts across both tasks, demonstrating an artificial intelligence capable of classifying skin cancer with a level of competence comparable to dermatologists. Outfitted with deep neural networks, mobile devices can potentially extend the reach of dermatologists outside of the clinic. It is projected that 6.3 billion smartphone subscriptions will exist by the year 2021 (ref. 13) and can therefore potentially provide low-cost universal access to vital diagnostic care.

25 top US Dermologists against AI system

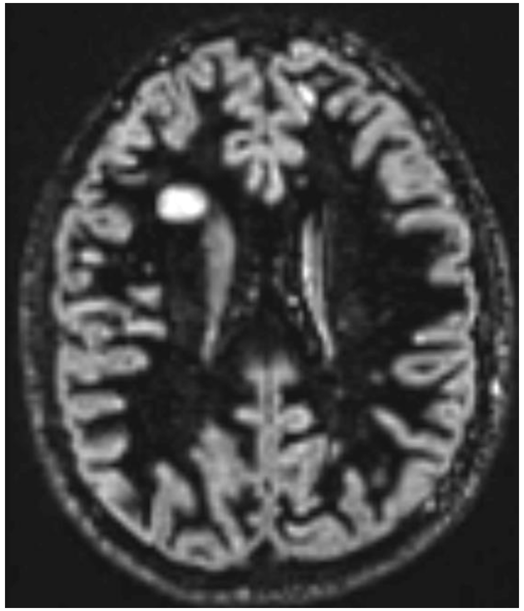


Und jetzt reden wir über Radiologie, weil ….

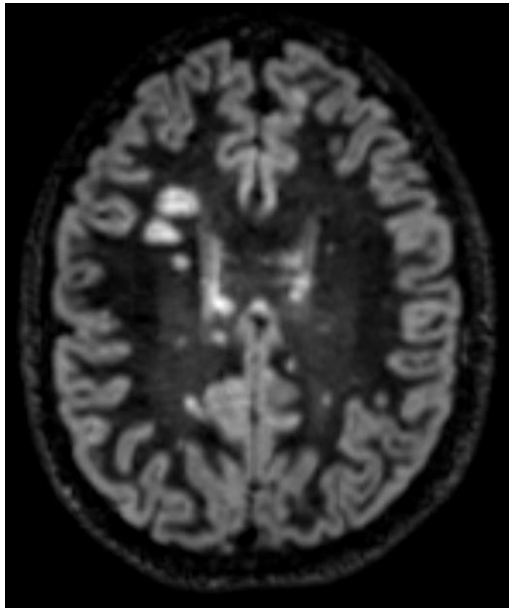




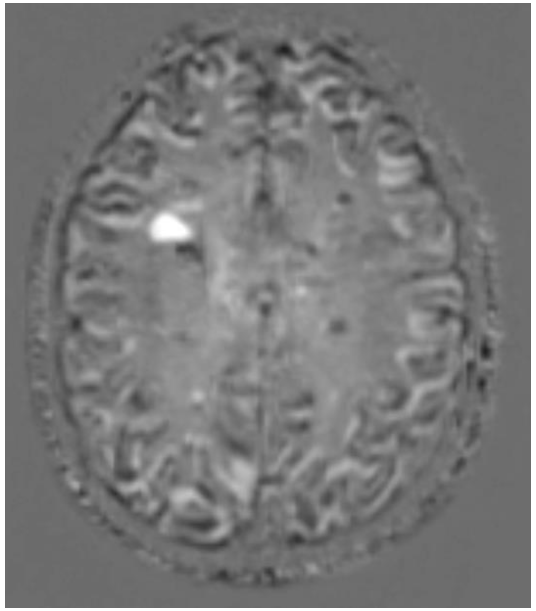
Subtraktion von Double Inversion Recovery-Aufnahmen für die MS-Verlaufsbeurteilung



23y/m mit
RRMS unter
Tecifidera

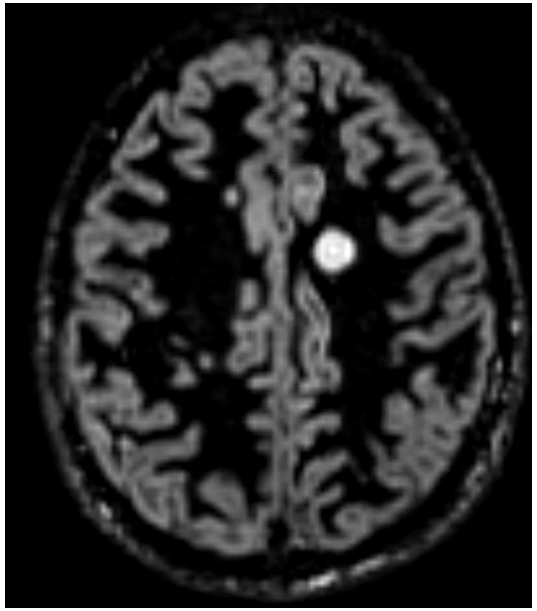


Mavenclad-
Therapiebeginn

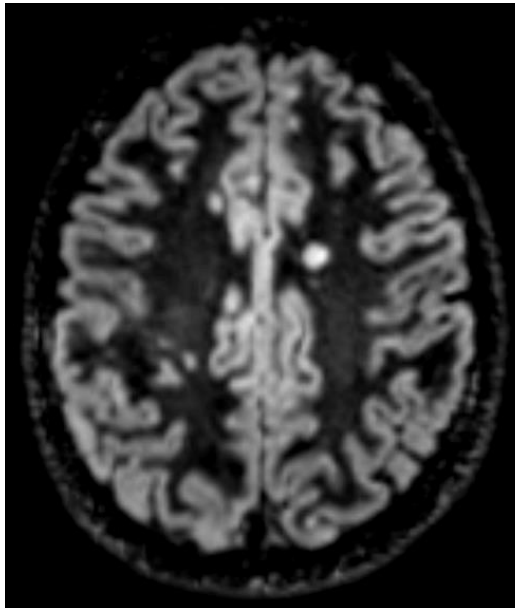


DIR-Subtraktion
Neuer MS-Herd
hyperintens

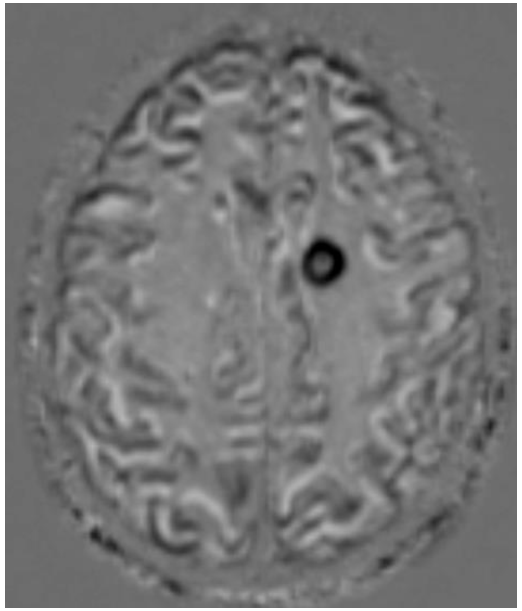
Subtraktion von 3D Double Inversion Recovery-Aufnahmen für die MS-Verlaufsbeurteilung



27.10.2017
23y/m mit
RRMS unter
Tecifidera



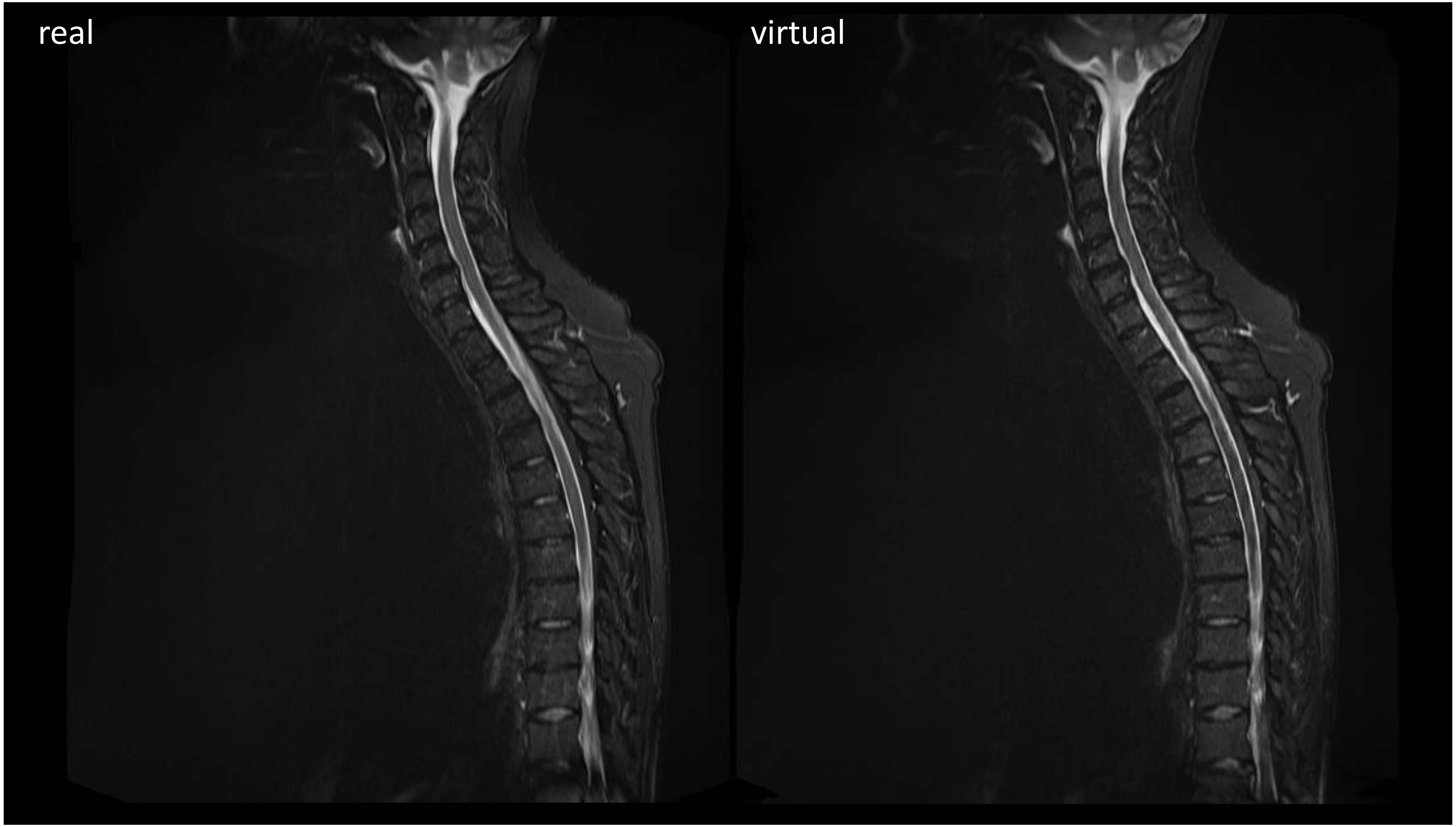
01.03.2018
Mavenclad-
Therapiebeginn



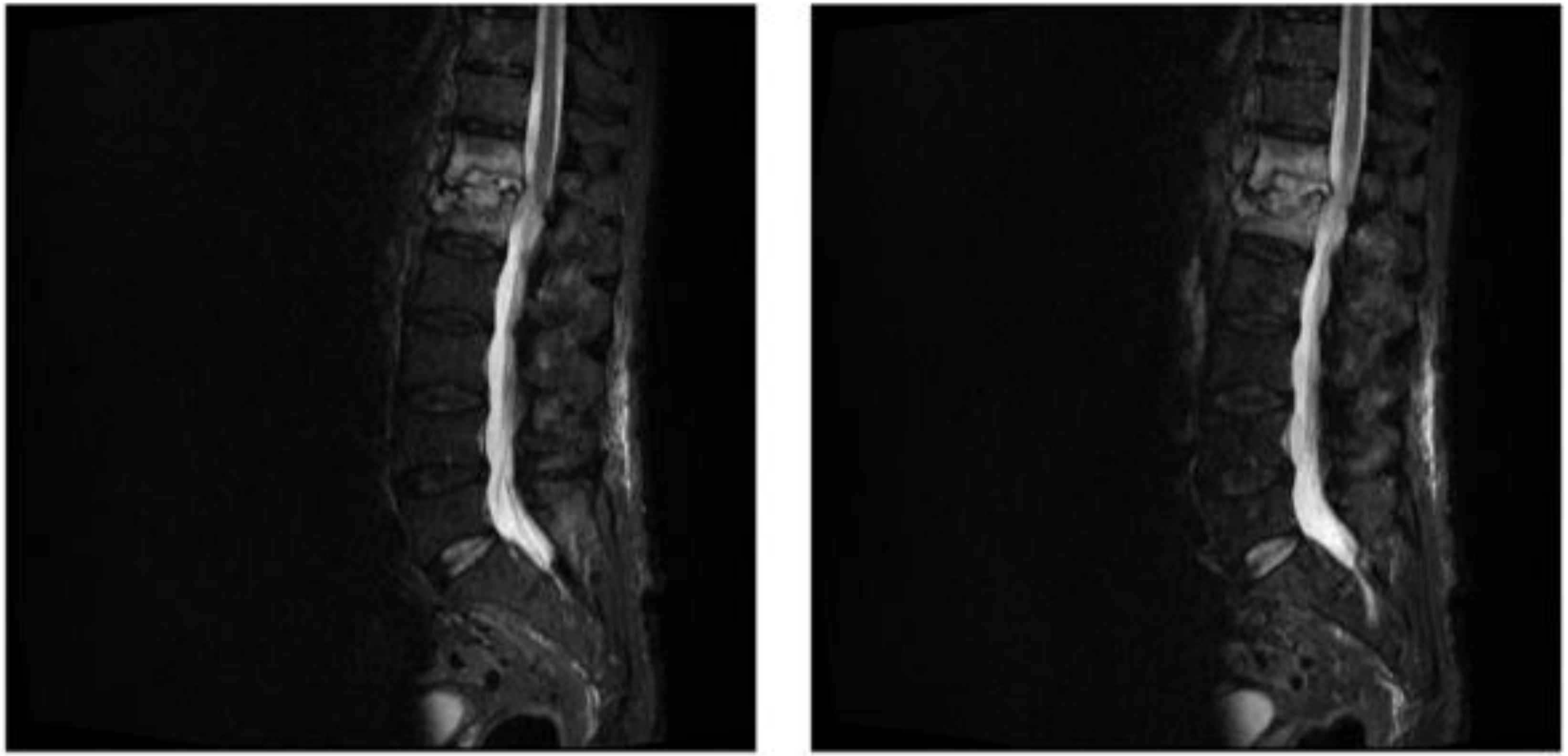
DIR-Subtraktion
abnehmender MS-Herd
hypointens

Was will der Neurologe wissen?

Virtual STIR

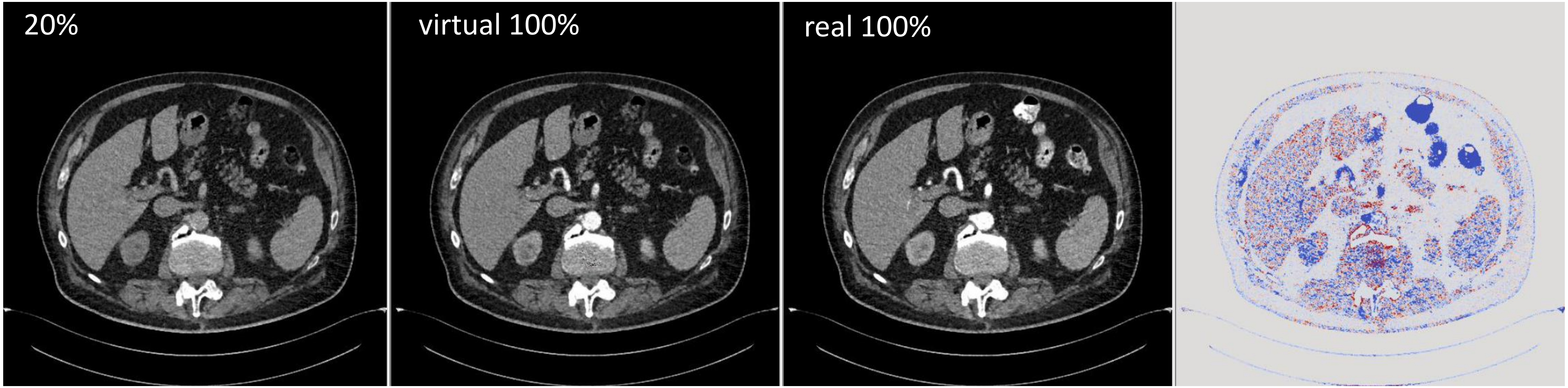


Virtual sequences (STIR): Reduction of time; reduction of problems in standardisation

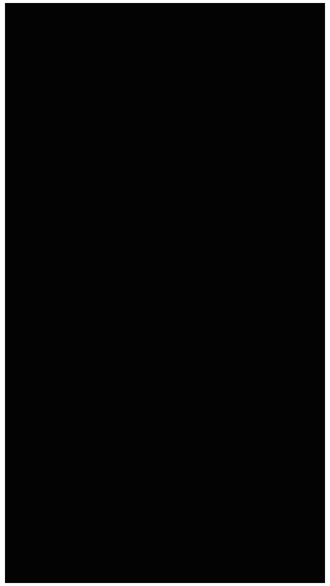


vSTIR: Which image is real?

Contrast media dose reduction

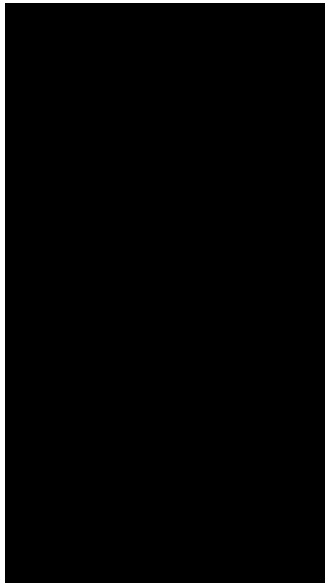


Virtual Contrast Agent Prediction Normal



T1w
native

ceT1w
groundtruth

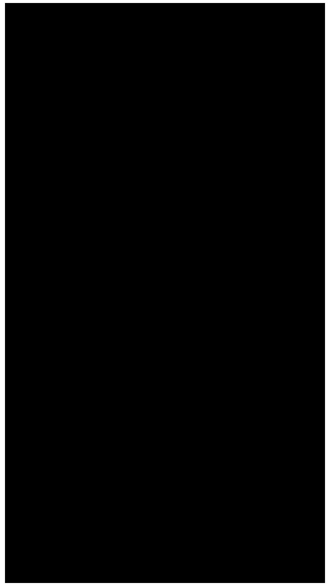


vT1w
predicted

Virtual Contrast Agent Prediction Tumor

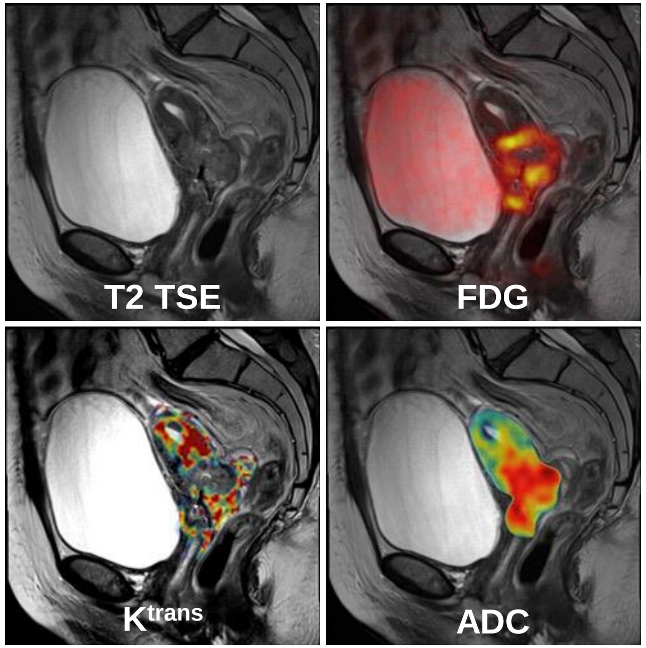
T1w
native

ceT1w
groundtruth



vT1w
predicted

PET/MRT Radiomics

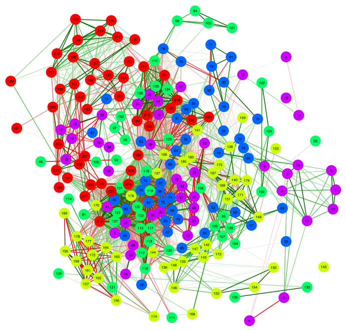


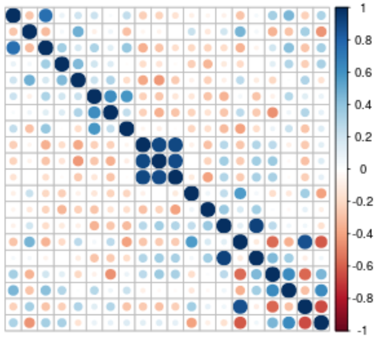
M-stage

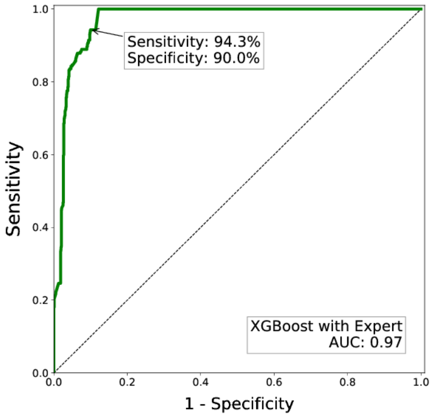
Name	P	P_{adj}
T2_TSE_glcm_differenceEntropy	0.0000	0.00
T1_Nativ_glcm_differenceEntropy	0.0002	0.02
T2_TSE_glcm_correlation	0.0005	0.05
Ktrans_1616_calc_min	0.0009	0.09
T2_TSE_glcm_energy	0.0019	0.19
T2_TSE_glrlm_LRE	0.0019	0.19
ADC_glcm_contrast	0.0083	0.82
PET_glcm_correlation	0.0083	0.82
PET_glcm_IDMN	0.0098	0.97
ADC_glcm_IDMN	0.0115	1.00
ADC_glcm_inverseVariance	0.0156	1.00
T1_Nativ_glcm_correlation	0.0210	1.00
Ktrans_1616_glcm_cShade	0.0210	1.00
T1_Nativ_glrlm_LRLGLE	0.0241	1.00
PET_glcm_contrast	0.0241	1.00
ADC_calc_uniformity	0.0277	1.00
PET_glrlm_LRE	0.0277	1.00
Ktrans_1616_calc_median	0.0299	1.00
T2_TSE_calc_skewness	0.0317	1.00
ADC_calc_entropy	0.0317	1.00

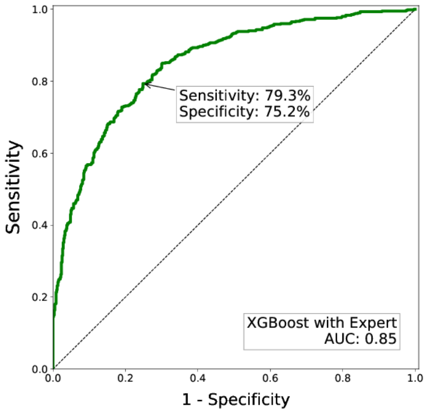
N-stage

Name	P	P_{adj}
PET_calc_median	0.0001	0.01
T2_TSE_glcm_correlation	0.0005	0.05
PET_calc_energy	0.0008	0.08
PET_glrlm_LRHGLE	0.0008	0.08
ADC_glcm_autoCorrelation	0.0013	0.13
Ktrans_1616_glcm_cShade	0.0013	0.13
ADC_calc_skewness	0.0016	0.16
T2_TSE_glcm_differenceEntropy	0.0023	0.22
ADC_glcm_cProminence	0.0023	0.22
ADC_glcm_mean	0.0027	0.26
ADC_glcm_cShade	0.0032	0.31
Ktrans_1616_glcm_maxProb	0.0037	0.37
Ktrans_1616_glrlm_GLN	0.0043	0.43
ADC_glrlm_LRLGLE	0.0059	0.58
PET_calc_uniformity	0.0068	0.68
PET_calc_max	0.0068	0.68
T1_Nativ_glrlm_LRLGLE	0.0079	0.78
T2_TSE_glrlm_LRE	0.0079	0.78
PET_glcm_IDMN	0.0079	0.78
Ktrans_1616_glrlm_LRLGLE	0.0079	0.78



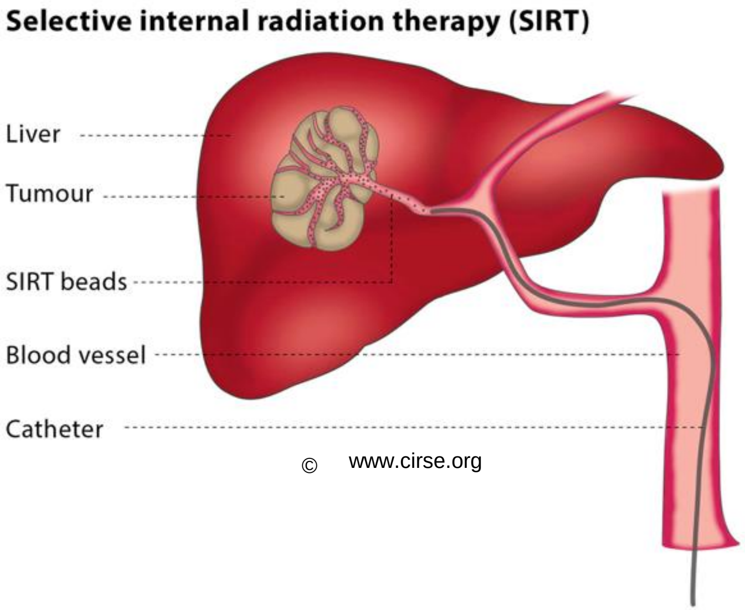






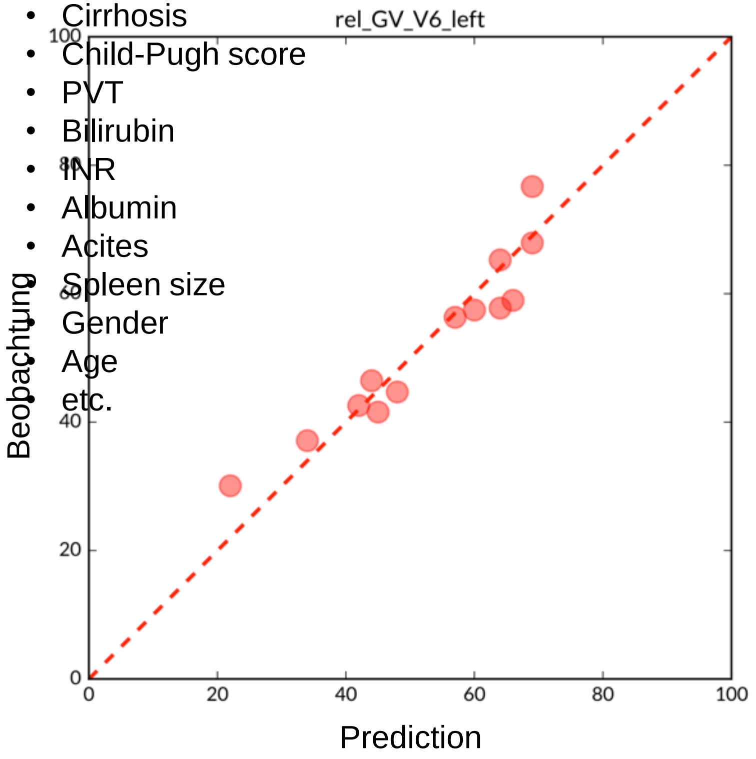
Umutlu L, Grüneisen J, Nensa F, et al. 2017. Unpublished

Selective internal radiation therapy

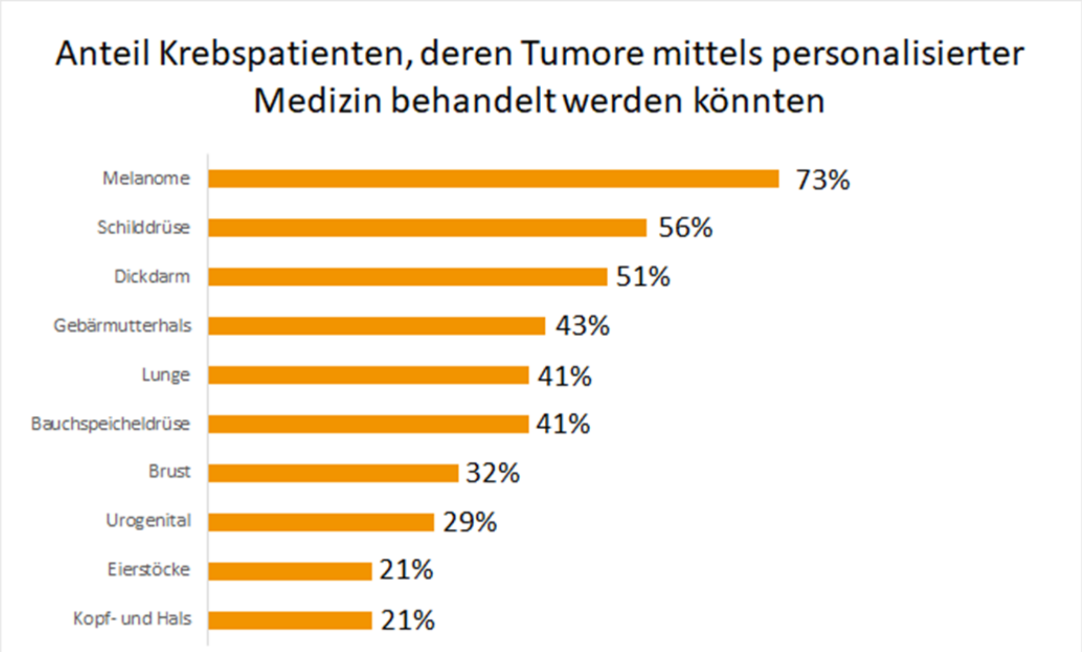


- Liver size
- Cirrhosis
- Child-Pugh score
- PVT
- Bilirubin
- INR
- Albumin
- Acites
- Spleen size
- Gender
- Age
- etc.

Growth of untreated liver part



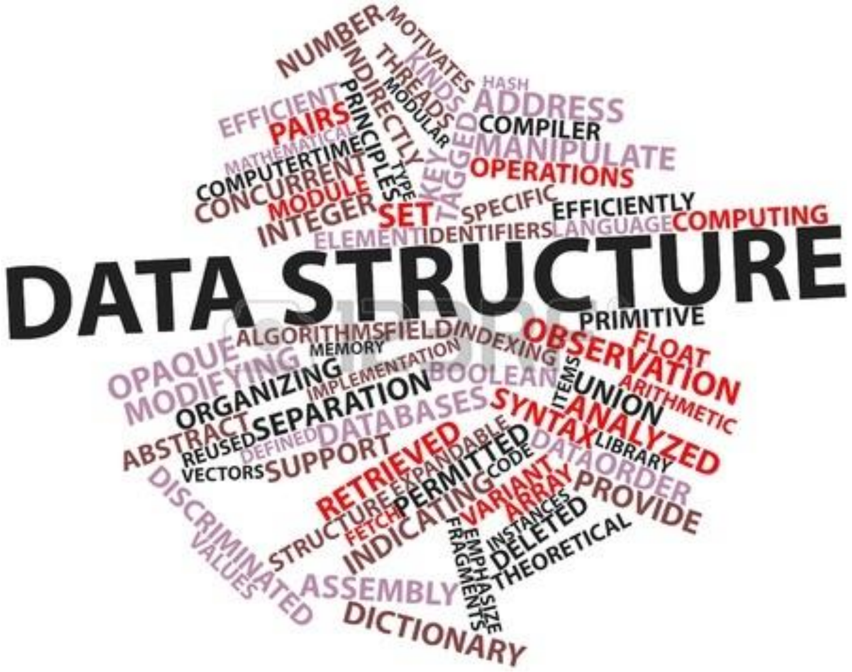
Megatrends in der Medizin



Modifiziert nach Statistika

Warum ist das alles noch nicht in der Medizin?

Daten und Wörter müssen zueinander finden



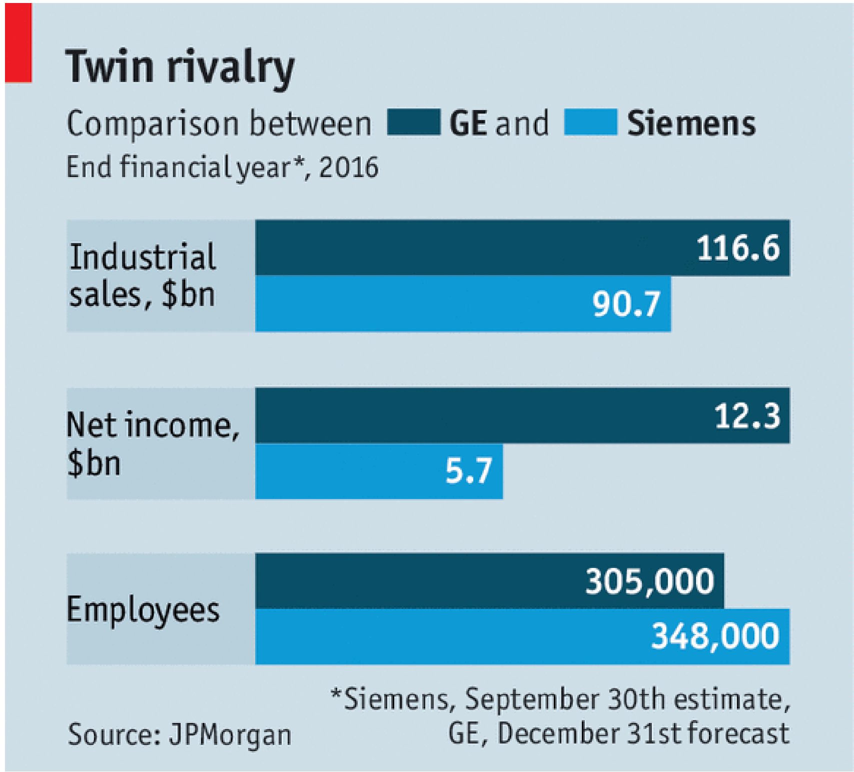
Wir brauchen valide Daten!

garbage in, garbage out

Like 20 Tweet 0 +1 0

Often abbreviated as GIGO, this is a famous computer axiom meaning that if invalid data is entered into a system, the resulting output will also be invalid. Although originally applied to computer software, the axiom holds true for all systems, including, for example, decision-making systems.

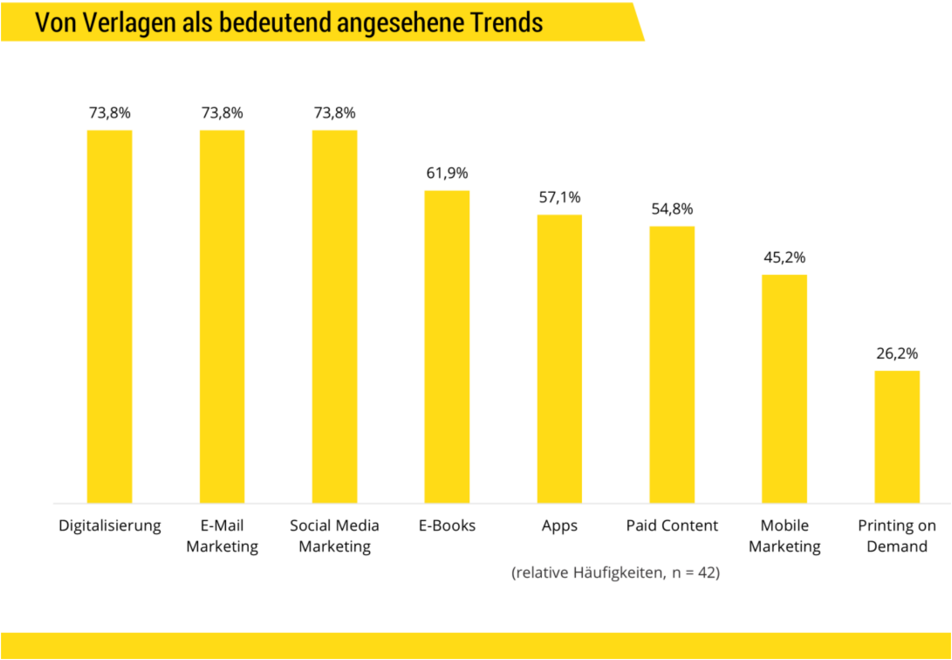
Wer wird es in die Medizin bringen?



Economist.com

- Siemens, GE….
  - Haben viele Daten, aber keine validen Daten

Wer wird es in die Medizin bringen?



- Thieme, Springer…
  - Haben Lehrbuchwissen, aber auch nicht genug Daten, um Systeme zu trainieren

Wer wird es in die Medizin bringen?

06

Qualität statt Quantität

Wenn wir von Informationen als dem neuen „schwarzen Gold" sprechen, könnte man meinen, das Ziel sei die Erschließung der meisten Ölquellen und damit der meisten Informationsquellen, die uns mit Informationen in großer Menge versorgen. Aber der Erfolg liegt sicherlich nicht in der Quantität, sondern in der Qualität - also der Auswertung, Filterung und Verfeinerung. Ohne diese „Raffinerie" ist das schwarze Gold nur sehr begrenzt nutzbar. Dies zeigt sich am Beispiel der Standard Oil Company, die nicht mehr bzw. nicht mehr in der gleichen Form und unter diesem Namen existiert. Exxon Mobile, eines der größten Unternehmen der Welt, ist das größte Überbleibsel der einst gigantischen Standard Oil Company.

Das Unternehmen wurde seinerzeit von John D. Rockefeller gegründet. Durch seinen Ehrgeiz und mithilfe des nötigen Glücks konnte Rockefeller für sich den amerikanischen Traum verwirklichen - vom praktisch mittellosen jungen Mann zum reichsten Mann der USA. Und natürlich hatte er ein besonderes Gespür für echte Chancen. Rockefeller war außerdem zur rechten Zeit am rechten Ort. Als die ersten Ölquellen erschlossen wurden, war er Buchhalter und arbeitete an den Hafenanlagen. So hatte er Einblick in die Art und Weise der Festsetzung von Ölpreisen.

Rockefeller war ein intelligenter junger Mann mit Geschäftssinn und erkannte die Chance, die

Grosse Kliniken

- Haben Daten
- Können die Daten annotieren
- Können Daten aggregieren
- Können Daten veredeln
- Können Systeme trainieren
- Kennen die Fragestellungen
- Brauchen strategische Partner

Ärzte Zeitung online, 28.02.2019

„Super-Diagnostics“

Uniklinikum Essen setzt voll auf Künstliche Intelligenz

Das Uniklinikum Essen richtet ein eigenes KI-Institut mit vier neuen Professuren ein. Mittelfristig will Essen kleineren Kliniken, aber auch onkologischen Praxen dienen – als Dienstleister onkologischer „Super-Diagnostics“.

Von Matthias Wallenfels

Kommentieren (0)

suchen...

Erfolgs-Rezept Praxis-Preis 2018

Welcher ihrer Favoriten hat gewonnen?
Die zehn besten Bewerber um den Erfolgs-Rezept Praxis-Preis stehen fest. Sie konnten mit über Ihren Favoriten online abstimmen. Lesen Sie, auf wen jetzt die Entscheidung gefallen ist. mehr »

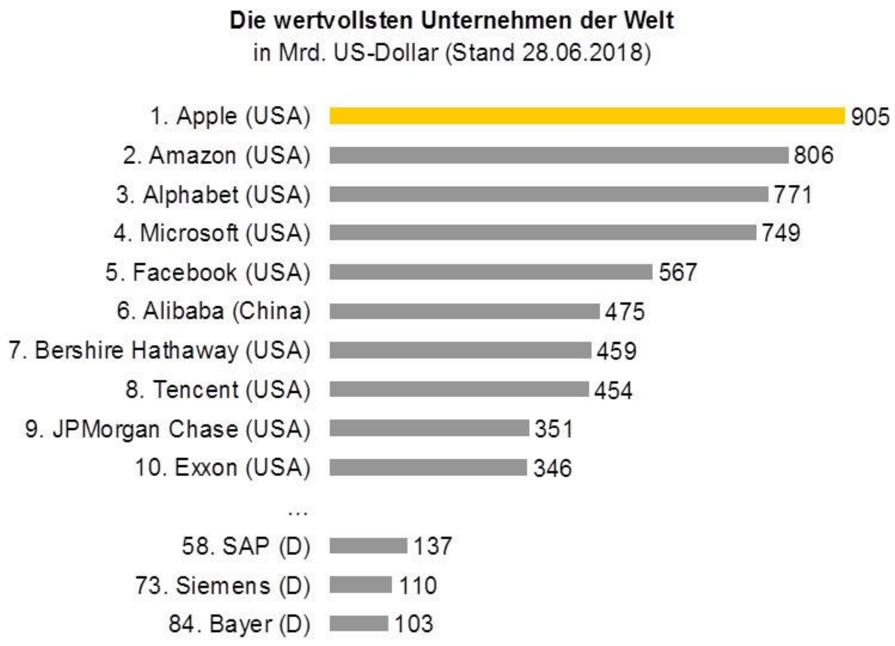
Geomarketing

Ärzte und Delegation

Die Neigung von Ärzten, Arbeit an NäPA, VERAH oder MFA zu delegieren, variiert stark. Wie groß die Unterschiede sind, zeigt unsere interaktive Karte. Ein kostenloser Service der „Ärzte Zeitung“.


SIEMENS
SOMATOM Force
ELEKTRONISCHE PATIENTENAKTE
POCKEN
BAKTERIEN
IMPFUNG
INDIEN
VIREN
THERAPIE
MEDIZIN
VIROLOGIE
ELEKTRONENMIKROSKOP
TABAKMOSAIKVIRUS
MIKROBIOLOGIE
PROTOZOEN
KRANKHEITSPRAEVENTION
fotolia
by Adobe

Wer wird es in die Medizin bringen?

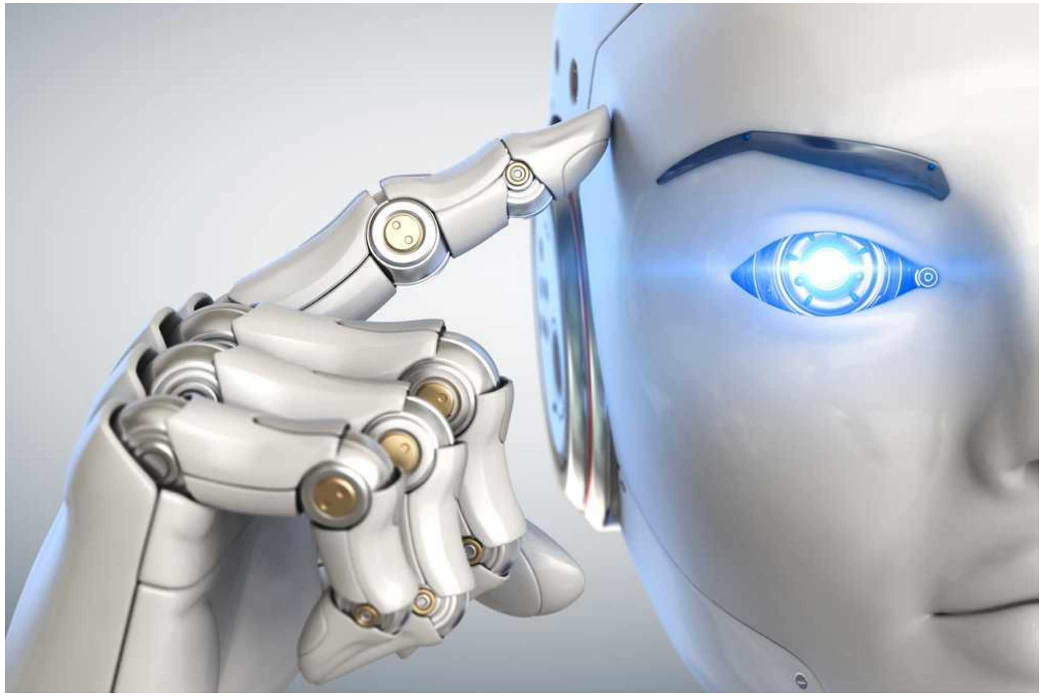


- Google, Amazon….
  - Haben bislang keine Daten zum Training der KI-Systeme
  - Werden Koalitionen eingehen, aber auch dann ein Qualitätsproblem mit den Daten haben
  - Werden selber in den Krankenhausmarkt einsteigen

Und nun?

- Technisierung hat sich in der Medizin in der Vergangenheit schon bewährt
- KI ist ausserhalb der Medizin schon oft Routine
- KI braucht Digitalisierung und wird deshalb zunächst in die technischen Fächer kommen
- KI wird die „sprechende Medizin“ deutlich mehr verbessern als die technische Medizin
- Spannend bleibt, wer die besten KI-Anwendungen in die Medizin bringt

Kurz und knapp



- Unmet needs definieren
- Ground truth ist entscheidend
- Anwendungen in die klinische Routine integrieren
- Produkte auf den Markt bringen