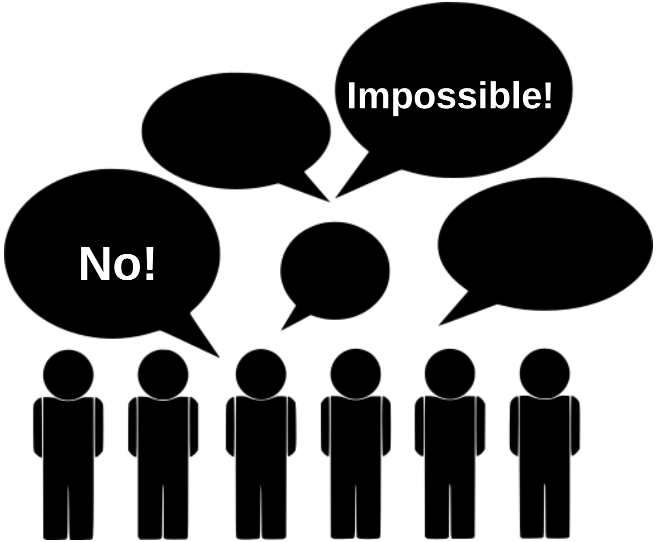
Künstliche Intelligenz in der Medizin: offene ethische Fragen einer Revolution

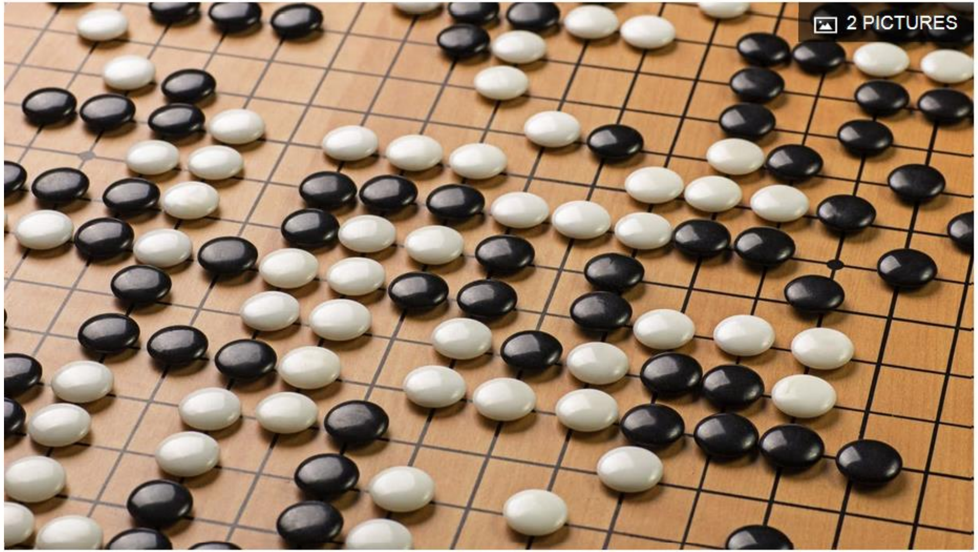
Markus Herrmann & Lena Maier-Hein
National Center for Tumor Diseases (NCT) &
German Cancer Research Center (DKFZ)

AI takes every hurdle – AlphaGo



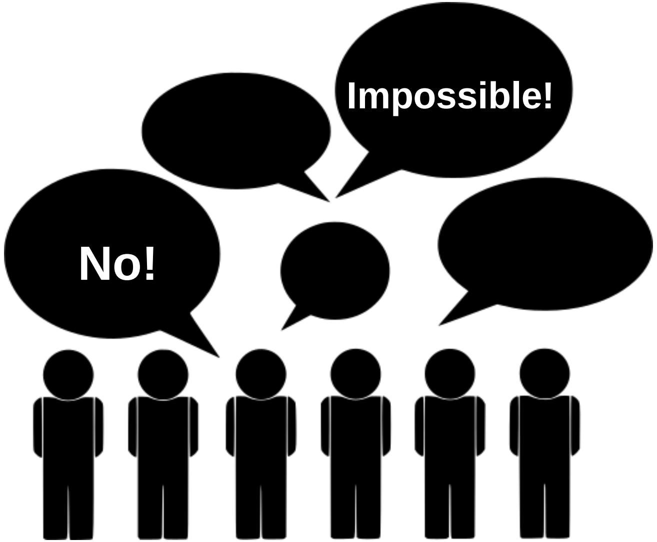
Google's AI beats world's top-ranking Go player

Michael Irving | May 24th, 2017

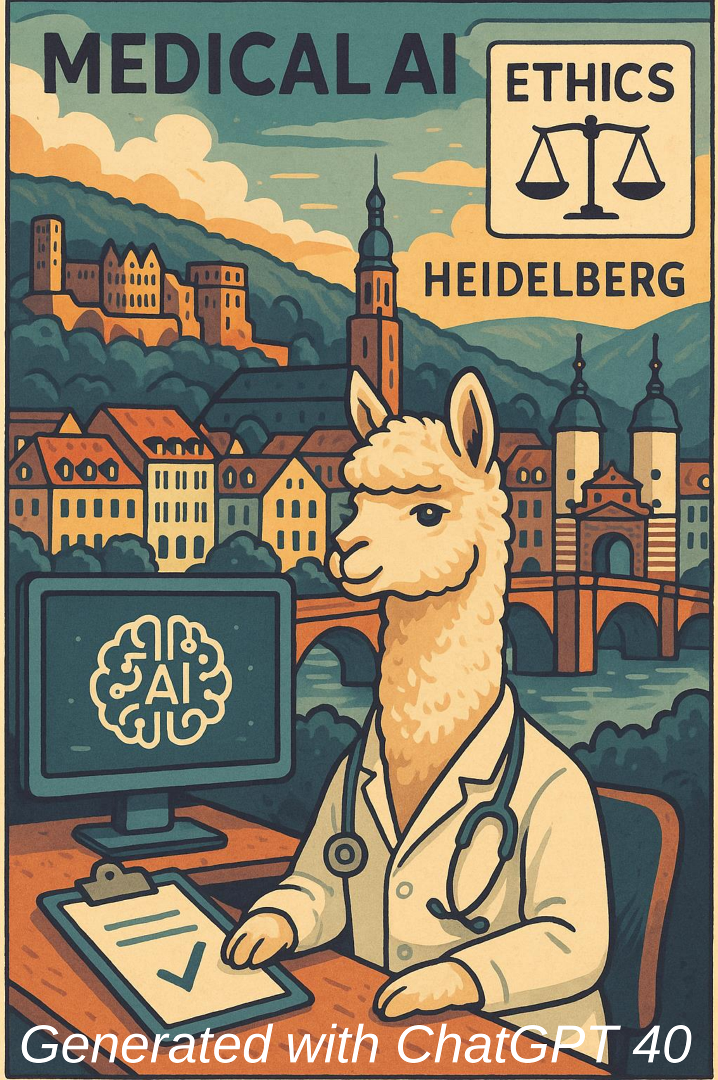


Google's AlphaGo AI system has beaten the world's top-ranking Go player in the first of three games (Credit: Zerbor/Depositphotos)

AI takes every hurdle – Creativity

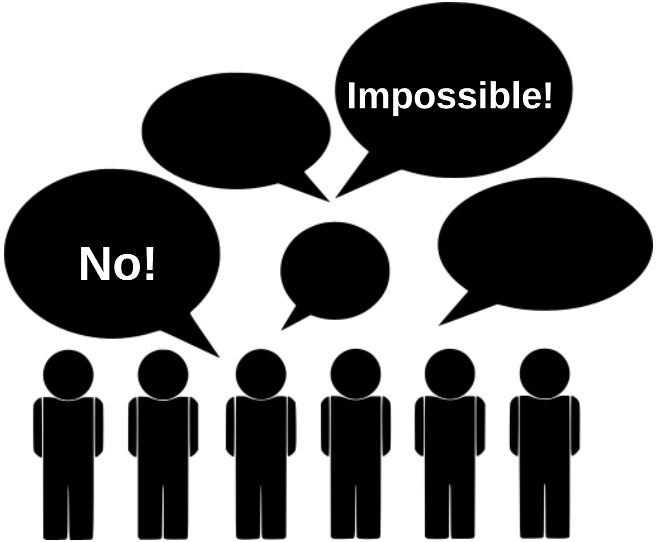


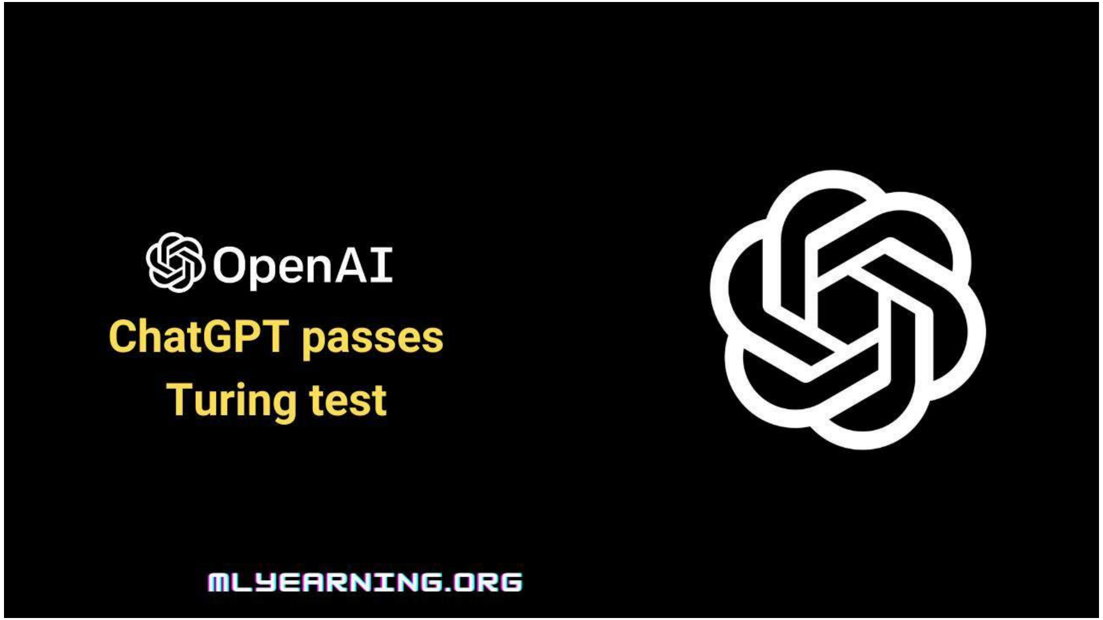
ChatGPT: Please create an artwork combining the themes of Medical AI, Alpaca, Heidelberg and Ethics



Generated with ChatGPT 4o

AI takes every hurdle – Turing Test





AI takes every hurdle – Turing Test

RESEARCH ARTICLE | ECONOMIC SCIENCES

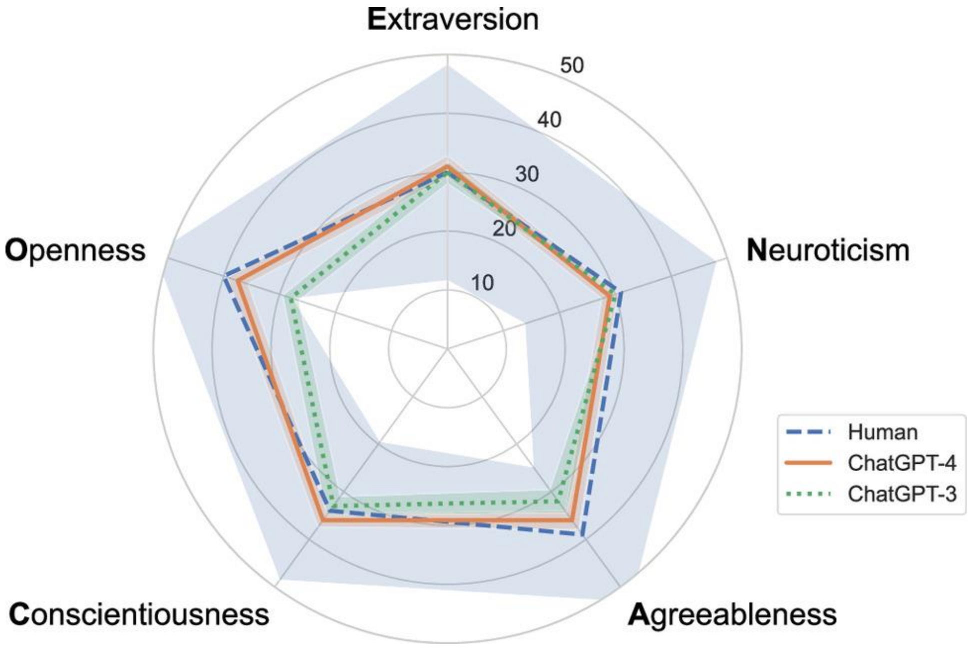
A Turing test of whether AI chatbots are behaviorally similar to humans

Qiaozhu Mei, Yutong Xie, Walter Yuan, and Matthew O. Jackson Authors Info & Affiliations

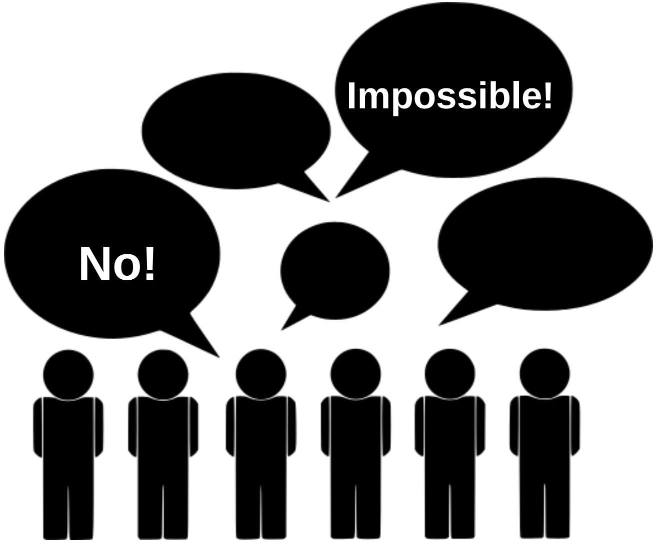
Contributed by Matthew O. Jackson; received August 12, 2023; accepted January 4, 2024; reviewed by Ming Hsu, Juanjuan Meng, and Arno Riedl

February 22, 2024 | 121 (9) e2313925121 | https://doi.org/10.1073/pnas.2313925121

“ChatGPT-4 exhibits behavioral and personality traits that are statistically indistinguishable from a random human from tens of thousands of human subjects from more than 50 countries.”



AI takes every hurdle – Emotions



AI has better 'bedside manner' than some doctors, study finds

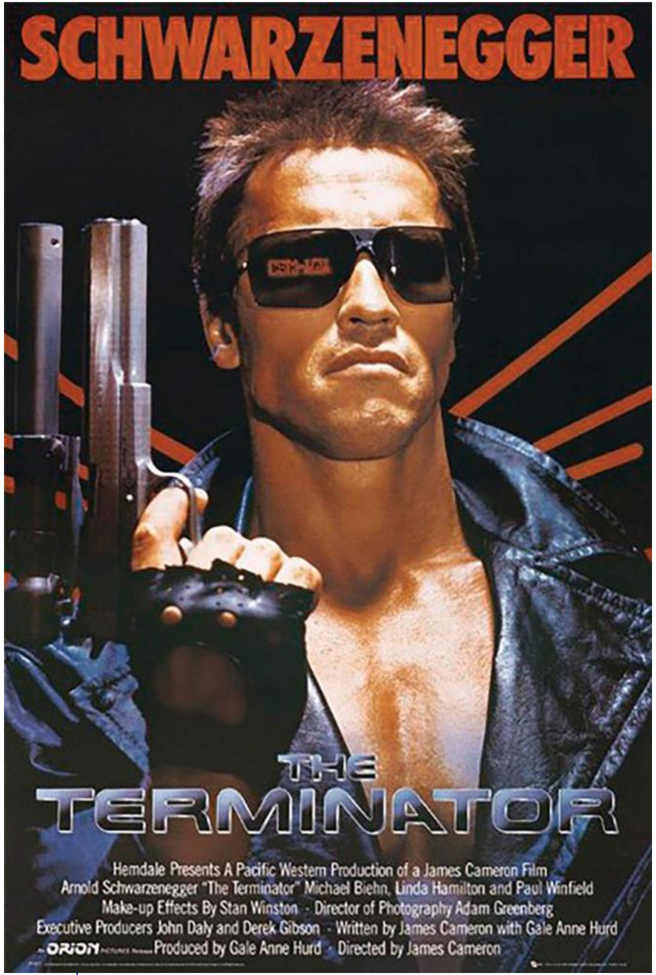
ChatGPT rated higher in quality and empathy of written advice, raising possibility of medical assistance role

A panel of healthcare professionals preferred ChatGPT's responses to medical questions over those of a doctor 79% of the time. Photograph: Ariel Skelley/Getty Images

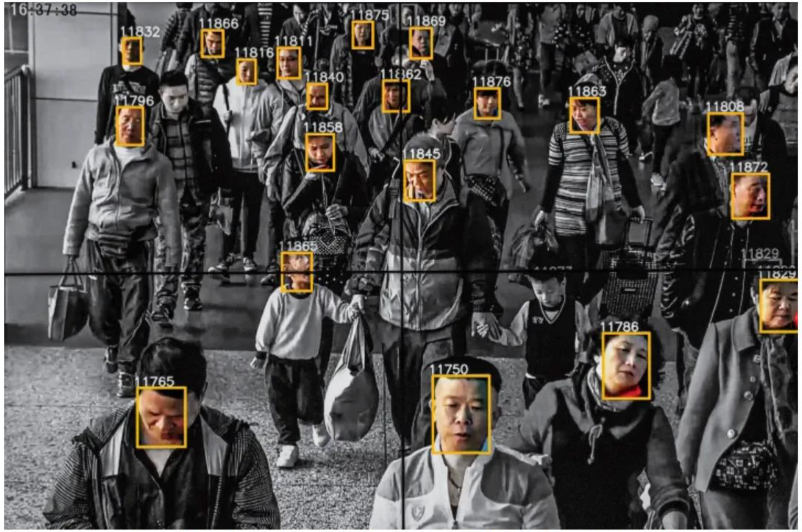
ChatGPT appears to have a better 'bedside manner' than some doctors - at least when their written advice is rated for quality and empathy, a study has shown.

Stop it?

Is it ethical to work in AI in the first place?


SCHWARZENEGGER
THE TERMINATOR
Hemdale Presents A Pacific Western Production of a James Cameron Film
Arnold Schwarzenegger "The Terminator" Michael Biehn, Linda Hamilton and Paul Winfield
Make-up Effects By Stan Winston · Director of Photography Adam Greenberg
Executive Producers John Daly and Derek Gibson · Written by James Cameron with Gale Anne Hurd
ORION
Produced by Gale Anne Hurd · Directed by James Cameron

TIME | DAVOS 2019



Facial recognition is one element of China's expanding tracking efforts Photo-Illustration by TIME: Source Photo: Gilles Sabrié—The New York Times/Redux

How China Is Using "Social Credit Scores" to Reward and Punish Its Citizens

By Charlie Campbell / Chengdu

NCT

?

ZEITUNG MEHR F.A.Z. Frankfurter Allgemeine Abo

Wissen > KI und Massenvernichtungswaffen: Wie real ist die Gefahr?

RISIKEN VON KI

ChatGPT, eine Biowaffe bitte!

Von Piotr Heller 20.11.2024, 07:23 Lesezeit: 8 Min.

The role of academia in AI?

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > nature index > article

NATURE INDEX | 18 September 2024

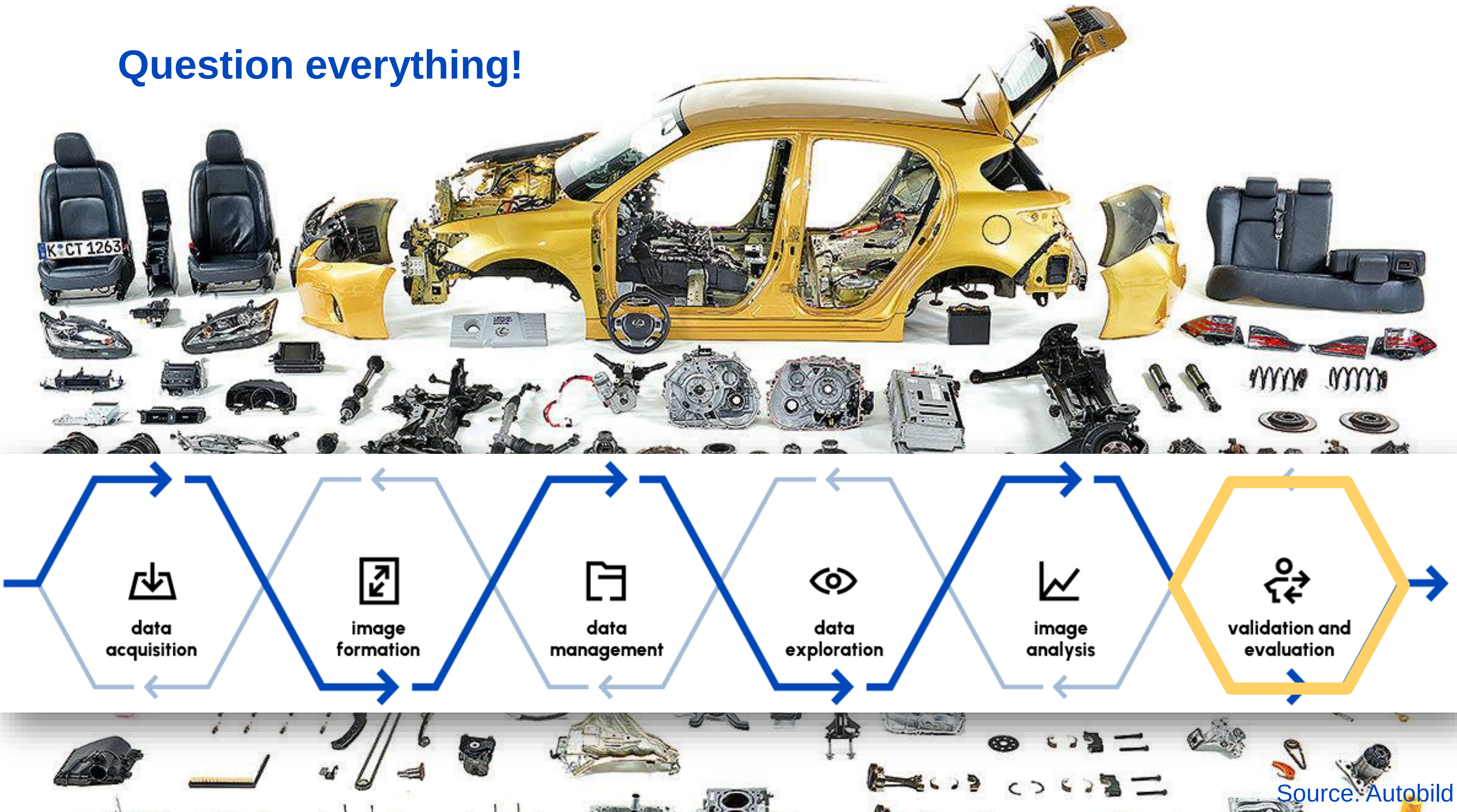
Rage against machine learning driven by profit

Industry research funding is vastly eclipsing academia's spend, but healthy development demands broad input.

> **"Academia is the only place where researchers still have the ability to work without an obvious roadmap to profit."**

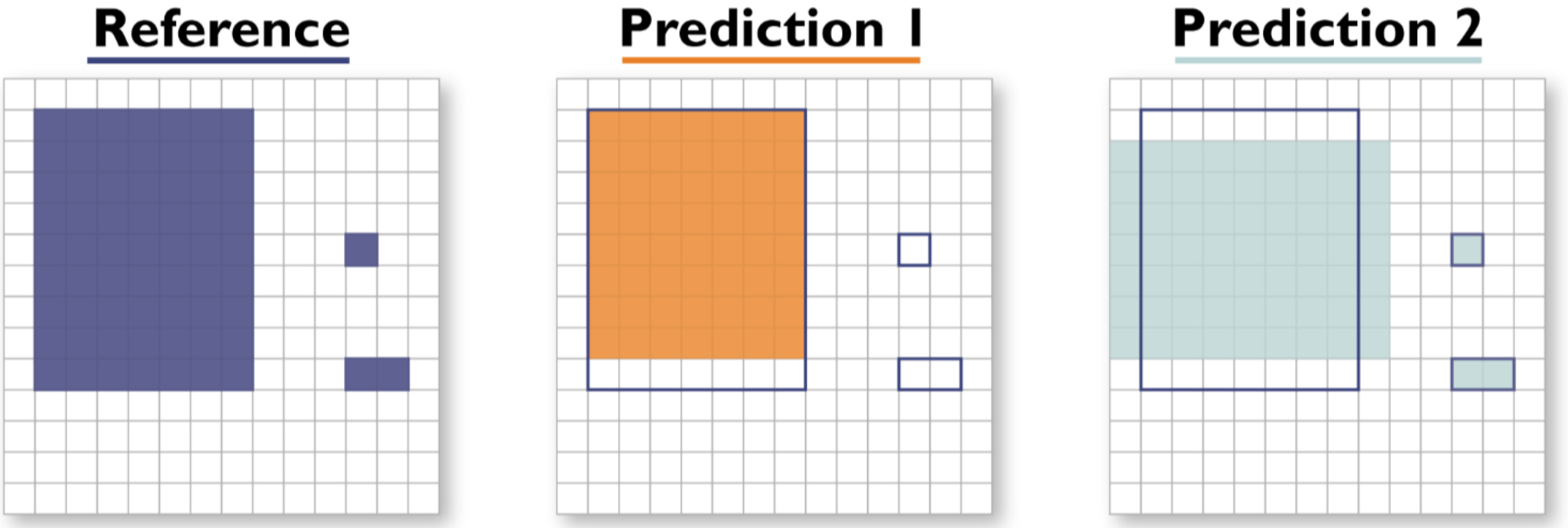
dkfz.

Question everything!



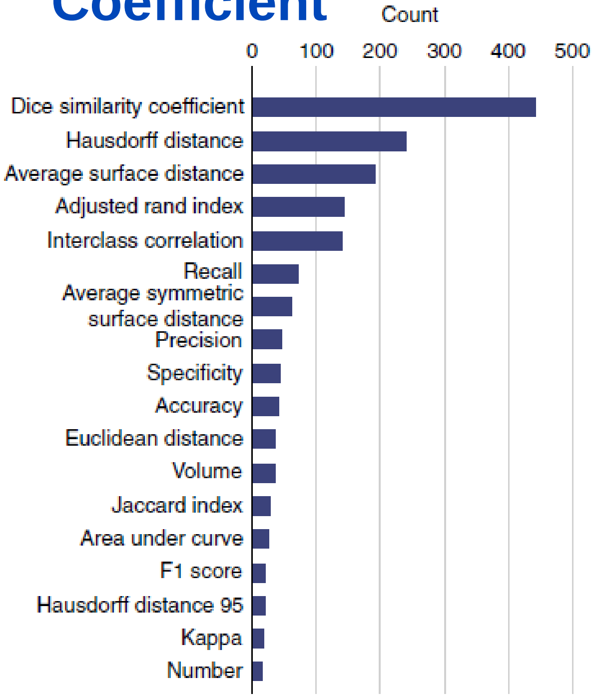
Source: Autobild

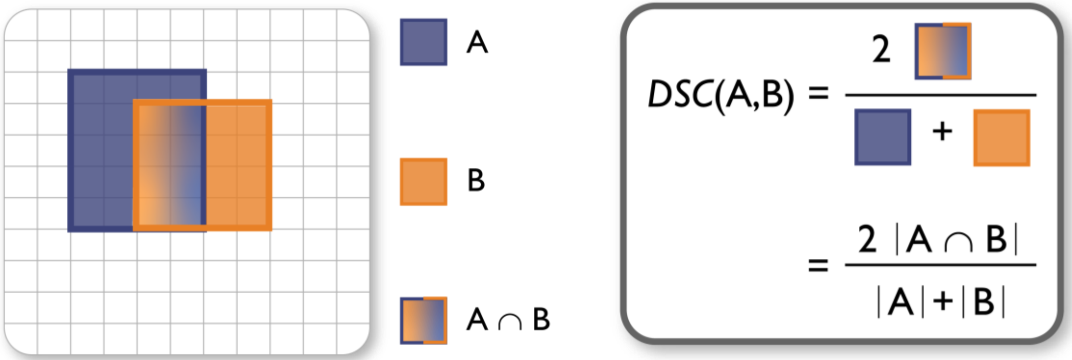
Which algorithm performs better (orange or light blue)?



Reinke/Tizabi, …, Maier-Hein. Understanding metric-related pitfalls in image analysis validation. **Nature Methods 2024**

Most widely used metric in challenges: Dice Similarity Coefficient

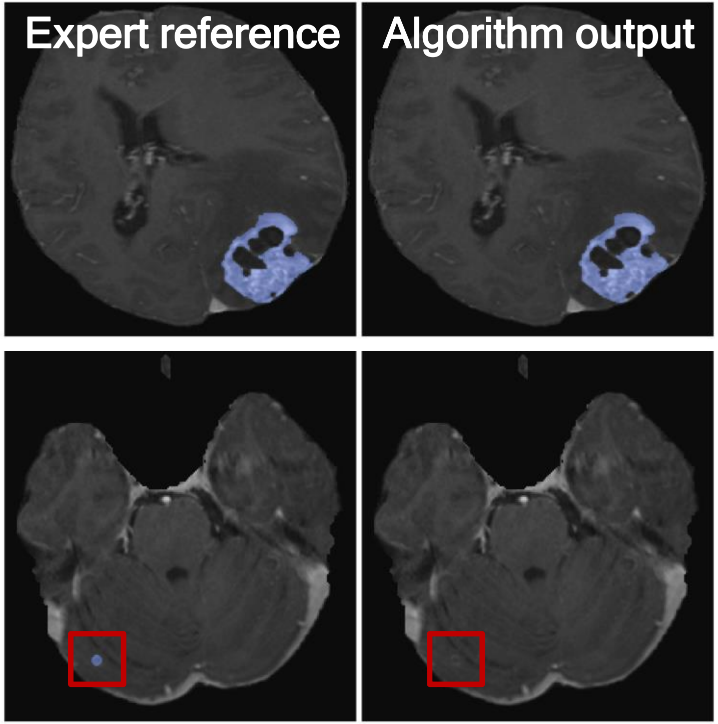




Maier-Hein et al. Why rankings of biomedical image analysis competitions should be interpreted with care **Nature Commun. 2018**
Reinke, …, Maier-Hein. Common Limitations of Image Processing Metrics: A Picture Story. **ArXiv 2021**

Flawed AI validation: A worldwide problem

Algorithm with expert performance according to common validation metric

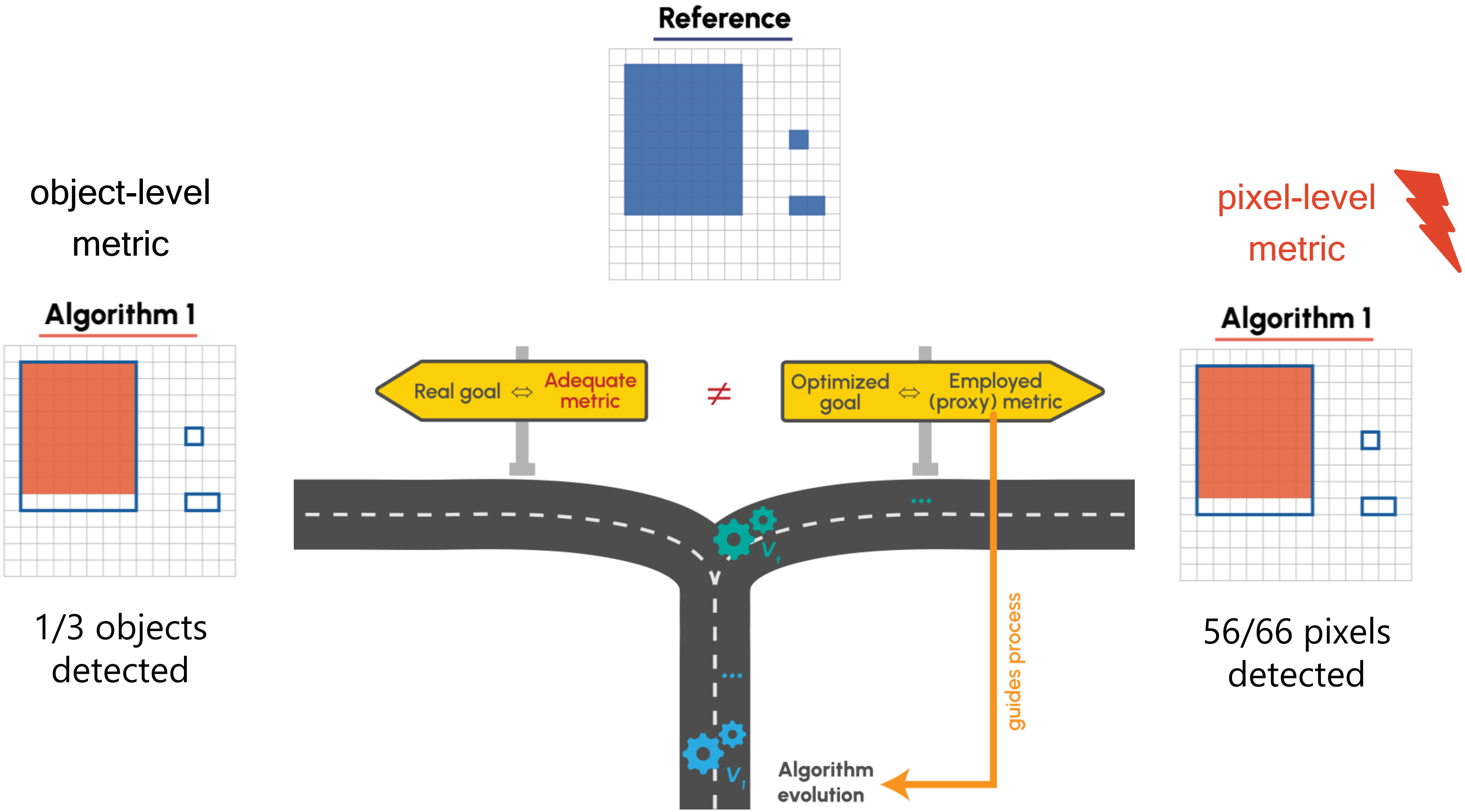


Most tumor pixels are detected...

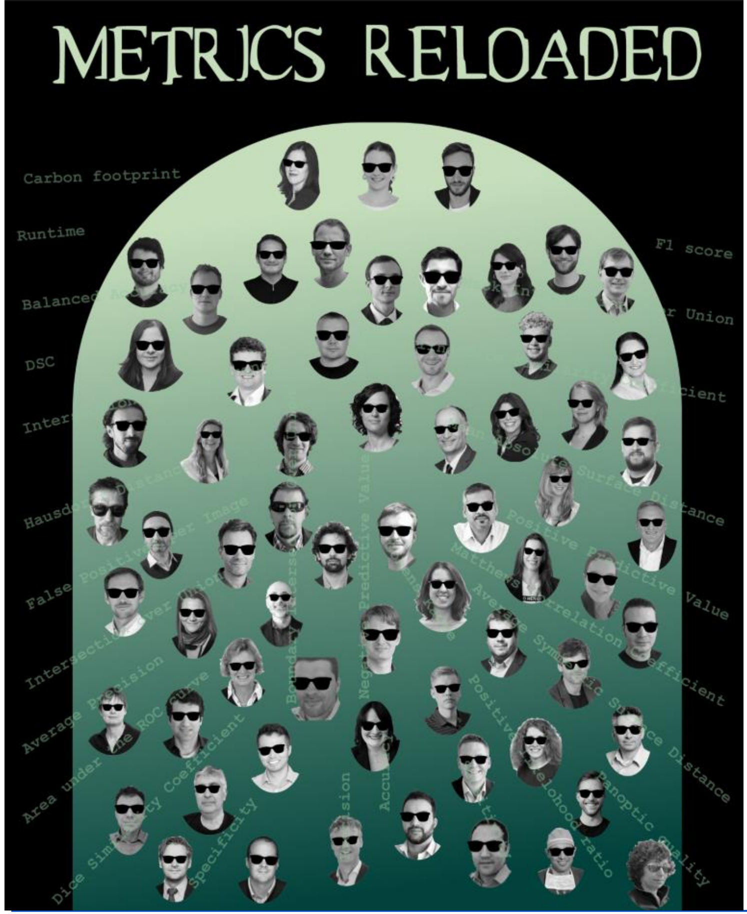
... but the small (new) metastases are missed!

Instance progress not detected in ~1/3 of the cases!

Reinke/Tizabi, ..., Maier-Hein. Understanding metric-related pitfalls in image analysis validation. **Nature Methods 2024**



Reference
object-level metric
Algorithm 1
1/3 objects detected
pixel-level metric
Algorithm 1
56/66 pixels detected
Real goal ⇔ Adequate metric
≠
Optimized goal ⇔ Employed (proxy) metric
guides process
Algorithm evolution



Metrics Reloaded

initiated by Helmholtz Imaging, MONAI and the MICCAI Society

Maier-Hein/Reinke et al. Metrics reloaded: Recommendations for image analysis validation. **Nature Methods 2024**

DALL-E, please create an image of a female president of a surgical society opening the annual congress

Generated with DALL-E

Beyond accuracy:
How to consider ethical issues (e.g. fairness) in medical imaging AI validation?

RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2*], Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5*†]

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

There is growing concern that algorithms may reproduce racial and gender disparities via the people building them or through the data used to train them (*1–3*). Empirical work is increasingly lending support to these concerns. For example, job search ads for highly paid positions are less likely to be presented to women (*4*), searches for distinctively Black-sounding names are more likely to trigger ads for arrest records (*5*), and image searches for professions such as CEO produce fewer images of women (*6*). Facial recognition systems increasingly used in law enforcement perform worse on recognizing faces of women and Black individuals (*7*, *8*), and natural language processing algorithms encode language in gendered ways (*9*).

Empirical investigations of algorithmic bias, though, have been hindered by a key constraint: Algorithms deployed on large scales are typically proprietary, making it difficult for independent researchers to dissect them. Instead, researchers must work "from the outside," often with great ingenuity, and resort to clever workarounds such as audit studies. Such efforts can document disparities, but understanding how and why they arise—much less figuring out what to do about them—is difficult without greater access to the algorithms themselves. Our understanding of a mechanism therefore typically relies on theory or exercises with researcher-created algorithms (*10–13*). Without an algorithm's training data, objective function, and prediction methodology, we can only guess as to the actual mechanisms for the important algorithmic disparities that arise.

In this study, we exploit a rich dataset that provides insight into a live, scaled algorithm deployed nationwide today. It is one of the largest and most typical examples of a class of commercial risk-prediction tools that, by industry estimates, are applied to roughly 200 million people in the United States each year. Large health systems and payers rely on this algorithm to target patients for "high-risk care management" programs. These programs seek to improve the care of patients with complex health needs by providing additional resources, including greater attention from trained providers, to help ensure that care is well coordinated. Most health systems use these programs as the cornerstone of population health management efforts, and they are widely considered effective at improving outcomes and satisfaction while reducing costs (*14–17*). Because the programs are themselves expensive—with costs going toward teams of dedicated nurses, extra primary care appointment slots, and other scarce resources—health systems rely extensively on algorithms to identify patients who will benefit the most (*18*, *19*).

Identifying patients who will derive the greatest benefit from these programs is a challenging causal inference problem that requires estimation of individual treatment effects. To solve this problem, health systems make a key assumption: Those with the greatest care needs will benefit the most from the program. Under this assumption, the targeting problem becomes a pure prediction policy problem (*20*). Developers then build algorithms that rely on past data to build a predictor of future health care needs.

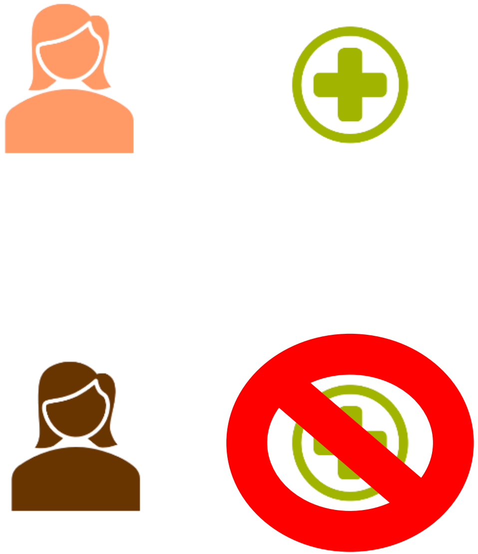
Our dataset describes one such typical algorithm. It contains both the algorithm's predictions as well as the data needed to understand its inner workings: that is, the underlying ingredients used to form the algorithm (data, objective function, etc.) and links to a rich set of outcome data. Because we have the inputs, outputs, and eventual outcomes, our data allow us a rare opportunity to quantify racial disparities in algorithms and isolate the mechanisms by which they arise. It should be emphasized that this algorithm is not unique. Rather, it is emblematic of a generalized approach to risk prediction in the health sector, widely adopted by a range of for- and non-profit medical centers and governmental agencies (*21*).

Our analysis has implications beyond what we learn about this particular algorithm. First, the specific problem solved by this algorithm has analogies in many other sectors: The predicted risk of some future outcome (in our case, health care needs) is widely used to target policy interventions under the assumption that the treatment effect is monotonic in that risk, and the methods used to build the algorithm are standard. Mechanisms of bias uncovered in this study likely operate elsewhere. Second, even beyond our particular finding, we hope that this exercise illustrates the importance, and the large opportunity, of studying algorithmic bias in health care, not just as a model system but also in its own right. By any standard—e.g., number of lives affected, life-and-death consequences of the decision—health is one of the most important and widespread social sectors in which algorithms are already used at scale today, unbeknownst to many.

Data and analytic strategy

Working with a large academic hospital, we identified all primary care patients enrolled in risk-based contracts from 2013 to 2015. Our primary interest was in studying differences between White and Black patients. We formed race categories by using hospital records, which are based on patient self-reporting. Any patient who identified as Black was considered to be Black for the purpose of this analysis. Of the remaining patients, those who self-identified as races other than White (e.g., Hispanic) were so considered (data on these patients are presented in table S1 and fig. S1 in the supplementary materials). We considered all remaining patients to be White. This approach allowed us to study one particular racial difference of social and historical interest between patients who self-identified as Black and patients who self-identified as White without another race or ethnicity; it has the disadvantage of not allowing for the study of intersectional racial

[1]School of Public Health, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, USA. [3]Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA. [4]Mongan Institute Health Policy Center, Massachusetts General Hospital, Boston, MA, USA. [5]Booth School of Business, University of Chicago, Chicago, IL, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: sendhil.mullainathan@chicagobooth.edu



RESEARCH

RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2*], Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5*†]

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

There is growing concern that algorithms may reproduce racial and gender disparities via the people building them or through the data used to train them (*1*–*3*). Empirical work is increasingly lending support to these concerns. For example, job search ads for highly paid positions are less likely to be presented to women (*4*), searches for distinctively Black-sounding names are more likely to trigger ads for arrest records (*5*), and image searches for professions such as CEO produce fewer images of women (*6*). Facial recognition systems increasingly used in law enforcement perform worse on recognizing faces of women and Black individuals (*7*, *8*), and natural language processing algorithms encode language in gendered ways (*9*).

Empirical investigations of algorithmic bias, though, have been hindered by a key constraint: Algorithms deployed on large scales are typically proprietary, making it difficult for independent researchers to dissect them. Instead, researchers must work "from the outside," often with great ingenuity, and resort to clever workarounds such as audit studies. Such efforts can document disparities, but understanding how and why they arise—much less figuring out what to do about them—is difficult without greater access to the algorithms themselves. Our understanding of a mechanism therefore typically relies on theory or exercises with researcher-created algorithms (*10*–*13*). Without an algorithm's training data, objective function, and prediction methodology, we can only guess as to the actual mechanisms for the important algorithmic disparities that arise.

In this study, we exploit a rich dataset that provides insight into a live, scaled algorithm deployed nationwide today. It is one of the largest and most typical examples of a class of commercial risk-prediction tools that, by industry estimates, are applied to roughly 200 million people in the United States each year. Large health systems and payers rely on this algorithm to target patients for "high-risk care management" programs. These programs seek to improve the care of patients with complex health needs by providing additional resources, including greater attention from trained providers, to help ensure that care is well coordinated. Most health systems use these programs as the cornerstone of population health management efforts, and they are widely considered effective at improving outcomes and satisfaction while reducing costs (*14*–*17*). Because the programs are themselves expensive—with costs going toward teams of dedicated nurses, extra primary care appointment slots, and other scarce resources—health systems rely extensively on algorithms to identify patients who will benefit the most (*18*, *19*).

Identifying patients who will derive the greatest benefit from these programs is a challenging causal inference problem that requires estimation of individual treatment effects. To solve this problem, health systems make a key assumption: Those with the greatest care needs will benefit the most from the program. Under this assumption, the targeting problem becomes a pure prediction policy problem (*20*). Developers then build algorithms that rely on past data to build a predictor of future health care needs.

Our dataset describes one such typical algorithm. It contains both the algorithm's predictions as well as the data needed to understand its inner workings: that is, the underlying ingredients used to form the algorithm (data, objective function, etc.) and links to a rich set of outcome data. Because we have the inputs, outputs, and eventual outcomes, our data allow us a rare opportunity to quantify racial disparities in algorithms and isolate the mechanisms by which they arise. It should be emphasized that this algorithm is not unique. Rather, it is emblematic of a generalized approach to risk prediction in the health sector, widely adopted by a range of for- and non-profit medical centers and governmental agencies (*21*).

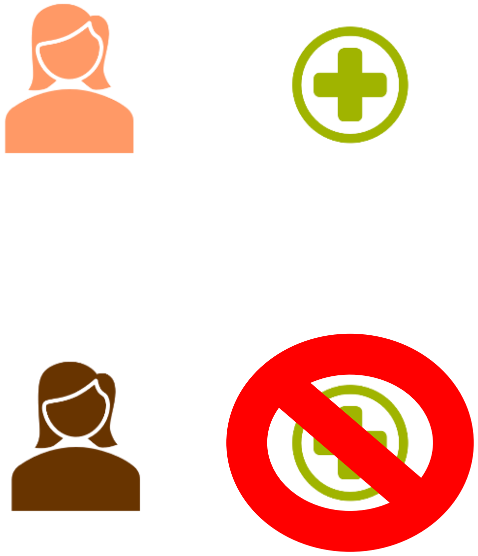
Our analysis has implications beyond what we learn about this particular algorithm. First, the specific problem solved by this algorithm has analogies in many other sectors: The predicted risk of some future outcome (in our case, health care needs) is widely used to target policy interventions under the assumption that the treatment effect is monotonic in that risk, and the methods used to build the algorithm are standard. Mechanisms of bias uncovered in this study likely operate elsewhere. Second, even beyond our particular finding, we hope that this exercise illustrates the importance, and the large opportunity, of studying algorithmic bias in health care, not just as a model system but also in its own right. By any standard—e.g., number of lives affected, life-and-death consequences of the decision—health is one of the most important and widespread social sectors in which algorithms are already used at scale today, unbeknownst to many.

Data and analytic strategy

Working with a large academic hospital, we identified all primary care patients enrolled in risk-based contracts from 2013 to 2015. Our primary interest was in studying differences between White and Black patients. We formed race categories by using hospital records, which are based on patient self-reporting. Any patient who identified as Black was considered to be Black for the purpose of this analysis. Of the remaining patients, those who self-identified as races other than White (e.g., Hispanic) were so considered (data on these patients are presented in table S1 and fig. S1 in the supplementary materials). We considered all remaining patients to be White. This approach allowed us to study one particular racial difference of social and historical interest between patients who self-identified as Black and patients who self-identified as White without another race or ethnicity; it has the disadvantage of not allowing for the study of intersectional racial

[1]School of Public Health, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, USA. [3]Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA. [4]Mongan Institute Health Policy Center, Massachusetts General Hospital, Boston, MA, USA. [5]Booth School of Business, University of Chicago, Chicago, IL, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: sendhil.mullainathan@chicagobooth.edu

Obermeyer *et al.*, *Science* 366, 447–453 (2019)   25 October 2019



Symptomlast

Anamese

~~Behandlung~~

PLZ

Versicherungsstatus

Keine Behandlung

Hautfarbe

Symptomlast

Anamese

~~Behandlung~~

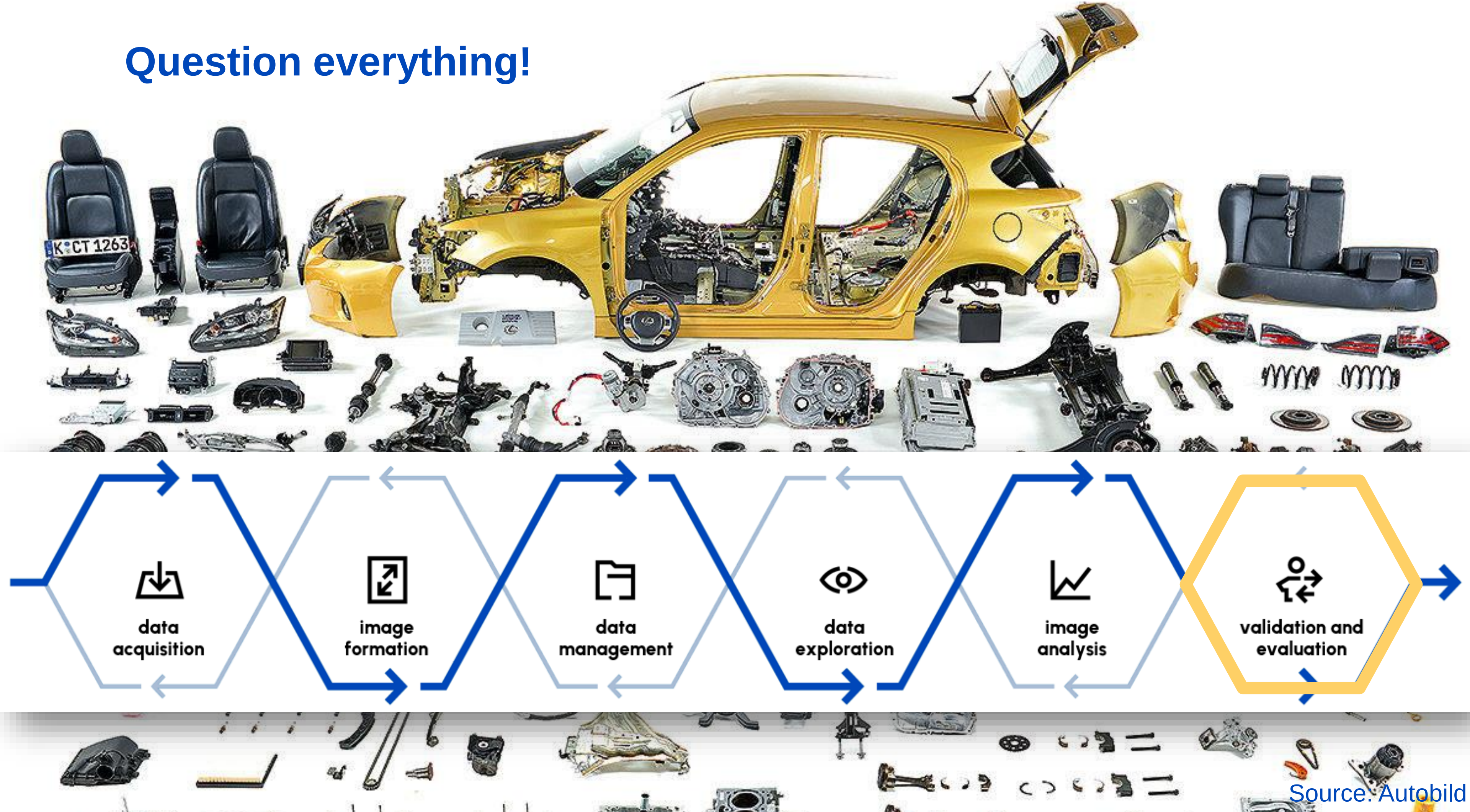
PLZ

Versicherungsstatus

Keine Behandlung

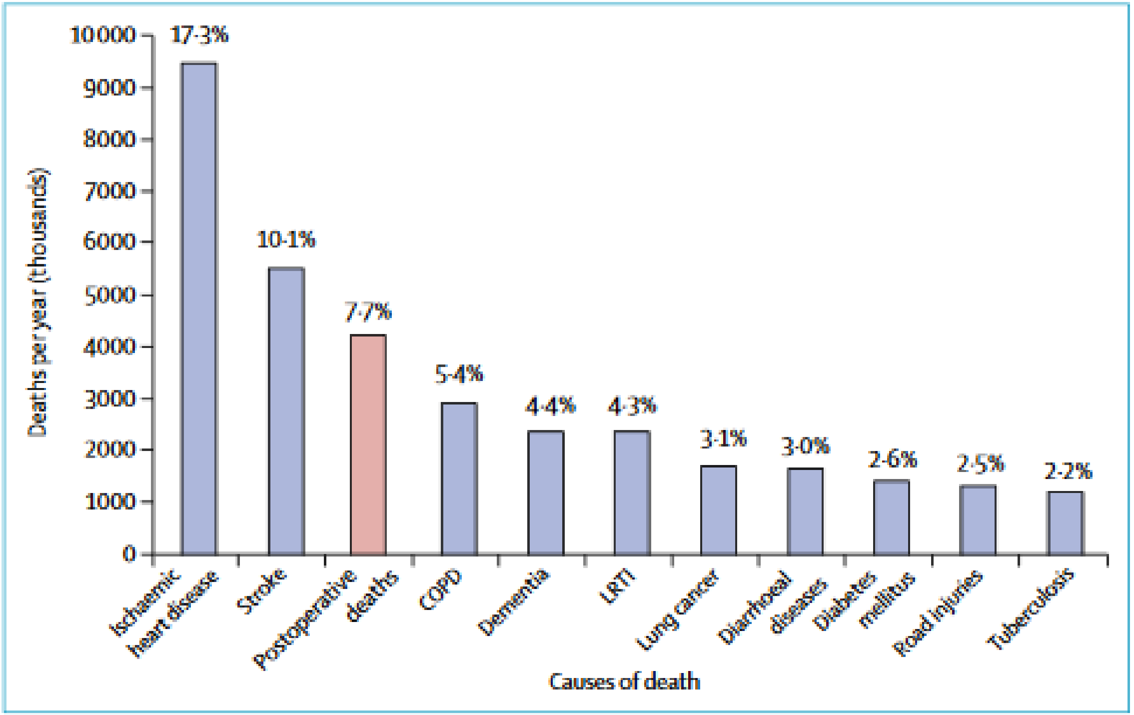
Hautfarbe

Question everything!



Source: Autobild

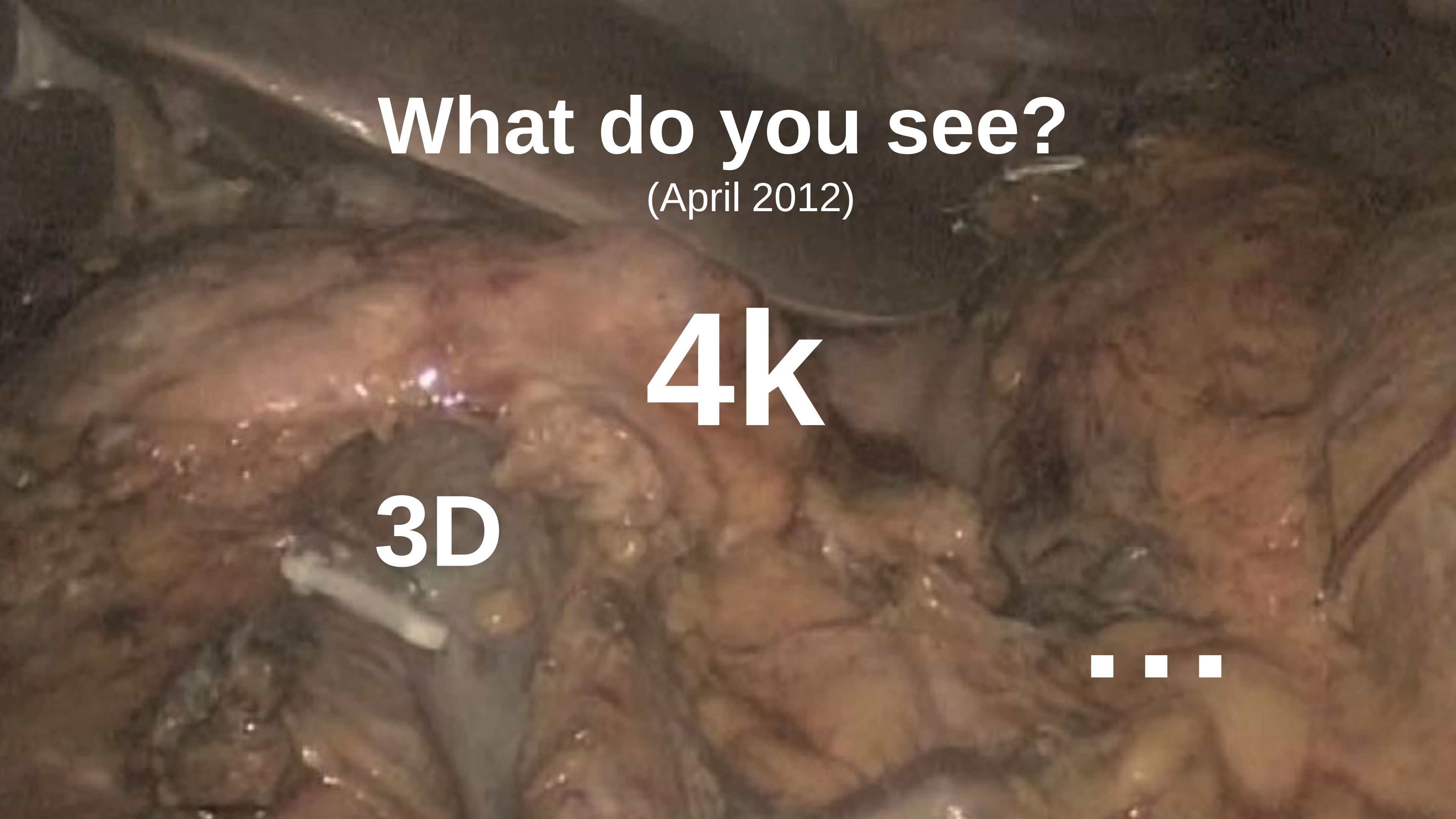
Surgery: A high stakes domain



Nepogodiev et al. Global burden of postoperative death, **Lancet 2019**

What do you see?

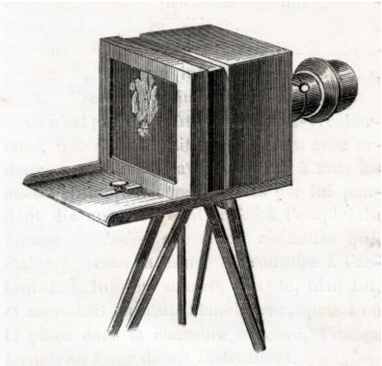
(April 2012)



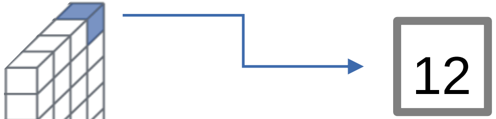
"If I had asked people what they wanted, they would have said faster horses."

— Henry Ford

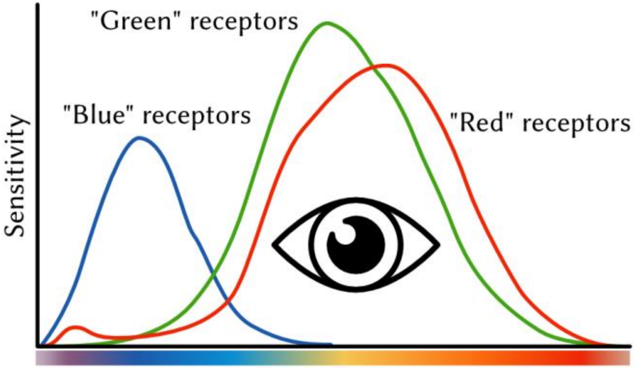
Recap: From *camera obscura* to RGB Cameras

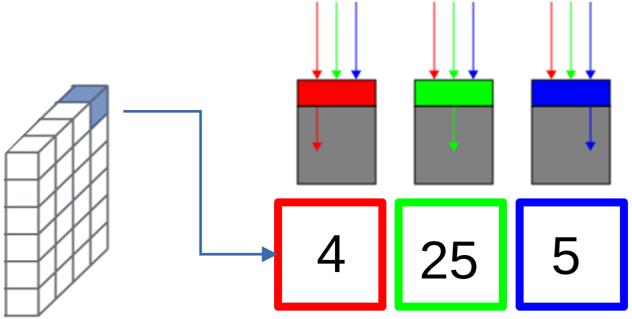


Source: SciencePhotoLibrary



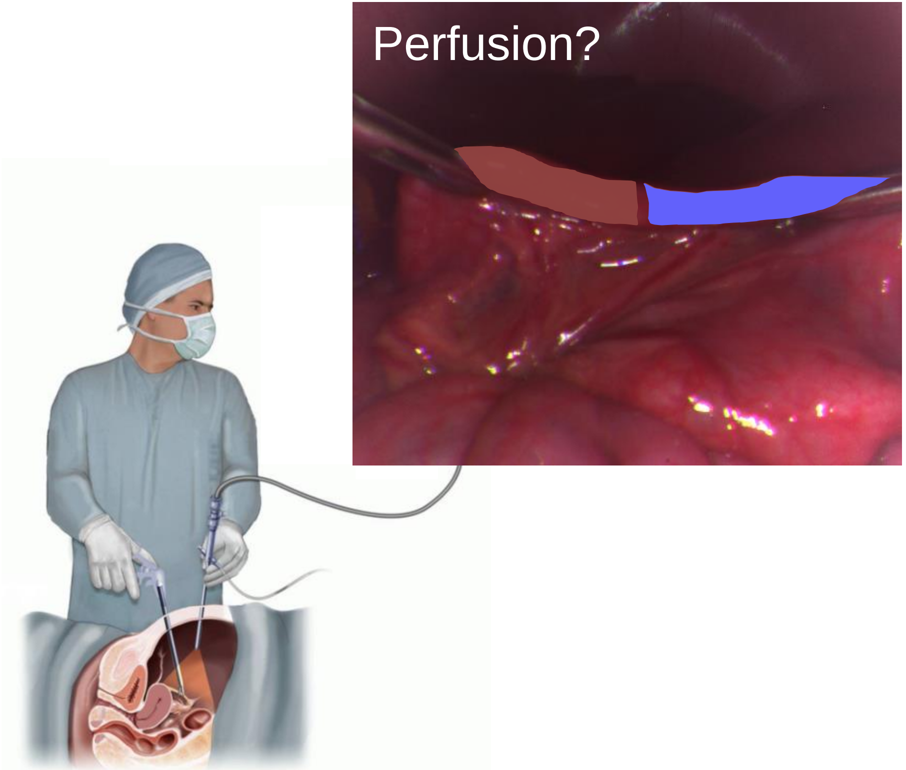
Single value / pixel
= Intensity of light



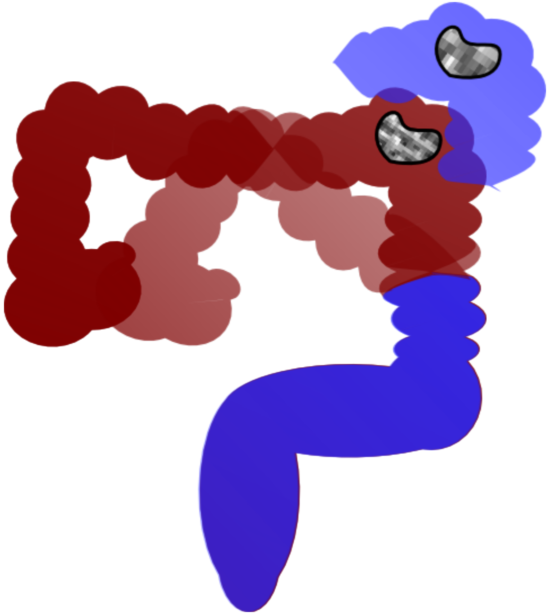


= Intensity for
Red/Green/Blue

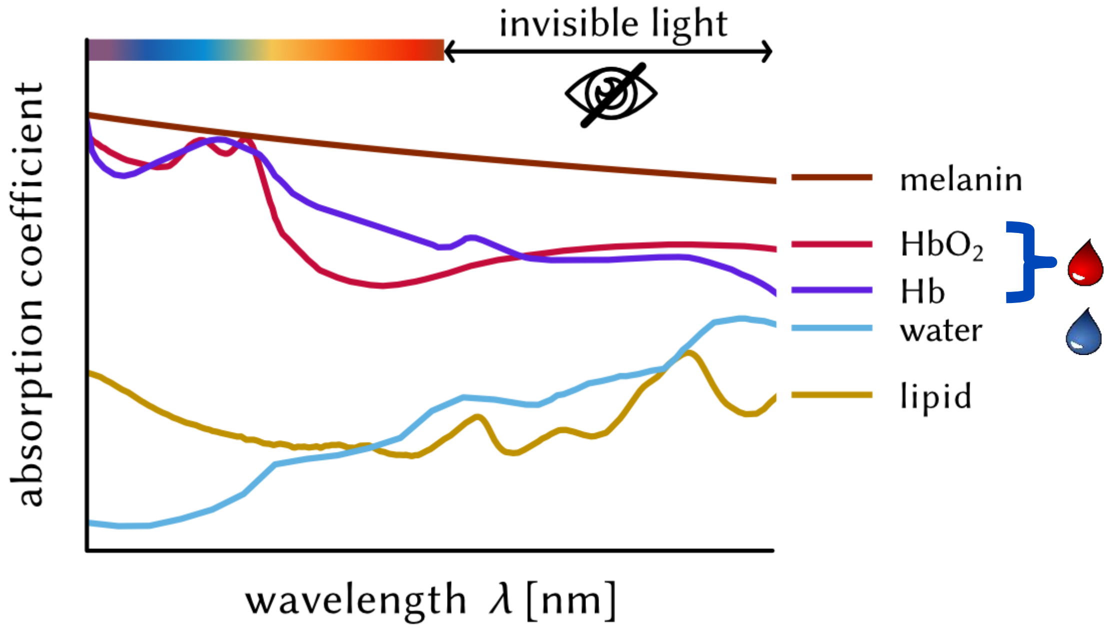
Limitations of human perception



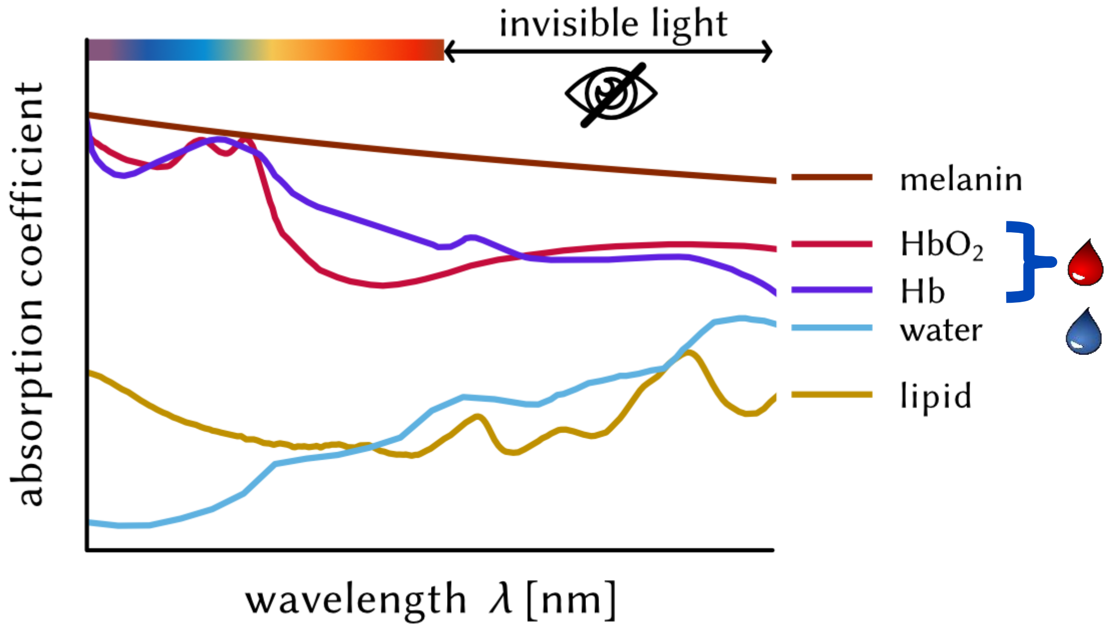
Source: thegastrocare.in

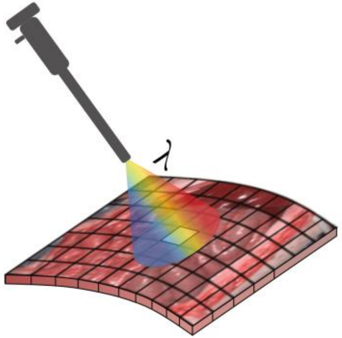


Beyond human perception: Spectral imaging



Beyond human perception: Spectral imaging





Funding granted

European Research Council

Established by the European Commission

Source: Reddit
@billionbackrecords

Priorities:

Is it ethical to work on *high tech* when the global needs are much more basic (e.g. sterilization)?


KETCHUP
Gera Gewürze
Thüringer Tomaten
KETCHUP
℮ 450 ml
Gera Gewürze
Tomatenmark aus EU und nicht-EU

npj digital medicine
www.nature.com/npjdigitalmed

REVIEW ARTICLE OPEN

Check for updates

Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review

Tadeusz Ciecierski-Holmes[1,2], Ritvij Singh[3], Miriam Axt[1], Stephan Brenner[1] and Sandra Barteit[1]

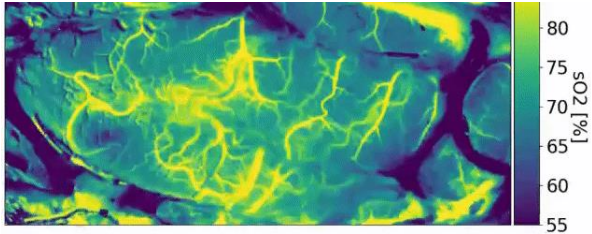
In low- and middle-income countries (LMICs), AI has been promoted as a potential means of strengthening healthcare systems by a growing number of publications. We aimed to evaluate the scope and nature of AI technologies in the specific context of LMICs. In this systematic scoping review, we used a broad variety of AI and healthcare search terms. Our literature search included records published between 1st January 2009 and 30th September 2021 from the Scopus, EMBASE, MEDLINE, Global Health and APA PsycInfo databases, and grey literature from a Google Scholar search. We included studies that reported a quantitative and/or qualitative evaluation of a real-world application of AI in an LMIC health context. A total of 10 references evaluating the application of AI in an LMIC were included. Applications varied widely, including: clinical decision support systems, treatment planning and triage assistants and health chatbots. Only half of the papers reported which algorithms and datasets were used in order to train the AI. A number of challenges of using AI tools were reported, including issues with reliability, mixed impacts on workflows, poor user friendliness and lack of adeptness with local contexts. Many barriers exists that prevent the successful development and adoption of well-performing, context-specific AI tools, such as limited data availability, trust and evidence of cost-effectiveness in LMICs. Additional evaluations of the use of AI in healthcare in LMICs are needed in order to identify their effectiveness and reliability in real-world settings and to generate understanding for best practices for future implementations.

npj Digital Medicine (2022)5:162 ; https://doi.org/10.1038/s41746-022-00700-y

vs.

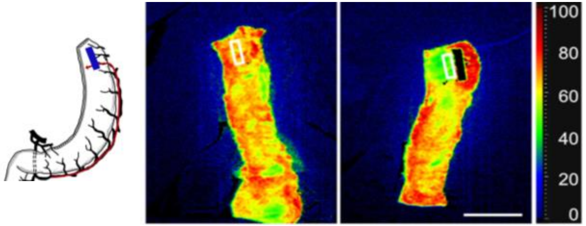
Back to the science: A new window into the body

Monitoring of hemodynamics for stroke treatment

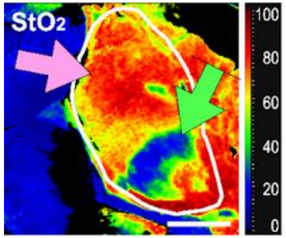


Automatic tissue differentiation

Optimization of surgical technique



Organ transplantation



...

Priorities:
Should we select use cases top-down rather than bottom-up to incorporate global needs

Thanks to our clinical collaborators…

Prof. Dr. Christoph Michalski
Univ. Hospital Heidelberg

Prof. Dr. Dogu Teber
Städtisches Klinikum Karlsruhe

Prof. Dr. M. Weigand
Univ. Hospital Heidelberg

Prof. Dr. Karl Kowalewki
Univ. Hospital Mannheim

Prof. Dr. Beat Müller
Universitätsspital Basel

PD Dr. Edgar Santos
Univ. Hospital Oldenburg

PD Dr. Felix Nickel
Univ. Hospital Hamburg-E.

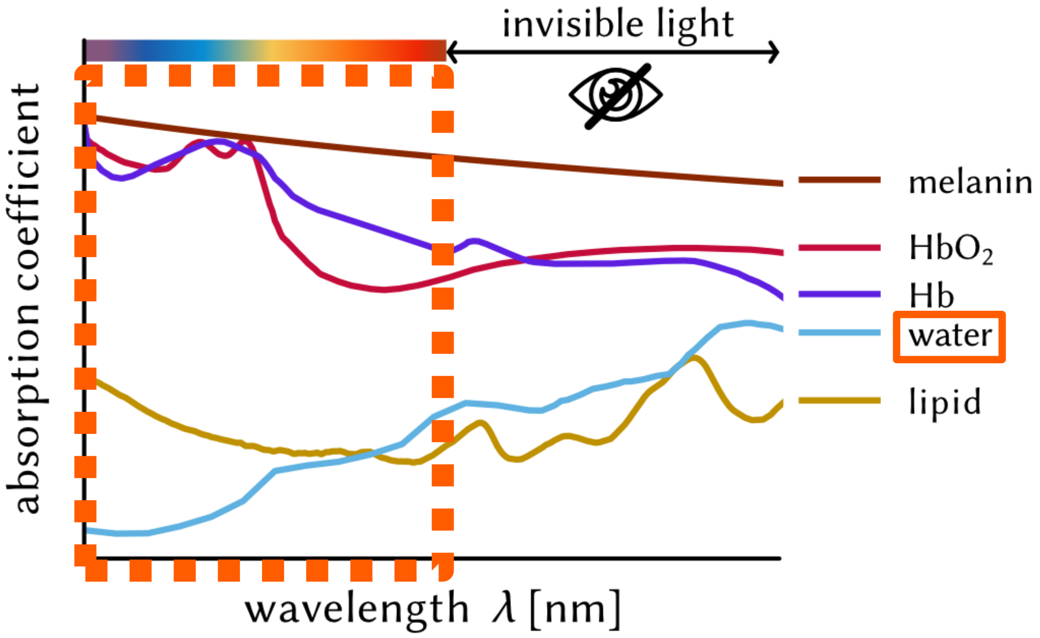
Dr. Alexander Studier-Fischer
Univ. Hospital Heidelberg

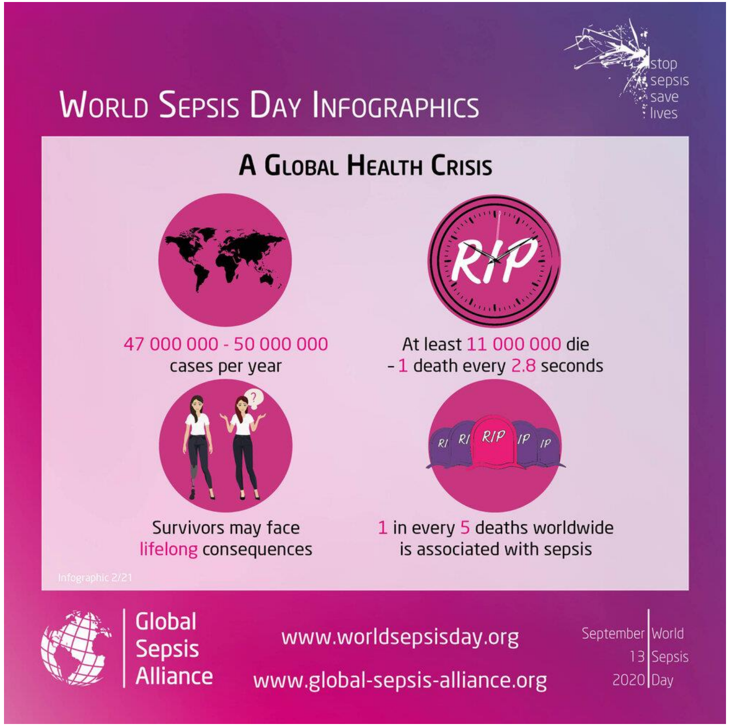
Dr. Maximilian Dietrich
Univ. Hospital Heidelberg

Dr. H. Götz Kenngott
Univ. Hospital Heidelberg

… and their teams

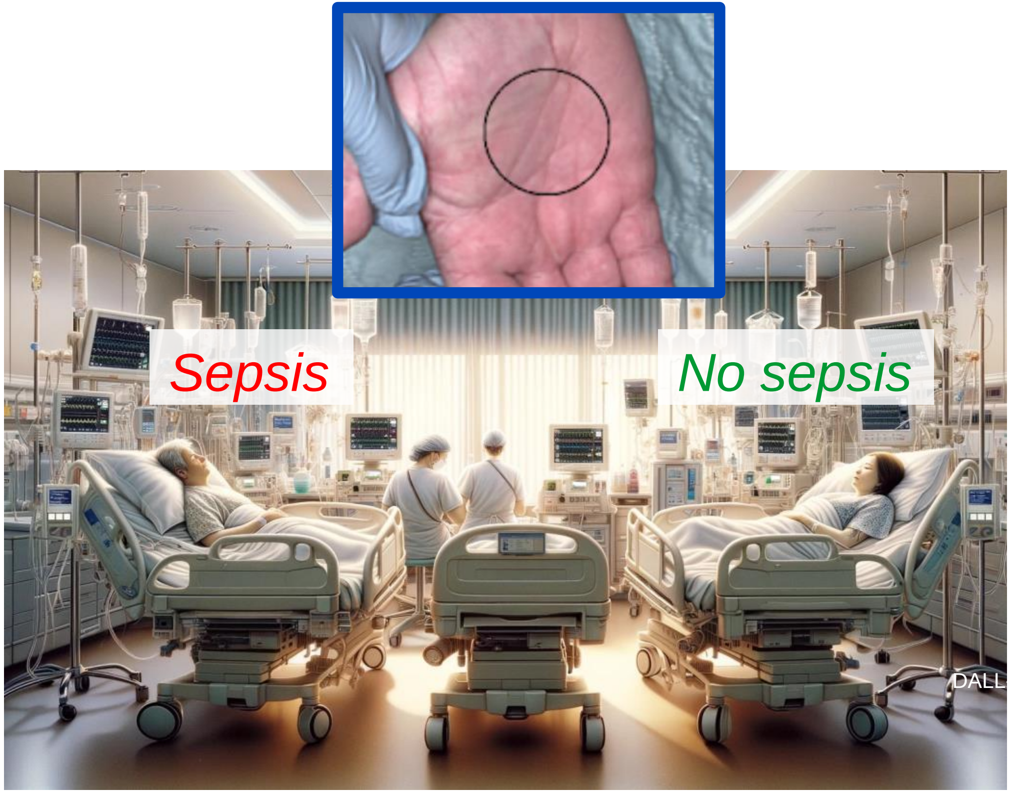
A *global AI* use case

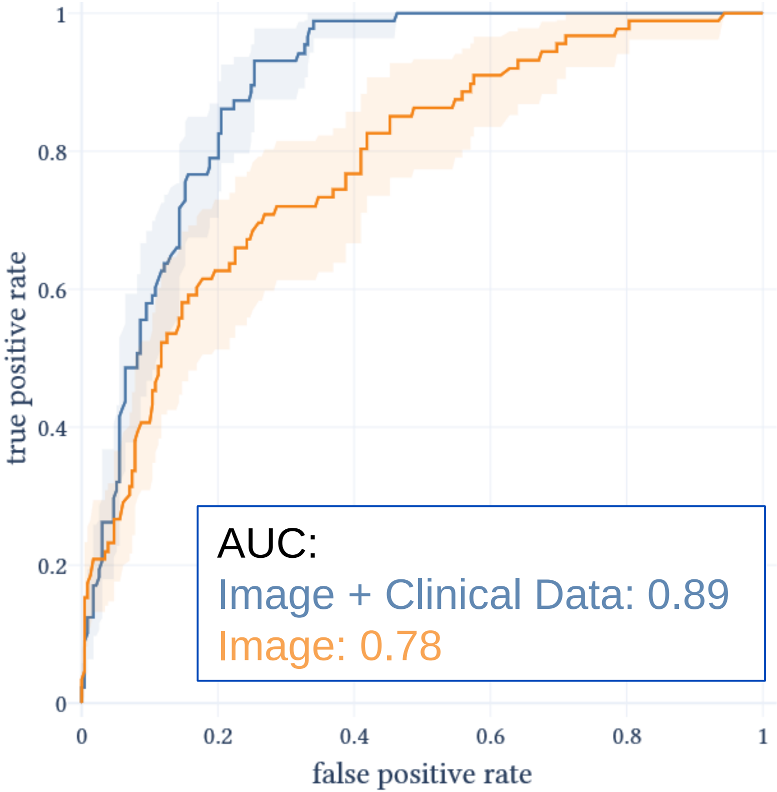




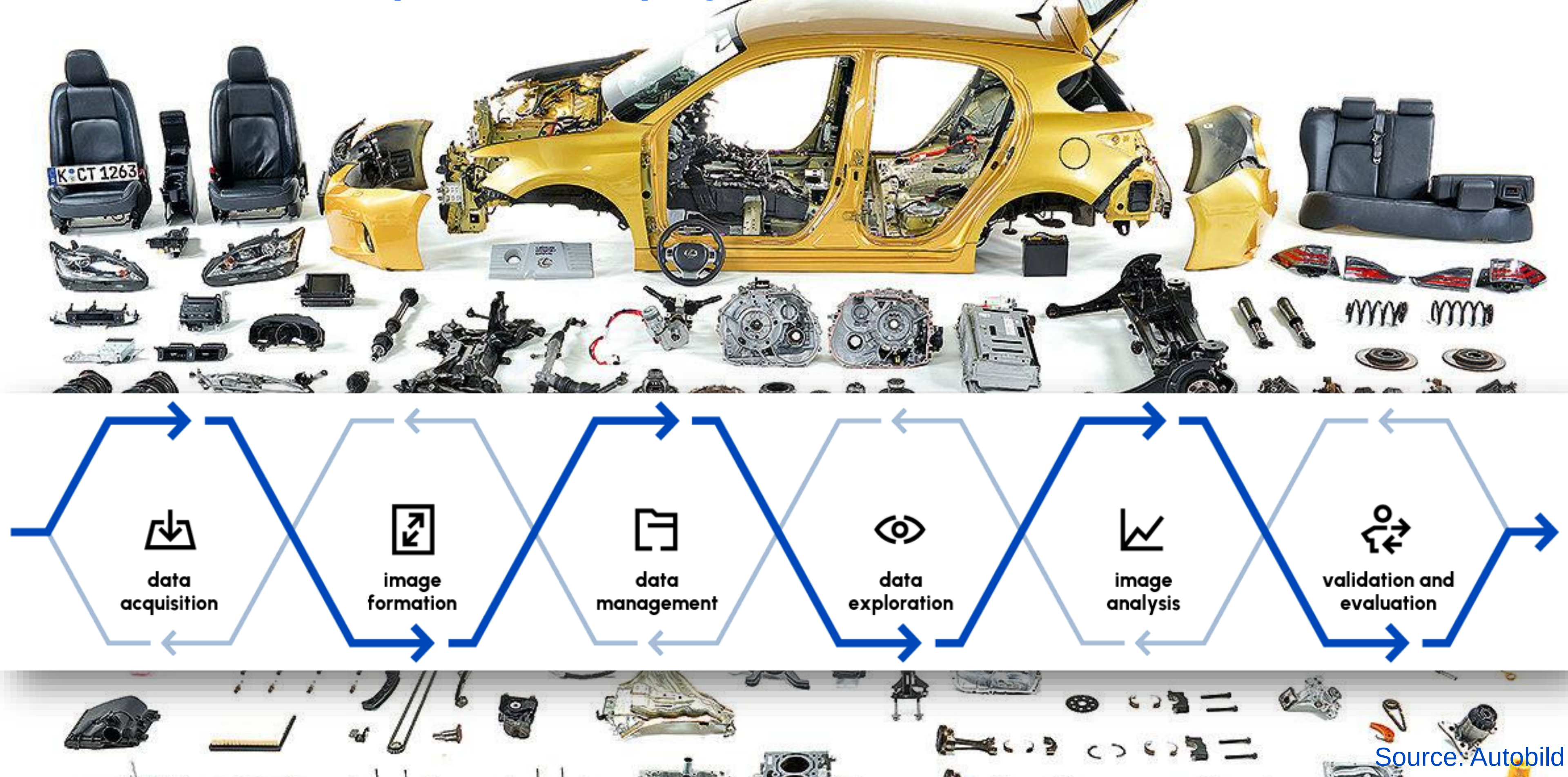
Source: global-sepsis-alliance.org

Unpublished: A new biomarker for sepsis

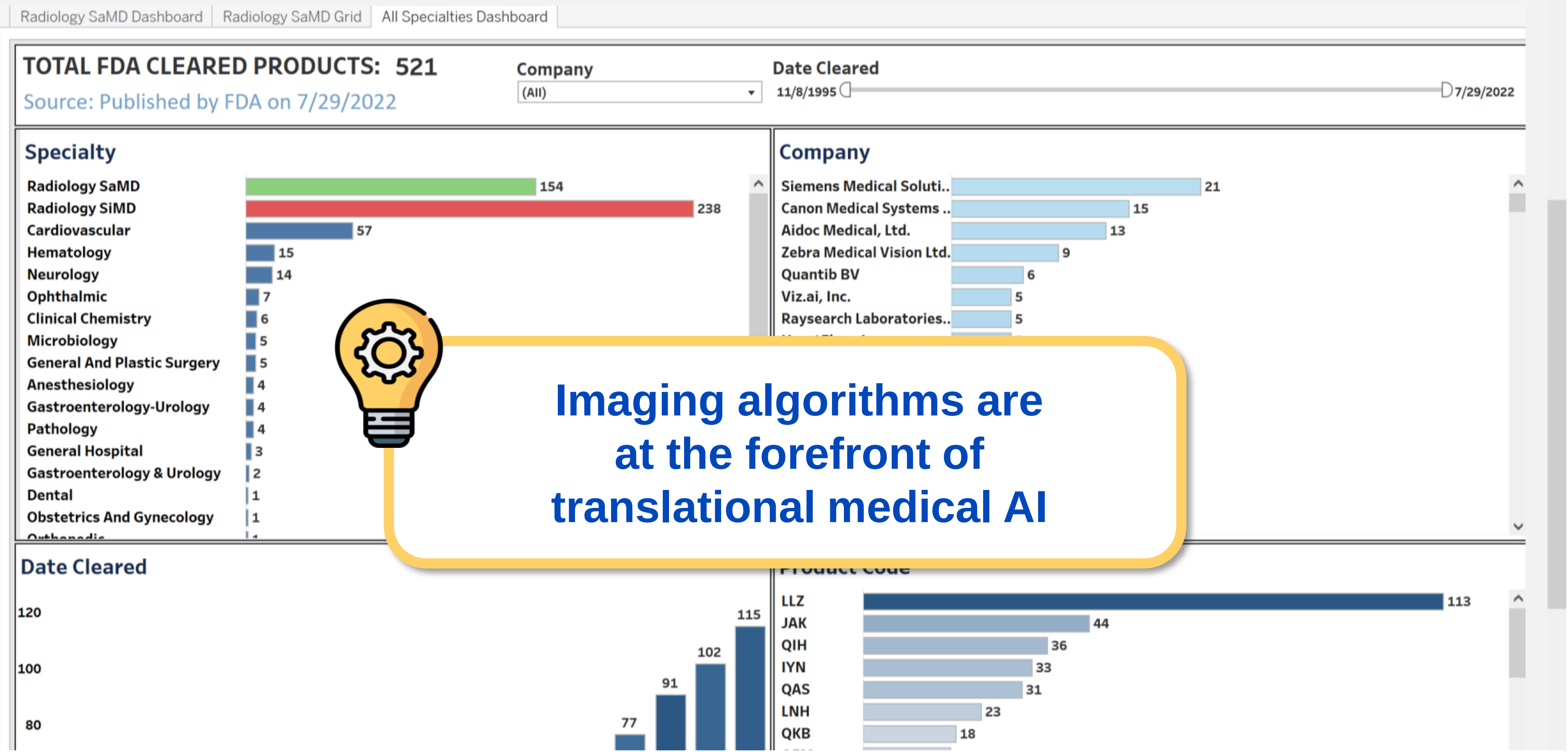




From development to deployment



Source: Autobild



Imaging algorithms are at the forefront of translational medical AI

https://aicentral.acrdsi.org (accessed November 13th, 2023)

Only few products are actually used (frequently) so far

Characterizing the Clinical Adoption of Medical AI Devices through U.S. Insurance Claims

Kevin Wu, M.S.,[1] Eric Wu, M.S.,[2] Brandon Theodorou,[3] Weixin Liang, M.S.,[4] Christina Mack, Ph.D.,[5] Lucas Glass, Ph.D.,[5] Jimeng Sun, Ph.D.,[3,6] and James Zou, Ph.D.[1,2,4]

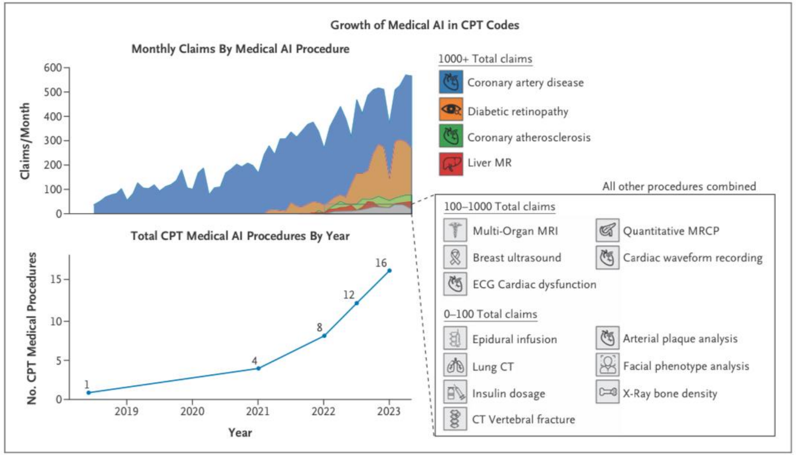


Table 1. Summary of AI CPT Codes.*

Total Claims	Condition or Medical AI Procedure	CPT Code(s)	Example Product Name	Effective Date
67,306	Coronary artery disease	0501T–0504T	HeartFlow Analysis[48]	June 1, 2018
15,097	Diabetic retinopathy	92229	LumineticsCore[49]	January 1, 2021
4,459	Coronary atherosclerosis	0623T–0626T	Cleerly[50]	January 1, 2021
2,428	Liver MR	0648T–0649T	Perspectum LiverMultiScan[51]	January 1, 2021
591	Multiorgan MRI	0697T–0698T	Perspectum CoverScan[52]	January 1, 2022
552	Breast ultrasound	0689T–0690T	Koios DS[53]	January 1, 2022
435	ECG cardiac dysfunction	0764T–0765T	Anumana[50]	January 1, 2023
331	Cardiac acoustic waveform recording	0716T	CADScor[50]	July 1, 2022
237	Quantitative MR cholangiopancreatography	0723T–0724T	Perspectum MRCP+[54]	July 1, 2022
67	Epidural infusion	0777T	CompuFlo[55]	January 1, 2023
4	Quantitative CT tissue characterization	0721T–0722T	Optellum Virtual Nodule Clinic[56]	July 1, 2022
1	Autonomous insulin dosage	0740T–0741T	d-Nav[57]	January 1, 2023
1	CT vertebral fracture assessment	0691T	HealthVCF[50]	January 1, 2022
1	Noninvasive arterial plaque analysis	0710T–0713T	ElucidVivo[50]	January 1, 2022
0	Facial phenotype analysis	0731T	Face2Gene[50]	July 1, 2022
0	X-ray bone density	0749T	OsteoApp[50]	January 1, 2023

Responsibility:

Who takes responsibility in case of failure?

Menschliche Letztverantwortung

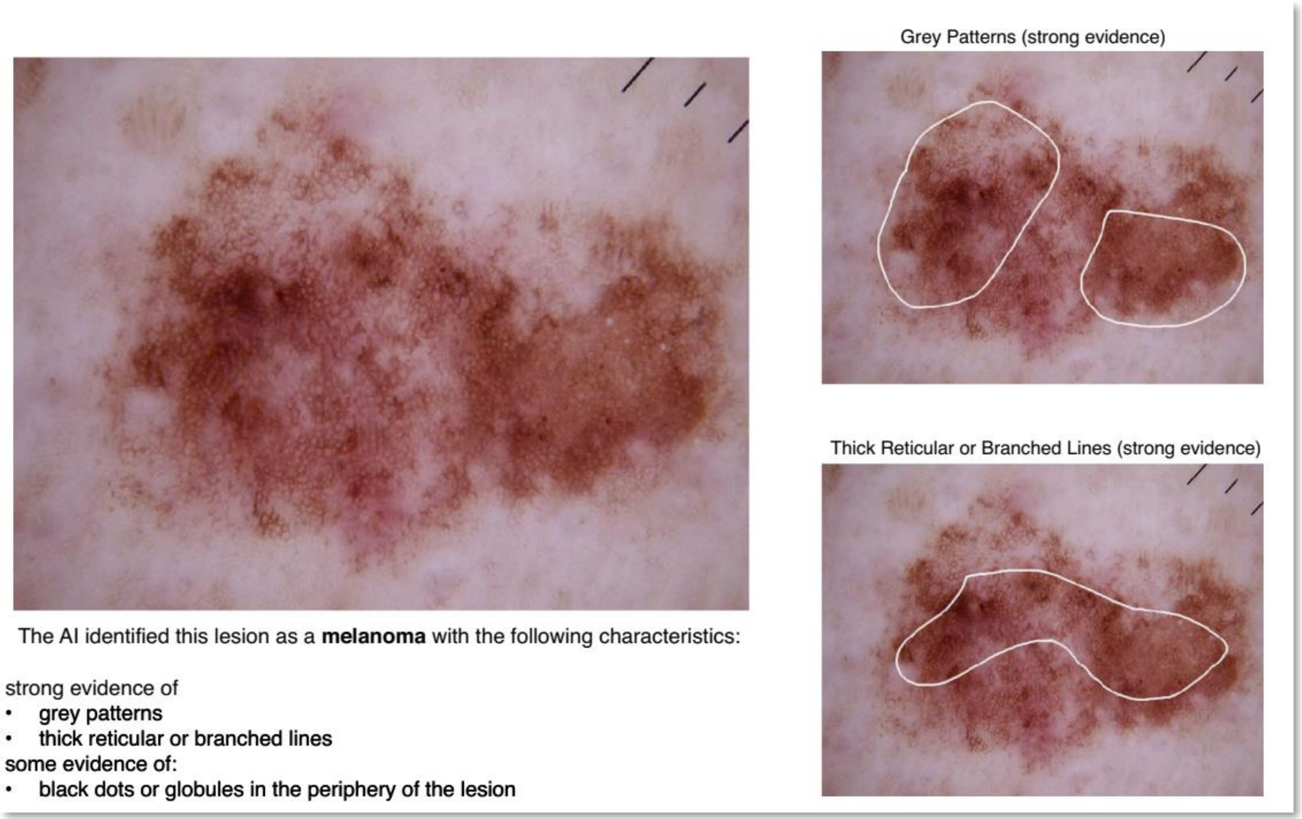
Human oversight

Human-in-the-loop

Bias Behind the Wheel: Fairness Testing of Autonomous Driving Systems

XINYUE LI, Peking University, China
ZHENPENG CHEN*, Nanyang Technological University, Singapore
JIE M. ZHANG, King's College London, United Kingdom
FEDERICA SARRO, University College London, United Kingdom
YING ZHANG, Peking University, China
XUANZHE LIU, Peking University, China


GTI
GF JN53H

Quelle: Chanda et al. 2023

Menschliche Letztverantwortung

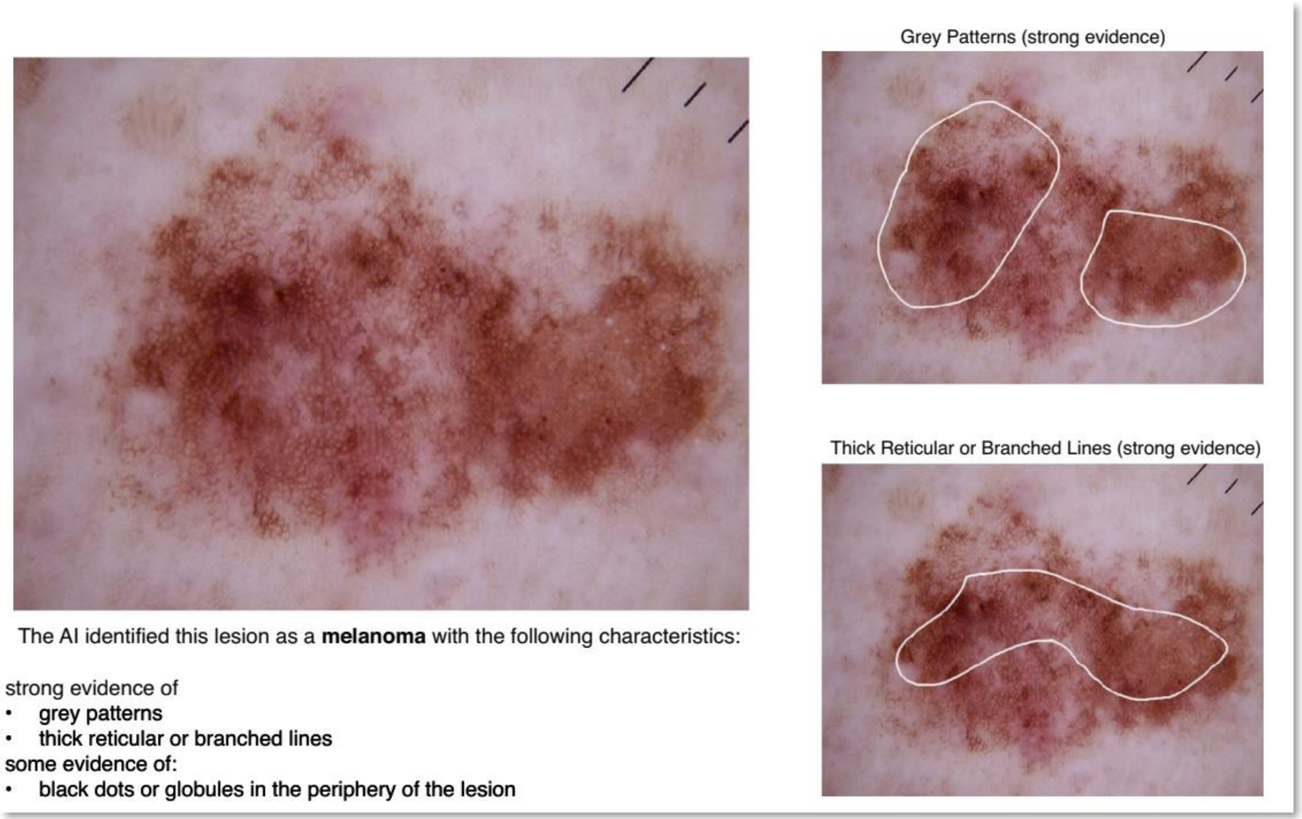
Human oversight

Human-in-the-loop

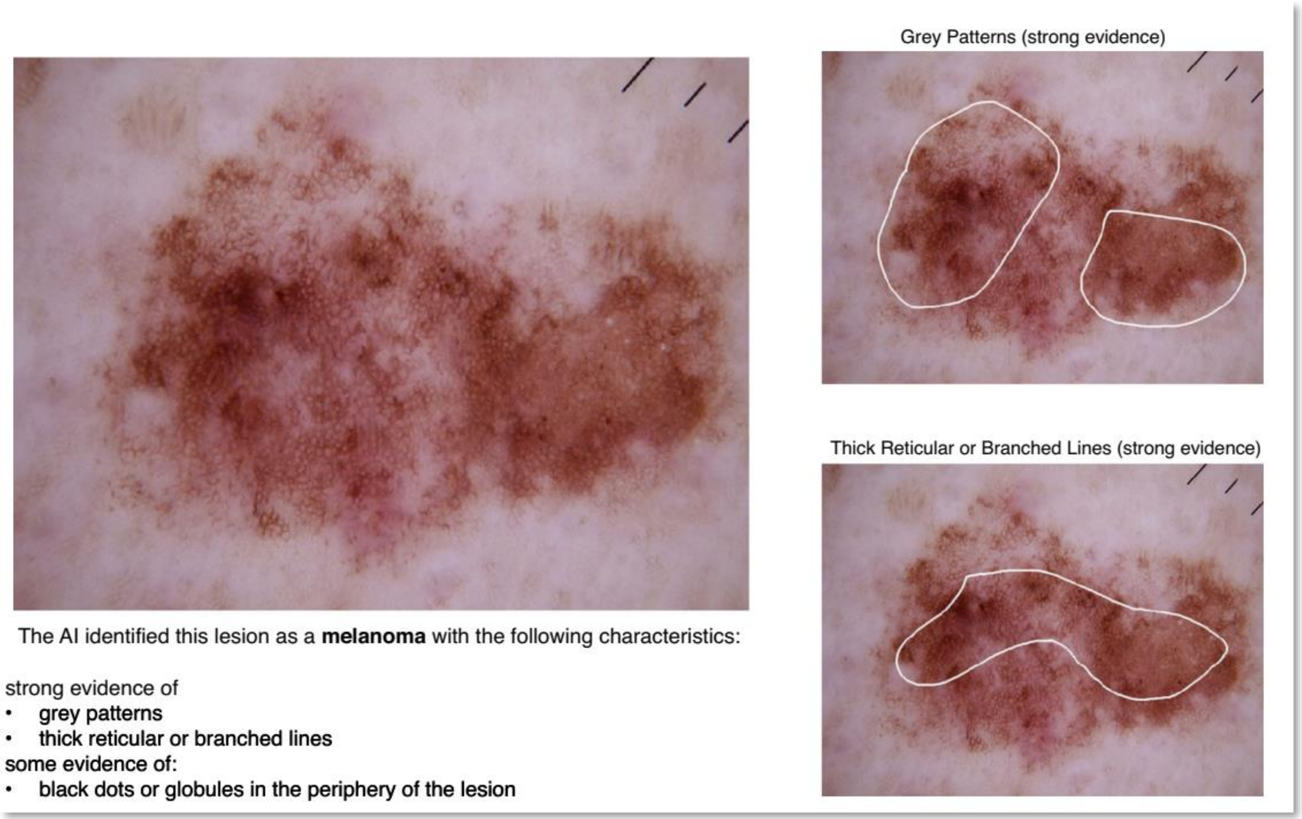
Ungewissheit

?


100m
60



Quelle: Chanda et al. 2023



Quelle: Chanda et al. 2023

International Journal of Computer Assisted Radiology and Surgery (2021) 16:2107–2117
https://doi.org/10.1007/s11548-021-02523-w

ORIGINAL ARTICLE

Check for updates

Point detection through multi-instance deep heatmap regression for sutures in endoscopy

Lalith Sharan[1] · Gabriele Romano[2] · Julian Brand[1] · Halvar Kelm[1] · Matthias Karck[2] · Raffaele De Simone[2] · Sandy Engelhardt[1]

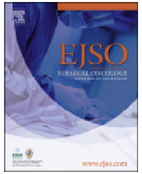
Received: 20 April 2021 / Accepted: 18 October 2021 / Published

ELSEVIER

Contents lists available at ScienceDirect

European Journal of Surgical Oncology

journal homepage: www.ejso.com



Optimization of anastomotic technique and gastric conduit perfusion with hyperspectral imaging and machine learning in an experimental model for minimally invasive esophagectomy

F. Nickel [a, b, 1], A. Studier-Fischer [a, c, 1], B. Özdemir [a], J. Odenthal [a], L.R. Müller [b, d, e], S. Knoedler [a], K.F. Kowalewski [f], I. Camplisson [g], M.M. Allers [a], M. Dietrich [h], K. Schmidt [i], G.A. Salg [a], H.G. Kenngott [a], A.T. Billeter [a], I. Gockel [j], C. Sagiv [k], O.E. Hadar [k], J. Gildenblat [k], L. Ayala [b, d, l], S. Seidlitz [b, d], L. Maier-Hein [b, d, e, l], B.P. Müller-Stich [a, b, *]

Many (more) open questions

Climate:
AGI benefits versus carbon footprint implications

Values:
Implicit encoding of subjective preferences

Much more:
- Job replacements
- Copyright issues
- Responsibility
- …

Take Home Messages:

1. AI is continuing to take every hurdle; AGI is the future

2. Numerous ethical questions remain

3. There is no one-size-fits-all in AI ethics

Intelligent Medical Systems, DKFZ

Institute for Medical and Data Ethcis, Heidelberg University

X @lena_maierhein
@DKFZ_IMSY_lab

HELMHOLTZ IMAGING
NCT
INTELLIGENT SYSTEMS IN SURGICAL ONCOLOGY
erc European Research Council

AI + ethics everywhere

The Washington Post
Democracy Dies in Darkness

Reddit slams ‘unethical experiment’ that deployed secret AI bots in forum

The platform’s chief legal officer called out the University of Zurich team that deployed bots on r/changemyview to study how AI can influence opinions.

April 30, 2025