Will AI make radiology more or less expensive?

January 20, 2022
Dutch Israeli Mini-Symposium on AI and Radiology

Bram van Ginneken

Diagnostic Image Analysis Group / Radboud AI for Health, Radboud University Medical Center; Fraunhofer MEVIS, Bremen; Thirona, Nijmegen

Radboudumc

Disclosures

- **Initial developer CAD4TB (Delft Imaging): royalties & funding**
- **Co-founder and CSO Thirona: stock, royalties & funding**
- **Co-Developer Veolity (MeVis Medical Solutions) & DynaCAD Lung (InVivo): royalties & funding**
- **I lead the Diagnostic Image Analysis Group at Radboud University Medical Center.**
  We receive royalties & funding from: Canon, Siemens Healthineers, Philips, ScreenPoint, Amazon Web Services, Elekta, Sectra, Novartis

AI in radiology

Radboudumc

1963 The dream of Gwilym Lodwick

VOL. 81 NO. 2

Radiology

AUGUST 1963

a monthly journal devoted to clinical radiology and allied sciences

PUBLISHED BY THE RADIOLOGICAL SOCIETY OF NORTH AMERICA, INC.

The Coding of Roentgen Images for Computer Analysis as Applied to Lung Cancer[1]

GWILYM S. LODWICK, M.D., THEODORE E. KEATS, M.D., and JOHN P. DORST, M.D.

THIS PAPER WILL DESCRIBE a concept of converting the visual images on roentgenograms into numerical sequences that can be manipulated and evaluated by the digital computer and will report the results of employing this system to determine the significance of certain radiographic findings in lung cancer. The cause, against a background of air density, the intimate details of the relationship between tumor and host may be faithfully reproduced roentgenographically. Parenthetically, it may be stated that similar density ranges exist in the relationships between bone and soft tissue and that an equally effective descriptive system

2012 AlexNet

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

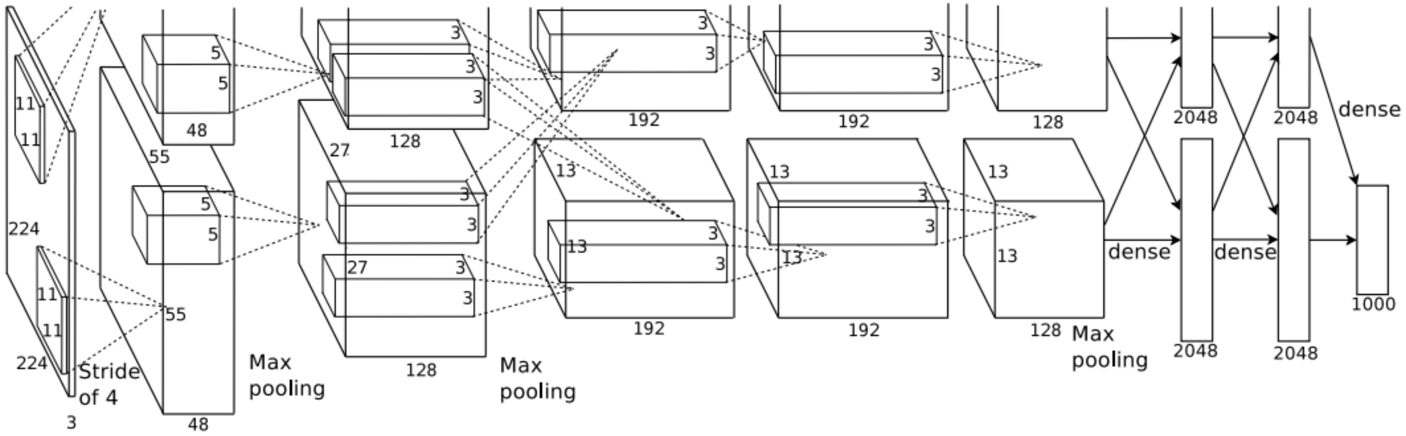


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

2017 Medical image analysis ruled by deep learning

Medical Image Analysis 42 (2017) 60–88

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Survey Paper

A survey on deep learning in medical image analysis

Geert Litjens*, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez

Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, The Netherlands

A R T I C L E   I N F O

Article history:
Received 27 February 2017
Revised 24 July 2017
Accepted 25 July 2017
Available online 26 July 2017

Keywords:
Deep learning
Convolutional neural networks
Medical imaging
Survey

A B S T R A C T

Deep learning algorithms, in particular convolutional networks, have rapidly become a methodology of choice for analyzing medical images. This paper reviews the major deep learning concepts pertinent to medical image analysis and summarizes over 300 contributions to the field, most of which appeared in the last year. We survey the use of deep learning for image classification, object detection, segmentation, registration, and other tasks. Concise overviews are provided of studies per application area: neuro, retinal, pulmonary, digital pathology, breast, cardiac, abdominal, musculoskeletal. We end with a summary of the current state-of-the-art, a critical discussion of open challenges and directions for future research.

Radboudumc

YouTube NL

Search

Geoff Hinton: On Radiology

Creative Destruction Lab

Subscribe 229

1,195 views

Add to Share More 10 3

Published on Nov 24, 2016

Geoff Hinton comments on radiology and deep learning at the 2016 Machine Learning and Market for Intelligence Conference in Toronto

Moderator: What do you think is the most exciting work to come?

Geoff Hinton: Let me start by just saying a few things that seem obvious.

I think if you work as a radiologist, you are like the coyote that is already over the edge of the cliff, but hasn’t yet looked down, so doesn’t realize there is no ground underneath him.

People should stop training radiologists now.

It is just completely obvious that within five years deep learning is going to do better than radiologists, because it is going to be able to obtain a lot more experience. It might be ten years, but we got plenty of radiologists already.

I said this to the hospital, and it didn’t go down too well.

[Hinton shrugs. Audience laughs.]

Was Hinton right or wrong?

- Hinton made two different statements
  - Within 5 (or 10) years deep learning is going to do better than radiologists
  - We should stop training radiologists now [in 2016/2017]
- Hinton was widely ridiculed and attacked for the second statement
- Far less discussion about the first statement

Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis

Urs J Muehlematter, Paola Daniore, Kerstin N Vokinger

There has been a surge of interest in artificial intelligence and machine learning (AI/ML)-based medical devices. However, it is poorly understood how and which AI/ML-based medical devices have been approved in the USA and Europe. We searched governmental and non-governmental databases to identify 222 devices approved in the USA and 240 devices in Europe. The number of approved AI/ML-based devices has increased substantially since 2015, with many being approved for use in radiology. However, few were qualified as high-risk devices. Of the 124 AI/ML-based devices commonly approved in the USA and Europe, 80 were first approved in Europe. One possible reason for approval in Europe before the USA might be the potentially relatively less rigorous evaluation of medical devices in Europe. The substantial number of approved devices highlight the need to ensure rigorous regulation of these devices. Currently, there is no specific regulatory pathway for AI/ML-based medical devices in the USA or Europe. We recommend more transparency on how devices are regulated and approved to enable and improve public trust, efficacy, safety, and quality of AI/ML-based medical devices. A comprehensive, publicly accessible database with device details for *Conformité Européene* (CE)-marked medical devices in Europe and US Food and Drug Administration approved devices is needed.

Lancet Digit Health 2021

Published Online January 18, 2021
https://doi.org/10.1016/S2589-7500(20)30292-2

Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Switzerland (U J Muehlematter MD); **and Institute of Law, University of Zurich, Switzerland** (P Daniore MSc, Prof K N Vokinger PhD)

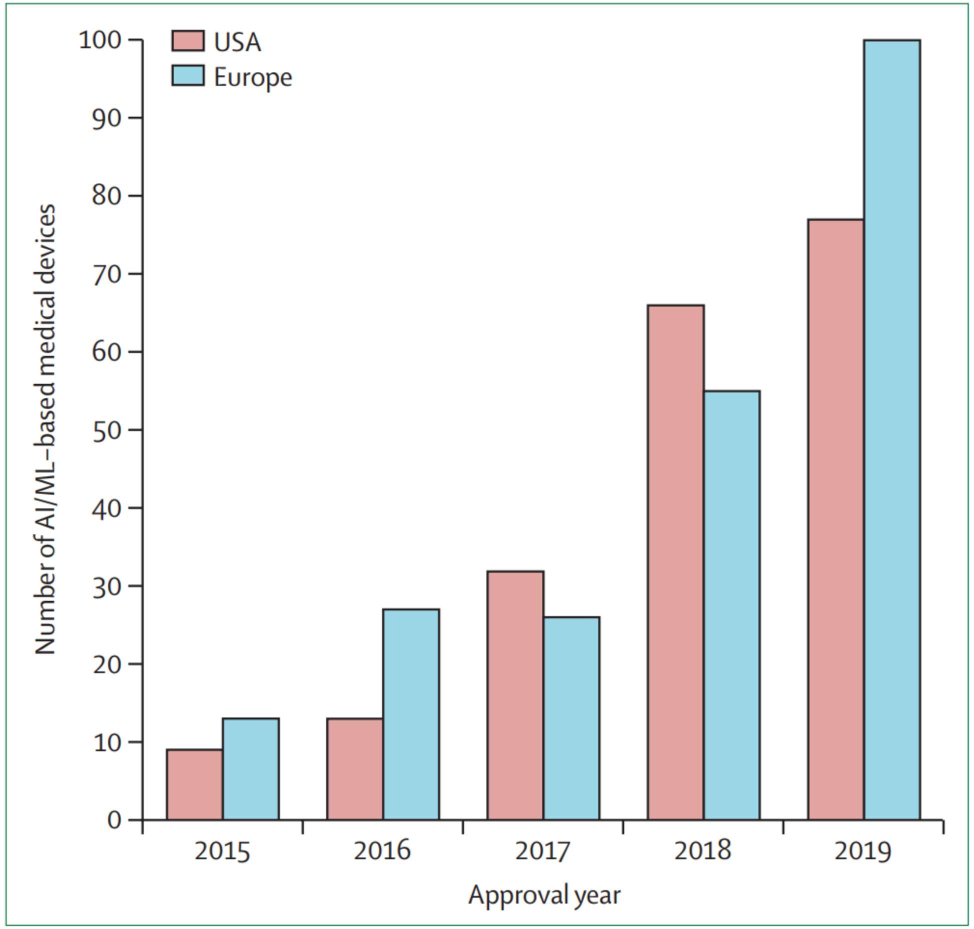


***Figure 2:* Number of approved (USA) and CE-marked (Europe) AI/ML-based medical devices between 2015 and 2019**

The CE-mark year is considered the approval year for devices in Europe. AI/ML=artificial intelligence and machine learning. CE=*Conformité Européenne*.

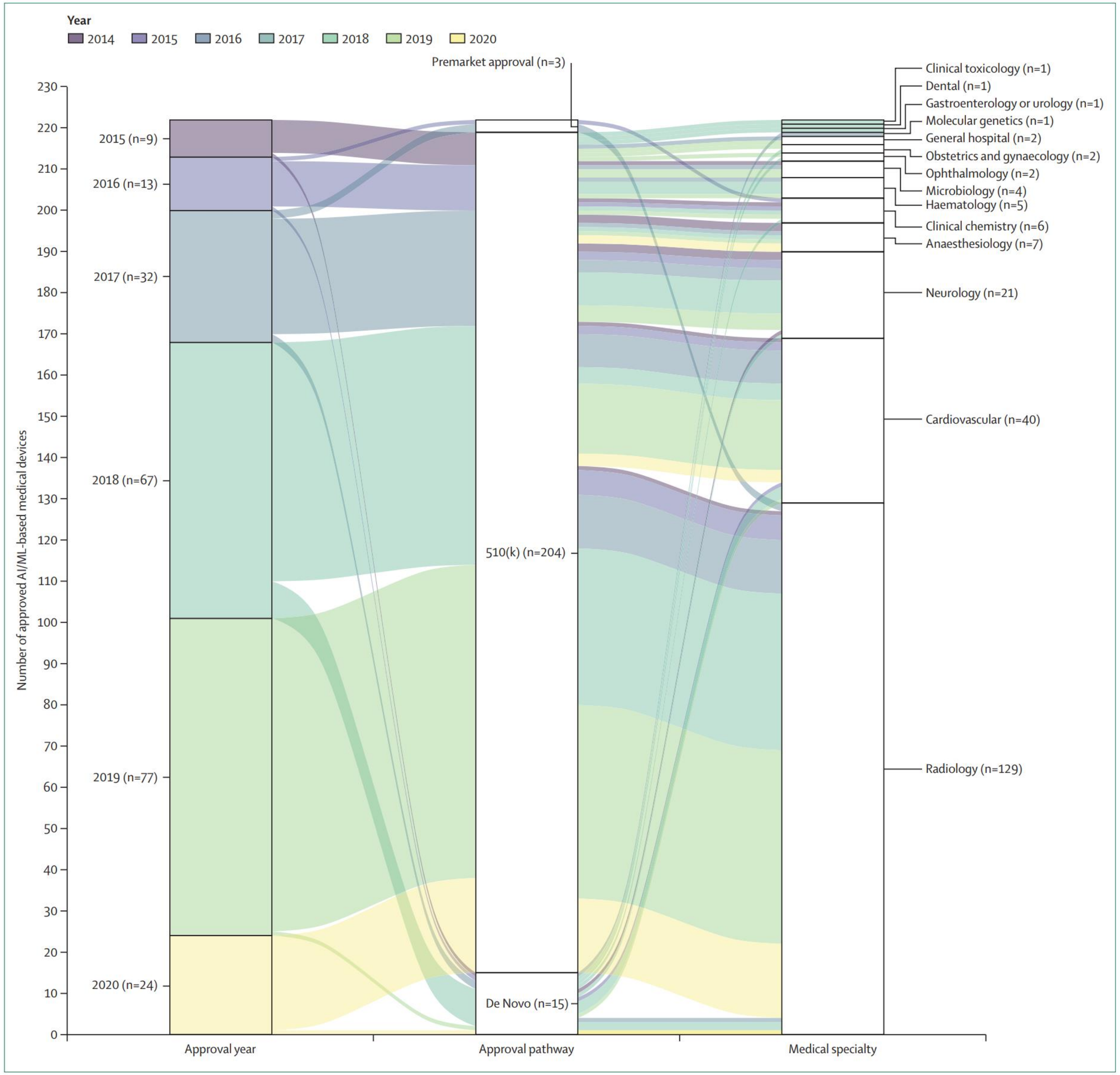


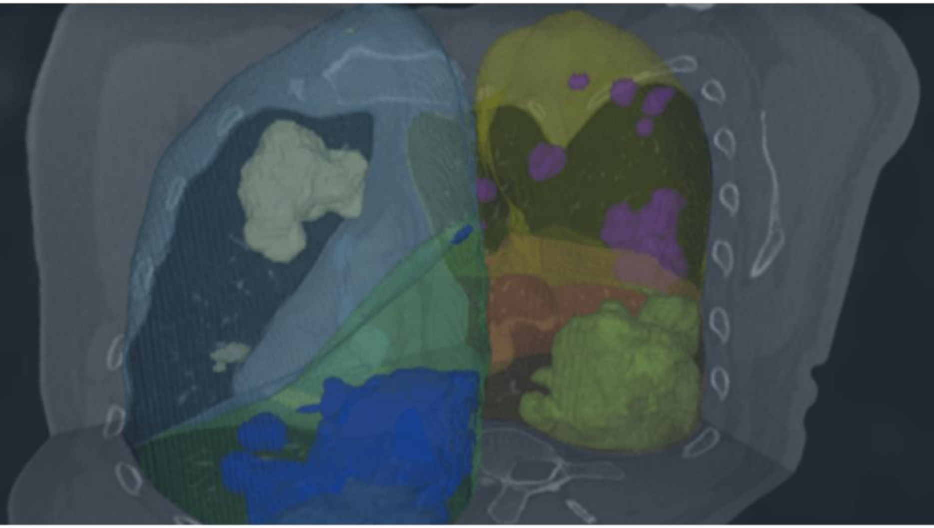
***Figure 3:* Sankey plot for AI/ML-based medical devices approved in the USA**

Data from Jan 1, 2015, to March 31, 2020. AI/ML=artificial intelligence and machine learning. NA=not applicable.

Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

68,000+ users | 290 challenges | 415 algorithms



Can you predict who will develop severe COVID-19 from a chest CT scan?

Last week, we opened STOIC2021: A COVID-19 AI challenge with 10,000 CT scans. Together with its participants, we aim to find the best solution for predicting who will develop severe COVID-19 from a chest CT scan. We will make the final solution easily accessible for everyone. In total, $20,000 in AWS Credits will be awarded to the winning…

FEATURED CHALLENGES

Participate in a challenge | Organize your own challenge

GC Grand Challenge

grand-challenge.org

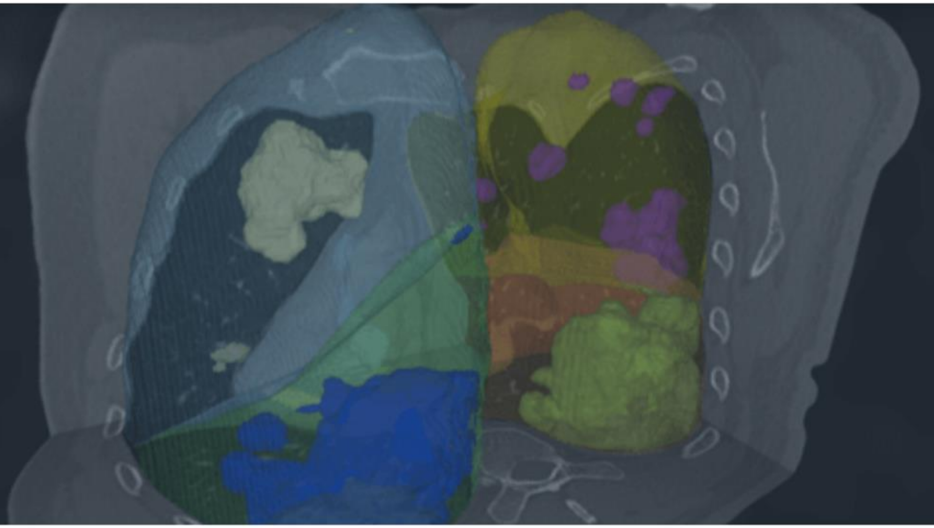
Grand Challenge Challenges Algorithms •••

Help

- Blog
- Reader Studies
- Archives
- Commercial products
- About
- Statistics
- Terms of Service

Grand Challeng

A platform for end-to-end development
solutions in biomedical imaging.



68,000+ users 290 challenges 415 algorithms

Can you predict who will develop severe COVID-19 from a chest CT scan?

Last week, we opened STOIC2021: A COVID-19 AI challenge with 10,000 CT scans. Together with its participants, we aim to find the best solution for predicting who will develop severe COVID-19 from a chest CT scan. We will make the final solution easily accessible for everyone. In total, $20,000 in AWS Credits will be awarded to the winning...

FEATURED CHALLENGES

Participate in a challenge Organize your own challenge

https://grand-challenge.org/aiforradiology/

Type here to search 4°C ENG US 10:13 AM 1/20/2022

Products

Find the artificial intelligence based software for radiology that you are looking for.
All products listed are available for the European market (CE marked).

Subspeciality: All | Modality: All | CE: All | CE class: All | FDA class: All

Search... Search

188/188 results

AI4MedImaging

AI4MedImaging

AI4CMR

Heart ventricles segmentation, quantification of ventricular volumes, myocardial mass, and ejection ...

The AI4CMR software performs a automatic cardiac segmentation and interpretation of Cardiac Magnetic Resonance (CMR), allowing quantification of various parameters without the need for intervention ...

Read more

Subspeciality: Cardiac
Modality: MR

CE: Class IIa - MDR
FDA:

Information source: Vendor
Certification verified: Yes

SIEMENS Healthineers

Siemens Healthineers

AI-Rad Companion Brain MR

brain volume quantification, segmentation, normative comparison, report generation

AI-Rad Companion Brain MR performs an automatic segmentation of the different brain areas, including an individual volumetric analysis. It compares the different volumes to a normative database and ...

Read more

Subspeciality: Neuro
Modality: MR

CE: Class IIa - MDR
FDA: Class II

Information source: Vendor
Certification verified: Yes

SIEMENS Healthineers

Siemens Healthineers

CE: Class IIa - MDR

Companies

Search... Search

List of companies providing the CE marked artificial intelligence solutions as listed in Products

87 results

Advantis

Advantis Medical Imaging makes advanced medical imaging more accessible, user-friendly and data-driven by merging it with cloud technology. Advantis is certified to ISO 13485, a quality management ...

Read more

HQ: Eindhoven, the Netherlands
Founded: 2016

Products:

Brainance MD

AI4MedImaging

The fully PACS/VNA integrated AI4CMR solution aims to automate the interpretation of cardiac magnetic resonance imaging (CMR). The AI4CMR cloud software (SaMD) performs a fully automatic cardiac ...

Read more

HQ: Braga, Portugal
Founded: 2019

Products:

AI4CMR

Aidence

Aidence rallies top-notch data scientists, medical professionals, and software engineers to bring deep-learning enabled solutions for automated medical image analysis. Our goal: empowering ...

Read more

HQ: Amsterdam, the Netherlands
Founded: 2015

Products:

Veye Lung Nodules

Aidoc

Aidoc develops the most advanced healthcare-grade AI based decision support software. Our technology analyzes medical imaging to provide the most comprehensive solution for detecting acute ...

Read more

Products:

Pulmonary embolism
C-Spine
Intracranial Hemorrhage
Large Vessel Occlusion
Incidental Pulmonary embolism

Benchmarking AI software

www.AIforRadiology.com/project-air

Project AIR

Multicenter data Dutch hospitals

Lung nodule detection
on
chest radiographs

OXIPIT InferVision Riverain TECHNOLOGIES MILVUE Lunit SAMSUNG JLK quibim qure.ai SIEMENS Healthineers VUNO

Bone age prediction
on
hand radiographs

BoneXpert image biopsy lab VUNO

Large vessel occlusion detection
on
brain CTA

aidoc Avicenna.AI NICOLAB viz.ai Brainomix iSchemaView RAPID

Radboudumc

Was Hinton right or wrong?

- Hinton made two different statements
  - Within 5 (or 10) years deep learning is going to do better than radiologists
  - We should stop training radiologists now [in 2016/2017]
- Hinton was widely ridiculed and attacked for the second statement
- Far less discussion about the first statement

- I believe the first statement is correct
  - Today, for almost all image interpretation tasks that radiologists perform, you can, fairly easily, build deep learning systems that performs this task as well

ORIGINAL RESEARCH • MUSCULOSKELETAL IMAGING

Radiology

Assessment of an AI Aid in Detection of Adult Appendicular Skeletal Fractures by Emergency Physicians and Radiologists: A Multicenter Cross-sectional Diagnostic Study

Loïc Duron, MD, MSc • Alexis Ducarouge, MSc • André Gillibert, MD, MSc • Julia Lainé, MD, MSc • Christian Allouche • Nicolas Cherel, MSc • Zekun Zhang, MSc • Nicolas Nitche, MSc • Elise Lacave, MSc • Aloïs Pourchot, MSc • Adrien Felter, MD • Louis Lassalle, MD, MSc • Nor-Eddine Regnard, MD, MSc • Antoine Feydy, MD, PhD

September 30, 2020; revision requested December 23;

This study was funded by Gleamer.

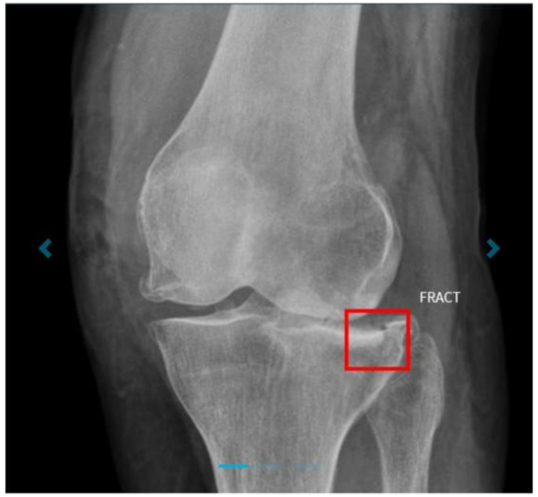
AI for Radiology
an implementation guide

Products   Companies

Products / BoneView

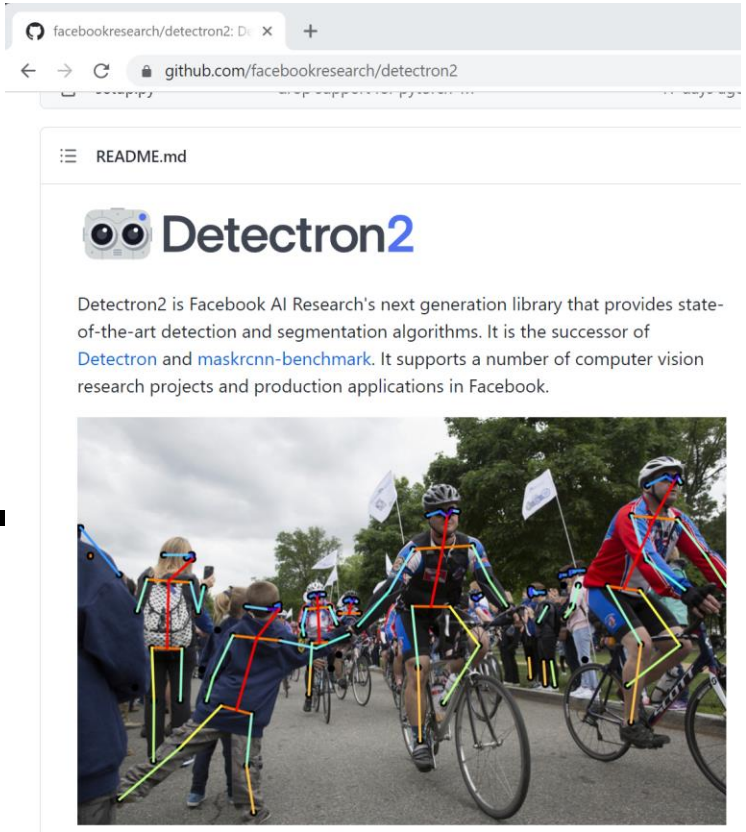
BoneView

GLEAMER



Model Building and Validation

We gathered a development data set of 60 170 radiographs in patients with trauma from 22 French public hospitals and private radiology departments from January 2011 to May 2019; this data set was randomly split into 70% training, 10% validation, and 20% internal test sets. A deep convolutional neural network based on the "Detectron 2" (31) framework was engineered, trained, optimized, and validated to detect and localize fractures on native resolution digital radiographs.



Was Hinton right or wrong?

- Hinton made two different statements
  - Within 5 (or 10) years deep learning is going to do better than radiologists
  - We should stop training radiologists now [in 2016/2017]
- Hinton was widely ridiculed and attacked for the second statement
- Far less discussion about the first statement

- I believe the first statement is correct
  - Today, for almost all image interpretation tasks that radiologists perform, you can, fairly easily, build deep learning systems that perform this task as well
  - Over time, building such systems will become easier

- The mistake Hinton made was to assume that statement 2 would be a logical and immediate consequence of statement 1

Healthcare

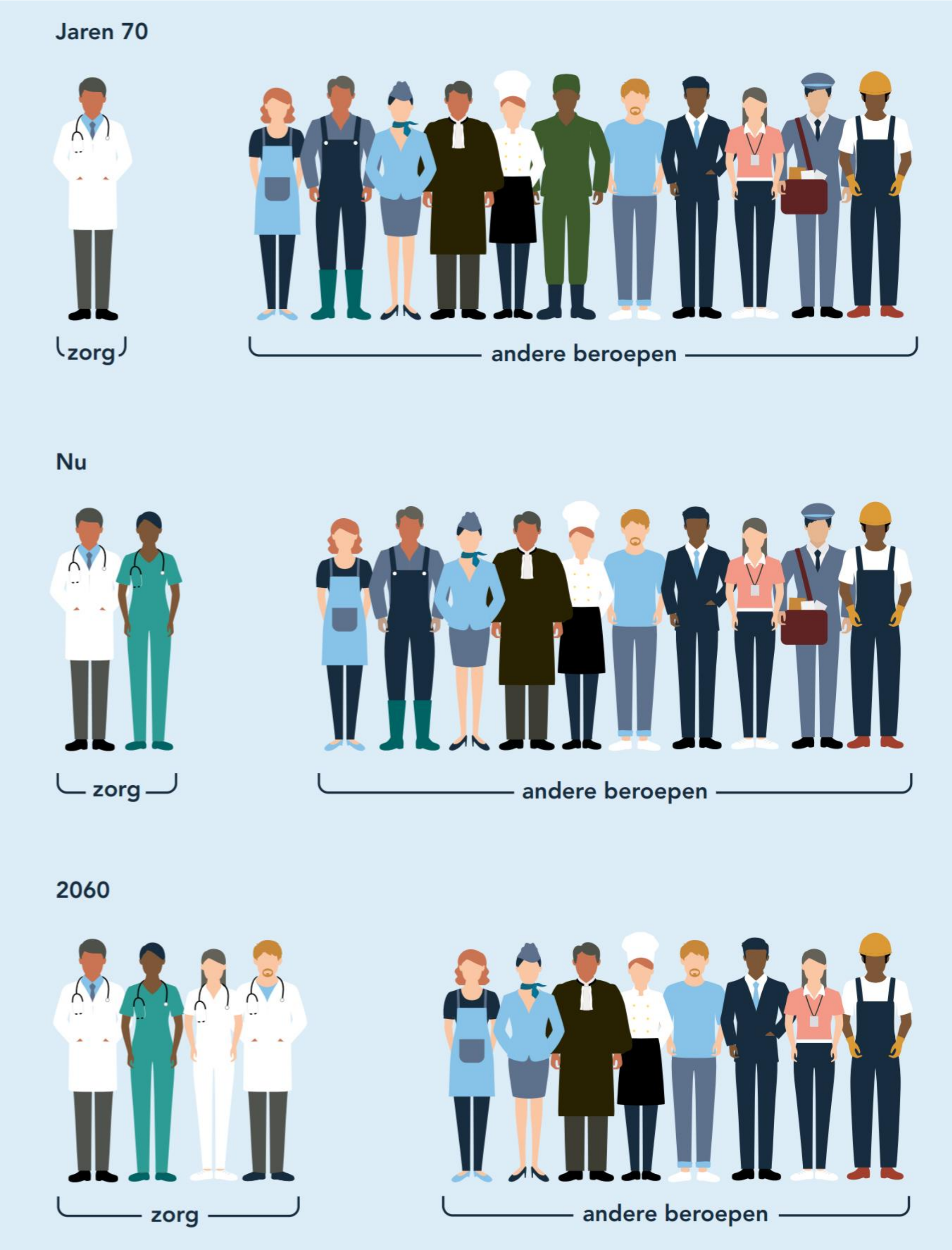
Radboudumc


Rapport
Kiezen voor houdbare zorg
Mensen, middelen en maatschappelijk draagvlak
WRR

1:12

1:6

1:3

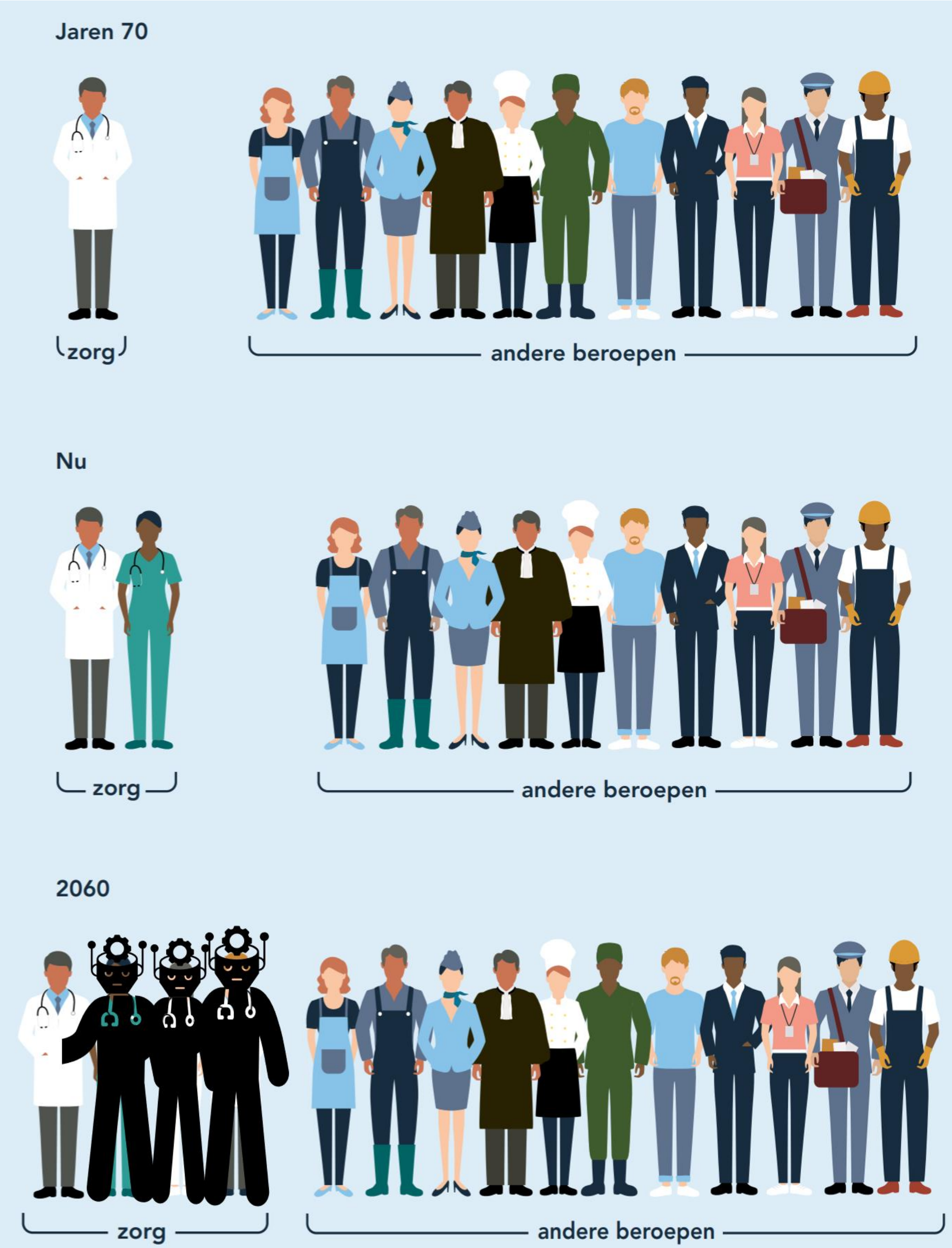

Jaren 70
zorg
andere beroepen
Nu
zorg
andere beroepen
2060
zorg
andere beroepen

Rapport
Kiezen voor houdbare zorg
Mensen, middelen en maatschappelijk draagvlak
WRR

1:12

1:6

1:12



Jaren 70
zorg
andere beroepen
Nu
zorg
andere beroepen
2060
zorg
andere beroepen

More affordable or more expensive?

- Most AI products for radiology on the market today aim to *assist* a radiologist
  - This will drive up costs as a hospital needs to buy the software **and** pay the salary of radiologists, unless the radiologists read faster (few studies address this; very few products targeting this)
  - Goal is quality improvement. In practice: modest sensitivity ↑ at equal specificity

- AI products that aim to *replace* tasks of human experts could potentially reduce costs
  - Goal is cost reduction and throughput increase; quality should not go down (too much)

- Why do companies with their products aim for assisting rather than replacing?
  - Business strategy: if you think of the radiologist as your customer, you are nice to them
  - Business strategy: start with assisting the customer, full automation could come later

- Strategy backfires: no reimbursement in Europe
- Alternative strategy: Focus on products that reduce workload for humans and save costs

IDx-DR

Close Care Gaps, Prevent Blindness

IDx-DR

IDx-DR is an AI diagnostic system that autonomously diagnoses patients for diabetic retinopathy and macular edema

With IDx-DR you get:

- Diagnostic results at the point-of-care
- No need for specialist overread or telemedicine call backs
- A simple user interface
- Customized workflow integration solutions

Evaluation of a System for Automatic Detection of Diabetic Retinopathy From Color Fundus Photographs in a Large Population of Patients With Diabetes

Michael D. Abràmoff, MD, PHD[1,2,3]
Meindert Niemeijer, PHD[3,4]
Maria S.A. Suttorp-Schulten, MD, PHD[5]
Max A. Viergever, PHD[4]
Stephen R. Russell, MD[1,2]
Bram van Ginneken, PHD[4]

OBJECTIVE — To evaluate the performance of a system for automated detection of diabetic retinopathy in digital retinal photographs, built from published algorithms, in a large, representative, screening population.

RESEARCH DESIGN AND METHODS — We conducted a retrospective analysis of 10,000 consecutive patient visits, specifically exams (four retinal photographs, two left and two right) from 5,692 unique patients from the EyeCheck diabetic retinopathy screening project imaged with three types of cameras at 10 centers. Inclusion criteria included no previous diagnosis of diabetic retinopathy, no previous visit to ophthalmologist for dilated eye exam, and both eyes photographed. One of three retinal specialists evaluated each exam as unacceptable quality, no referable retinopathy, or referable retinopathy. We then selected exams with sufficient image quality and determined presence or absence of referable retinopathy. Outcome measures included area under the receiver operating characteristic curve (number needed to miss one case [NNM]) and type of false negative.

RESULTS — Total area under the receiver operating characteristic curve was 0.84, and NNM was 80 at a sensitivity of 0.84 and a specificity of 0.64. At this point, 7,689 of 10,000 exams had sufficient image quality, 4,648 of 7,689 (60%) were true negatives, 59 of 7,689 (0.8%) were false negatives, 319 of 7,689 (4%) were true positives, and 2,581 of 7,689 (33%) were false positives. Twenty-seven percent of false negatives contained large hemorrhages and/or neovascularizations.

CONCLUSIONS — Automated detection of diabetic retinopathy using published algorithms cannot yet be recommended for clinical practice. However, performance is such that evaluation on validated, publicly available datasets should be pursued. If algorithms can be improved, such a system may in the future lead to improved prevention of blindness and vision loss in patients with diabetes.

Diabetes Care 31:193–198, 2008

Diabetic retinopathy blinds ~25,000 patients with diabetes annually in the U.S. alone and is the main cause of blindness in the U.S. and Europe working-age populations (1). Almost 50% of the 18 million patients with diabetes in the U.S do not undergo any form of regular documented dilated eye exam (2). This is in spite of overwhelming scientific evidence that this, if combined with appropriate management, can prevent up to 95% of cases of vision loss and blindness (3–10) and also in spite of guidelines by the American Diabetes Association and the American Academy of Ophthalmology that advise an annual dilated eye exam for most patients with diabetes (11). Digital photography of the retina examined by ophthalmologists or other qualified readers has been shown to have sensitivity and specificity comparable with or better than indirect ophthalmoscopy by an ophthalmologist (12,13) and has been proposed as an approach to make the dilated eye exam available to underserved populations that do not receive regular exams by ophthalmologists. If all of these populations were to be served with digital imaging, the number of retinal images to be evaluated annually is 32 million (~50% of patients with di-

Radiologists to FDA: Autonomous AI not ready for prime time

Posted 02 July 2020 | By Kari Oakes

Artificial intelligence is not ready for autonomy in radiology, according to two radiological professional associations who asked the US Food and Drug Administration (FDA) to wait for more rigorous testing and surveillance of the modality before authorizing its autonomous implementation in medical imaging.

In follow-up to a February 2020 workshop focused on artificial intelligence (AI) in medical imaging, the chairs of the American College of Radiology (ACR) and the Radiological Society of North America (RSNA) said in a joint letter that they have "some concerns with the approaches suggested at the workshop by a number of researcher/developer presentations with respect to FDA authorization pathways for autonomously functioning AI algorithms in medical imaging."

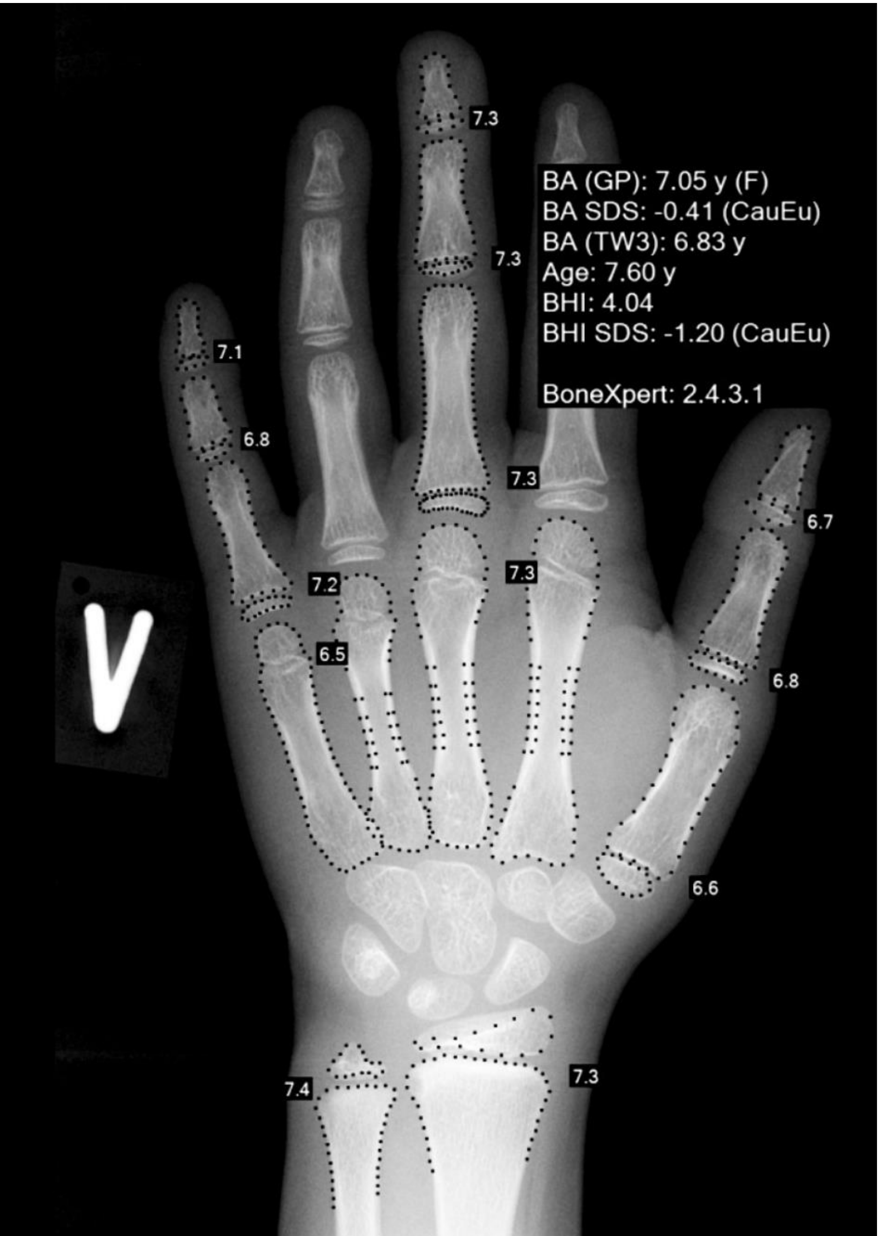
The two organizations "believe it is unlikely FDA could provide reasonable assurance of the safety and effectiveness of autonomous AI in radiology patient care without more rigorous testing, surveillance, and other oversight mechanisms throughout the total product life cycle," according to Howard B. Fleishon, MD, of ACR and Bruce G. Haffty, MD, of RSNA, who added that having AI perform autonomous image interpretation at a safe level "is a long way off."

Fleishon and Haffty advocated for FDA to hold off on further approvals until supervised AI algorithms in current use have broader market penetrance, so the agency can reach a better understanding of the efficacy and safety of these systems. This information can be used by FDA to formulate both the premarket approval and post-market surveillance processes for autonomous AI, they said.

Specifically, the letter calls for AI algorithms to be tested using multi-site heterogeneous data sets, "to ensure a minimum level of generalizability across diverse patient populations as well as variable imaging equipment and imaging protocols." Postmarket oversight by FDA should make sure that AI algorithms are working as expected over the long term, and labeling should be clear about what equipment and protocols are validated for use with the AI, Fleishon and Haffty said.

The associations' letter also referenced a 2019 discussion paper from FDA that proposes a regulatory framework for AI- and machine learning-based software as medical devices (SaMD). Some SaMD is "locked," whereas others use machine learning techniques to be "continuously learning." With regard to continuously adaptive algorithms, Fleishon and Haffty said, "we believe that without the safeguards provided by direct physician-expert oversight

BoneXpert



How radiologists use BoneXpert

- 10% performs full manual rating first and uses BoneXpert as a second reading
- 40% relies on BoneXpert's rating and looks at the image only for other findings (dysplasias, syndromes, fractures). Saves a lot of time
- 50% no longer looks at the images - instead, the referring physician looks up the BoneXpert bone age value shortly after the image has been recorded. Gives excellent workflow for the paediatricians

Source: Hans Henrik Thodberg, 2017

Radboudumc

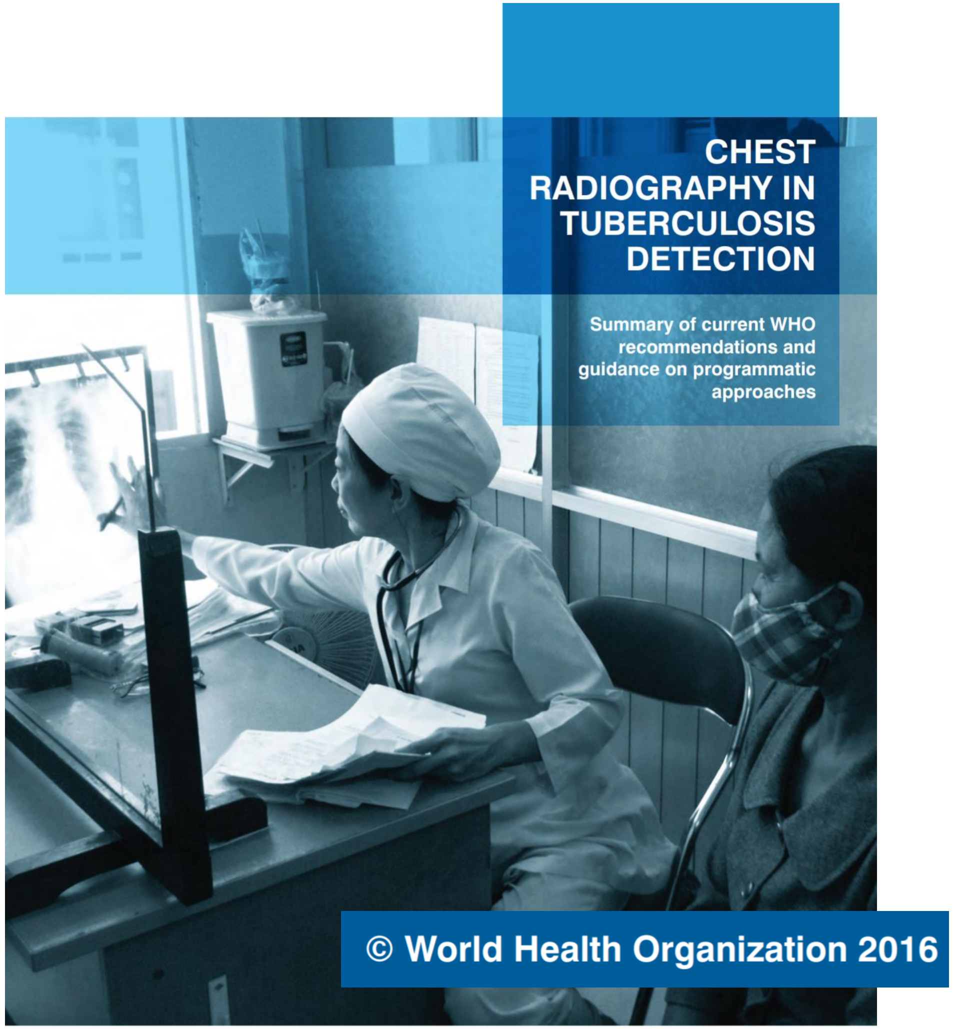
2014

CAD4TB

Artificial intelligence for the detection of tuberculosis

- CE certified
- 40+ publications
- Activated in 30+ countries
- Screening 5,000+ people per day

2016



5.3 Computer-aided detection of TB

New technologies for analyzing the results of CXR evaluations are being developed, including computer-aided detection (CAD) software that can analyze digital CXR images for abnormalities and the likelihood of TB being present. Such technology could help reduce interreader variability and delays in reading radiographs when skilled personnel are scarce.

As of 2016, WHO provides no recommendations on using CAD for TB. A systematic review of five peer reviewed articles published in 2016 concluded that the evidence of CAD's diagnostic accuracy is limited by the small number of studies of the single commercially available CAD software (CAD4TB, Delft Imaging

THE END TB STRATEGY

World Health Organization

Radboudumc

Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms

Zhi Zhen Qin, Shahriar Ahmed, Mohammad Shahnewaz Sarker, Kishor Paul, Ahammad Shafiq Sikder Adel, Tasneem Naheyan, Rachael Barrett, Sayera Banu, Jacob Creswell**

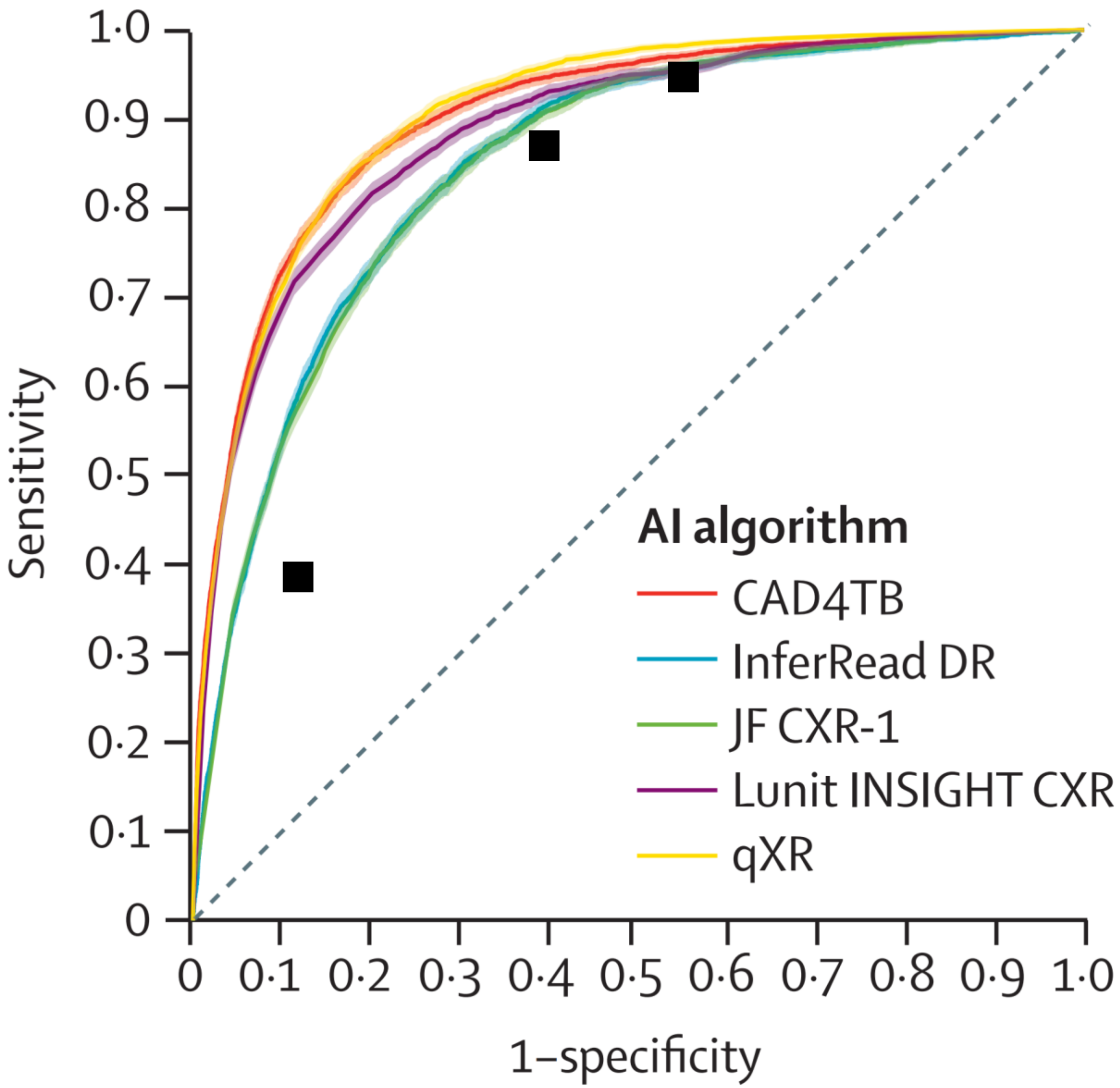
Summary

Background Artificial intelligence (AI) algorithms can be trained to recognise tuberculosis-related abnormalities on chest radiographs. Various AI algorithms are available commercially, yet there is little impartial evidence on how their performance compares with each other and with radiologists. We aimed to evaluate five commercial AI algorithms for triaging tuberculosis using a large dataset that had not previously been used to train any AI algorithms.

Lancet Digit Health 2021; 3: e543–54

See **Comment** page e535

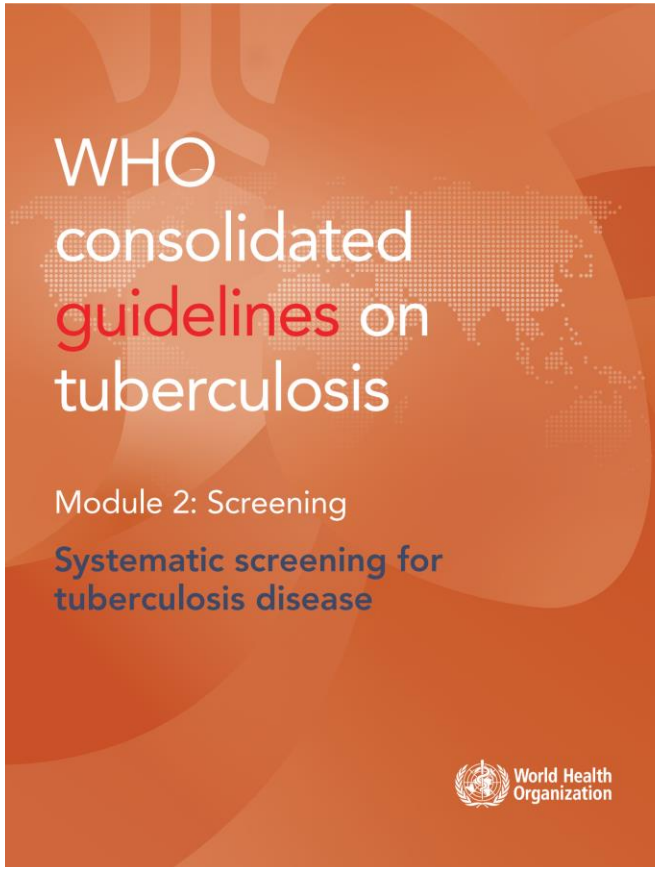
For the Bengali translation of the abstract see **Online** for appendix 1



■ Average of three radiologists at low, medium and high sensitivity settings

Radboudumc

2021



3.2 Use of computer-aided detection software for automated reading of digital chest radiographs

10. Among individuals aged 15 years and older in populations in which TB screening is recommended, computer-aided detection software programmes may be used in place of human readers for interpreting digital chest X-rays for screening and triage for TB disease

(new recommendation: conditional recommendation, low certainty of evidence).

Additional requirements for autonomous AI

- AI has to decide which cases it can process
  - Reject low quality images
  - Delegate out-of-distribution data to human experts
  - Delegate borderline/difficult cases to human experts
- AI needs to be monitored for (catastrophic) mistakes
  - If there is an accident you should learn from this (mistakes should become training data)