Artificial intelligence in medicine development – without the hot air

Lygature Partnership Meetup
Utrecht
October 29, 2019

Bram van Ginneken

Diagnostic Image Analysis Group, Radboud University Medical Center; Fraunhofer MEVIS, Bremen; Thirona, Nijmegen

Radboudumc

Diagnostic Image Analysis Group

The Diagnostic Image Analysis Group is part of the Departments of Radiology and Nuclear Medicine, Pathology, and Ophthalmology of Radboud University Medical Center. We develop computer algorithms to aid clinicians in the interpretation of medical images and thereby improve the diagnostic process.

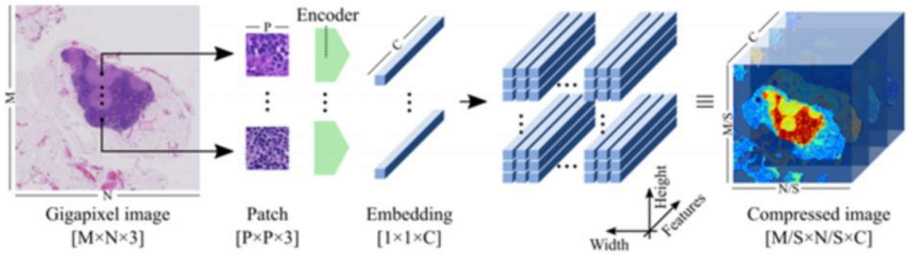
The group has its roots in computer-aided detection of breast cancer in mammograms, and we have expanded to automated detection and diagnosis in breast MRI, ultrasound and tomosynthesis, chest radiographs and chest CT, prostate MRI, neuro-imaging and the analysis of retinal and digital pathology images. The technology we primarily use is deep learning.

It is our goal to have a significant impact on healthcare by bringing our technology to the clinic. We are therefore fully certified to develop, maintain, and distribute software for analysis of medical images in a quality controlled environment (MDD Annex II and ISO 13485).

On this site you find information about the history of the group and our collaborations, an overview of people in DIAG, current projects, publications and theses, contact information, and info for those interested to join our team.

Highlights

August, 2019



David Tellez et al published a new method to train neural networks on gigapixel whole-slide images directly, avoiding the need for fine-grained annotations. This approach allows the neural network to discover new predictive features by using automatically derived 'annotations' such as molecular biomarkers or patient outcome.

In their work, published in IEEE Transactions on Pattern Analysis and Machine Intelligence, the authors present the neural image compression (NIC, top image)

Disclosures

- Developer CAD4TB (Delft Imaging Systems): royalties & funding
- Co-founder and CSO Thirona: stock, royalties & funding
- Developer Veolity (MeVis Medical Solutions) & DynaCAD Lung (InVivo): royalties & funding
- ScreenPoint is a spin-off from my group; my group receives royalties
- Group funding: Canon, Siemens Healthineers, Philips Medical, Sectra, Novartis

Artificial Intelligence in medicine development - without the hot air

- Developments in Artificial Intelligence (AI) are (…) drivers for a fourth industrial revolution, which will impact all human activities, including medicine development and health care.

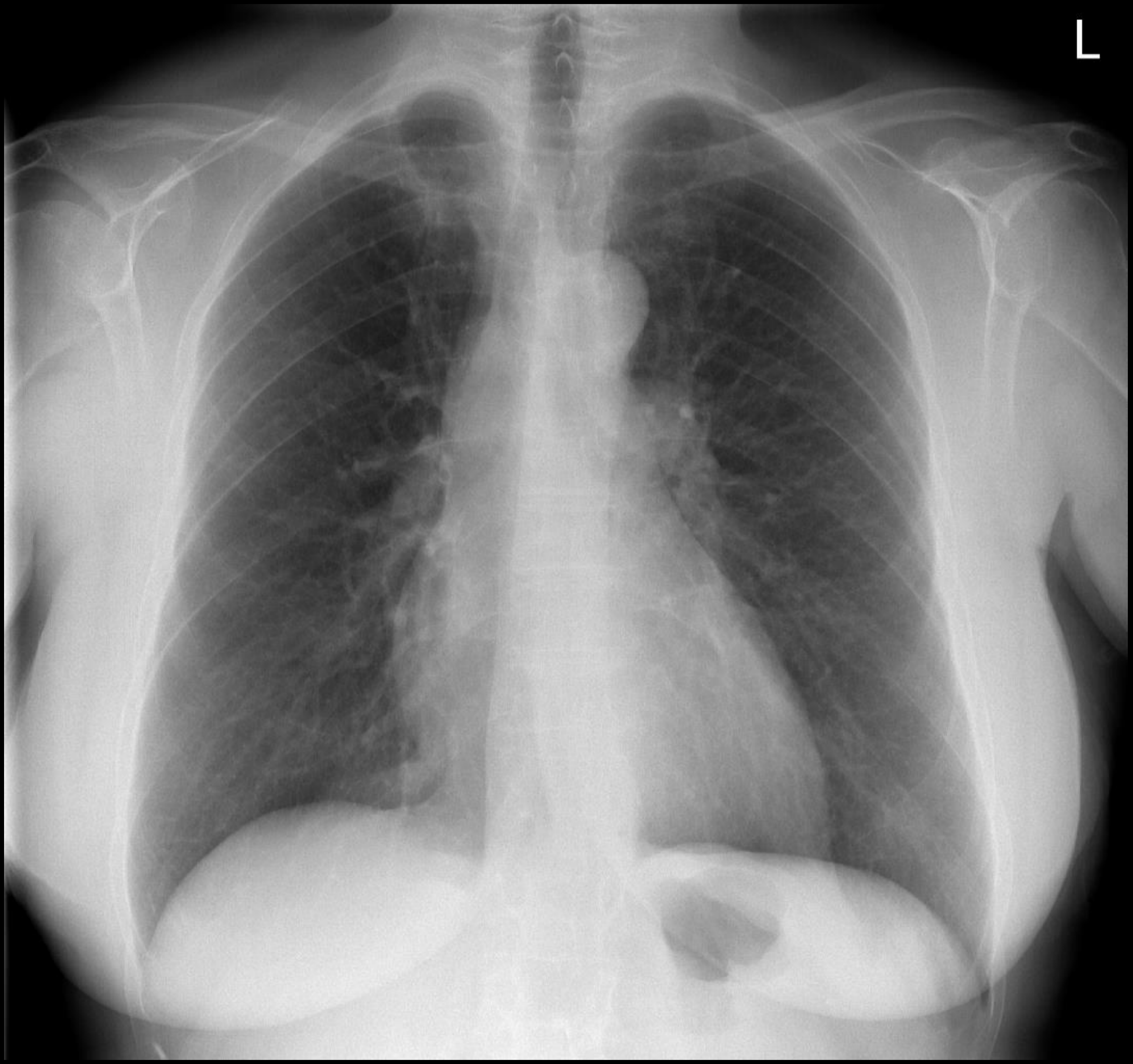
- There are obvious benefits for patients
  - More data about the patient that can be assessed
  - Freeing up time for health care professionals for human interaction
  - Including the full body of medical knowledge to identify the optimal treatment for a patient

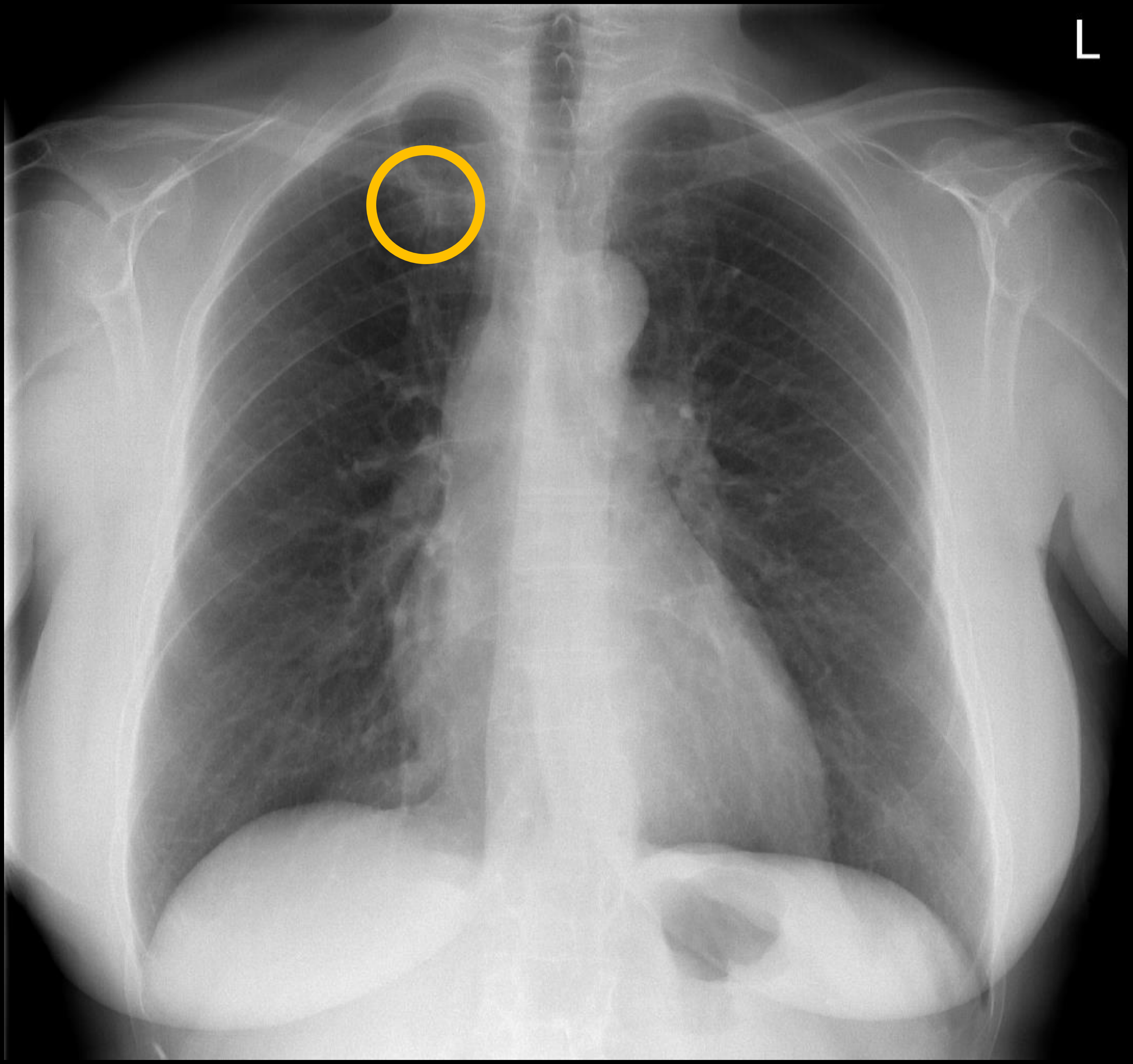
Video from ECR 2018 Wiro Niessen on how AI can help you

334 views • Mar 1, 2018 0 0 SHARE SAVE

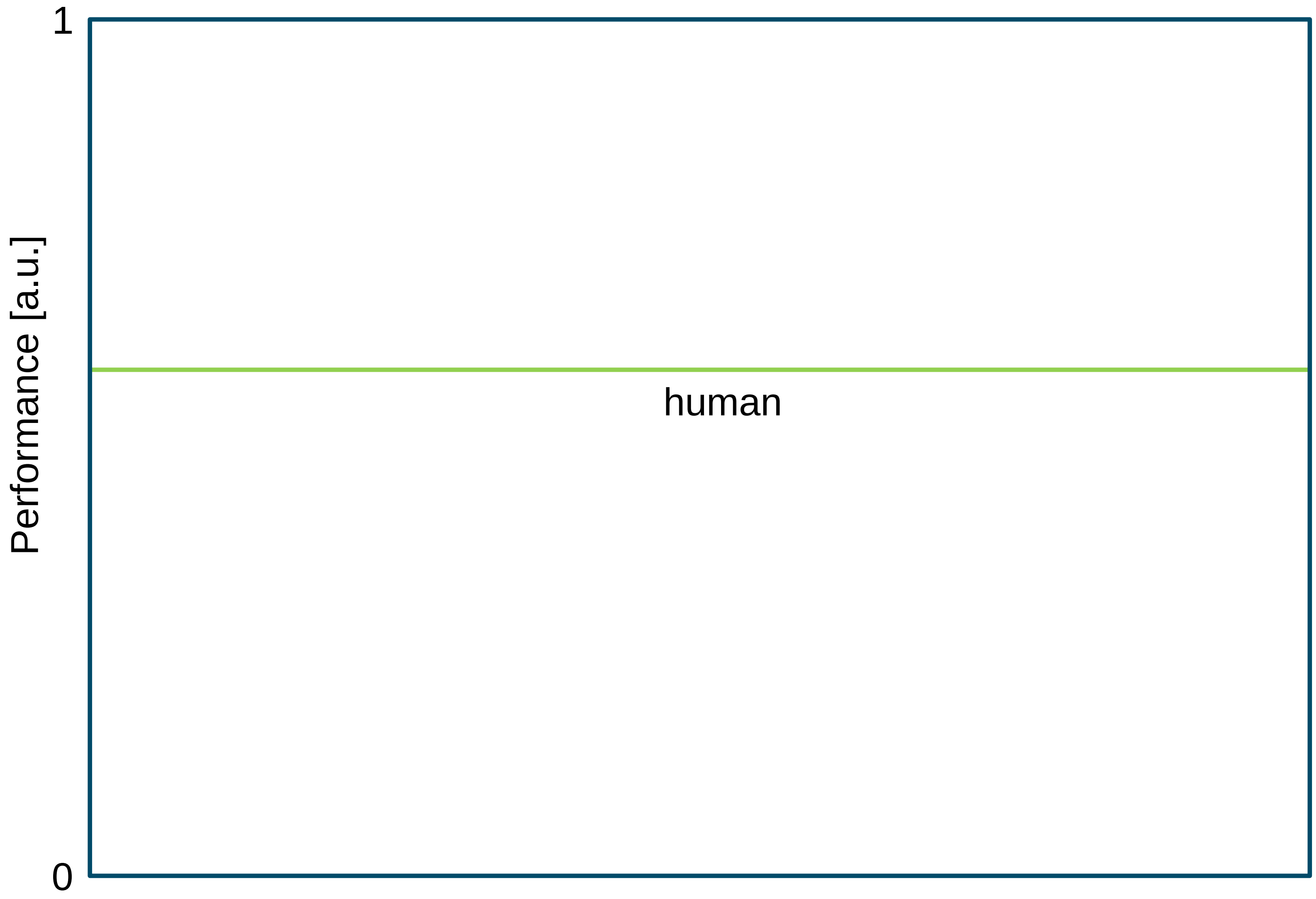
AI is an enormous opportunity for the field. We can learn so much from medical image data. Really the promise of precision medicine, to treat someone in the optimal way, we can try to reach that by incorporating AI. But we're not there yet …

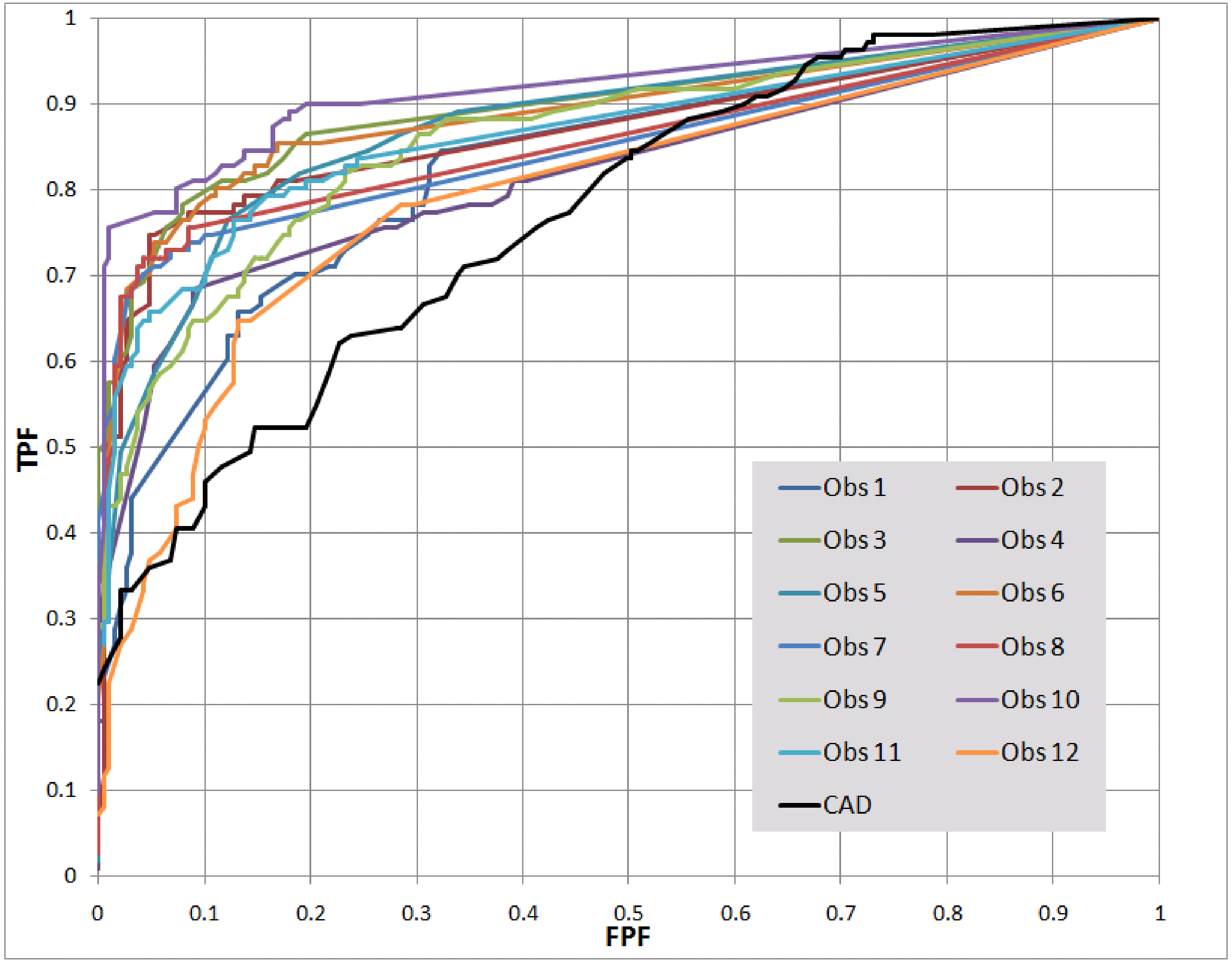
Radboudumc


L


L

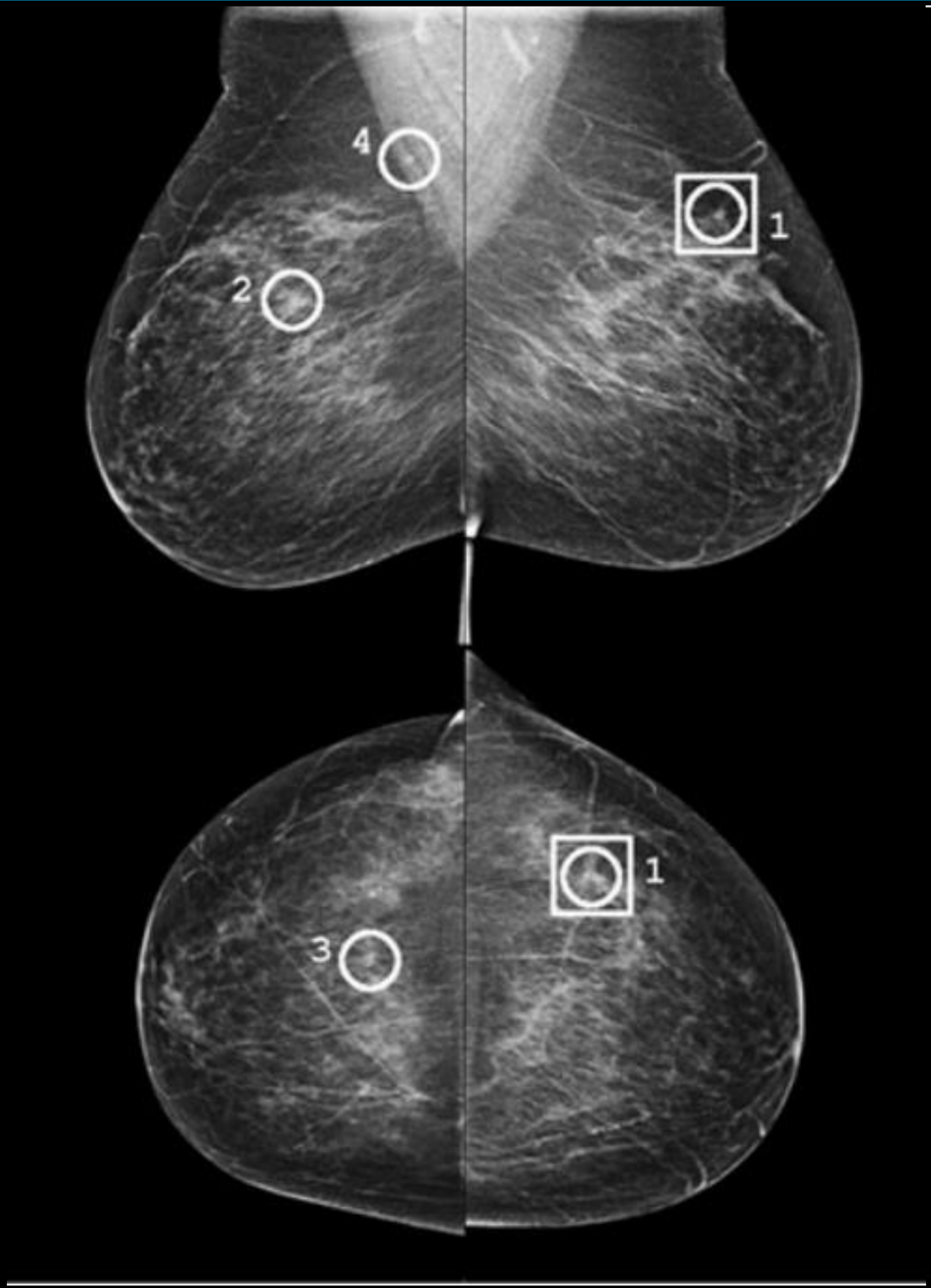
A detection/diagnosis/quantification task involving medical images

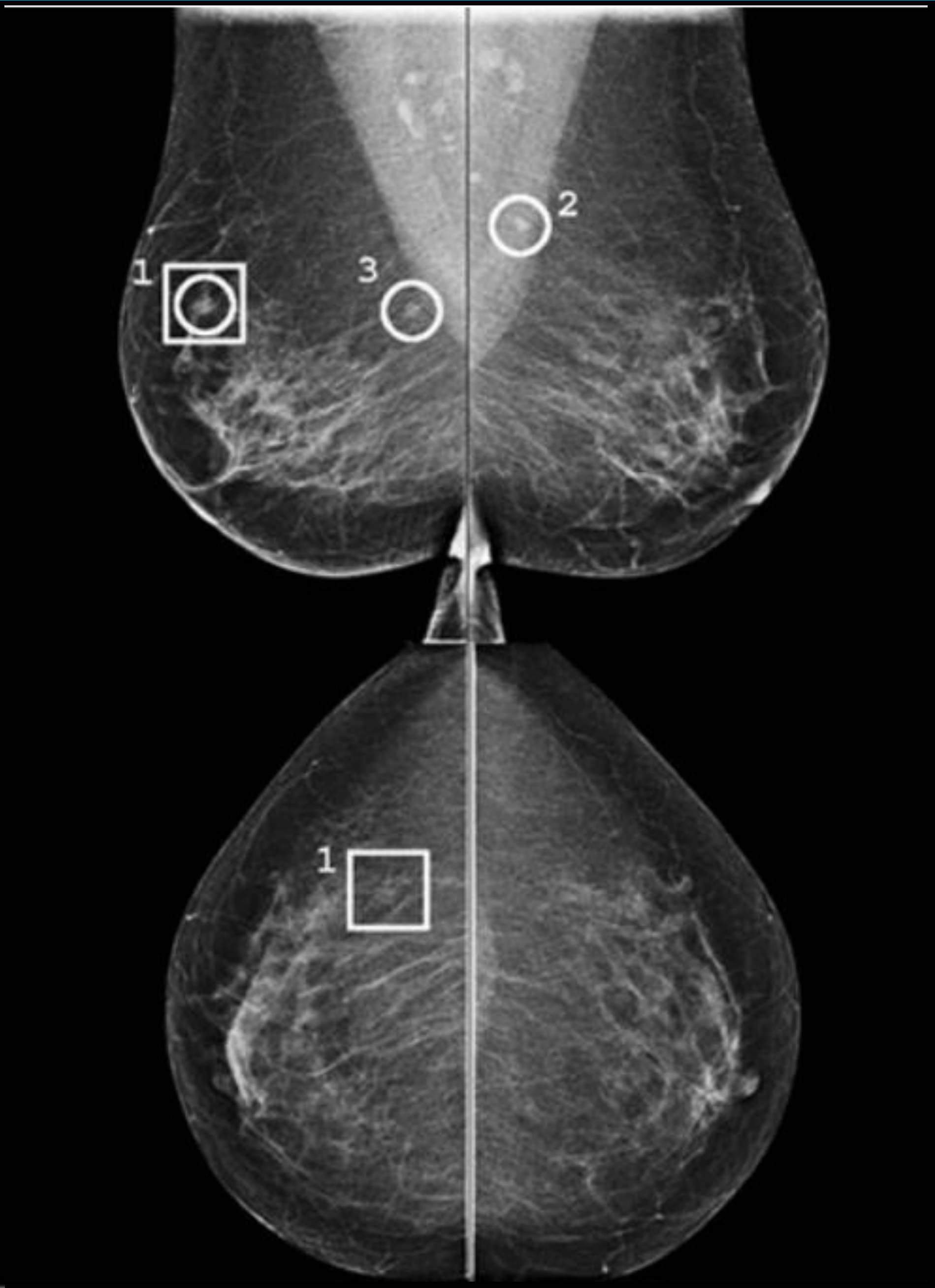


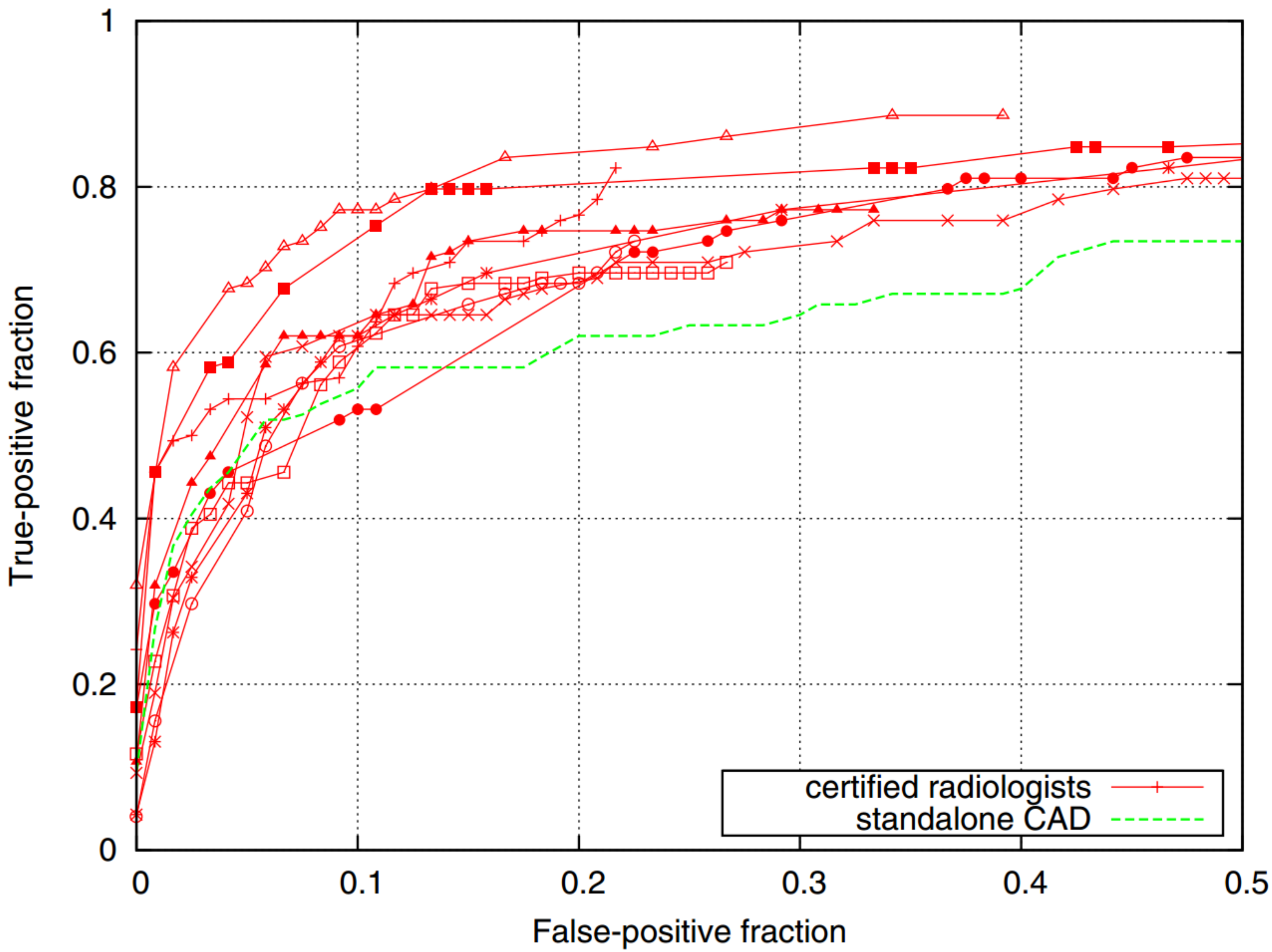


S. Schalekamp, B. van Ginneken, E. Koedam et al. Radiology 2014;272:252-261

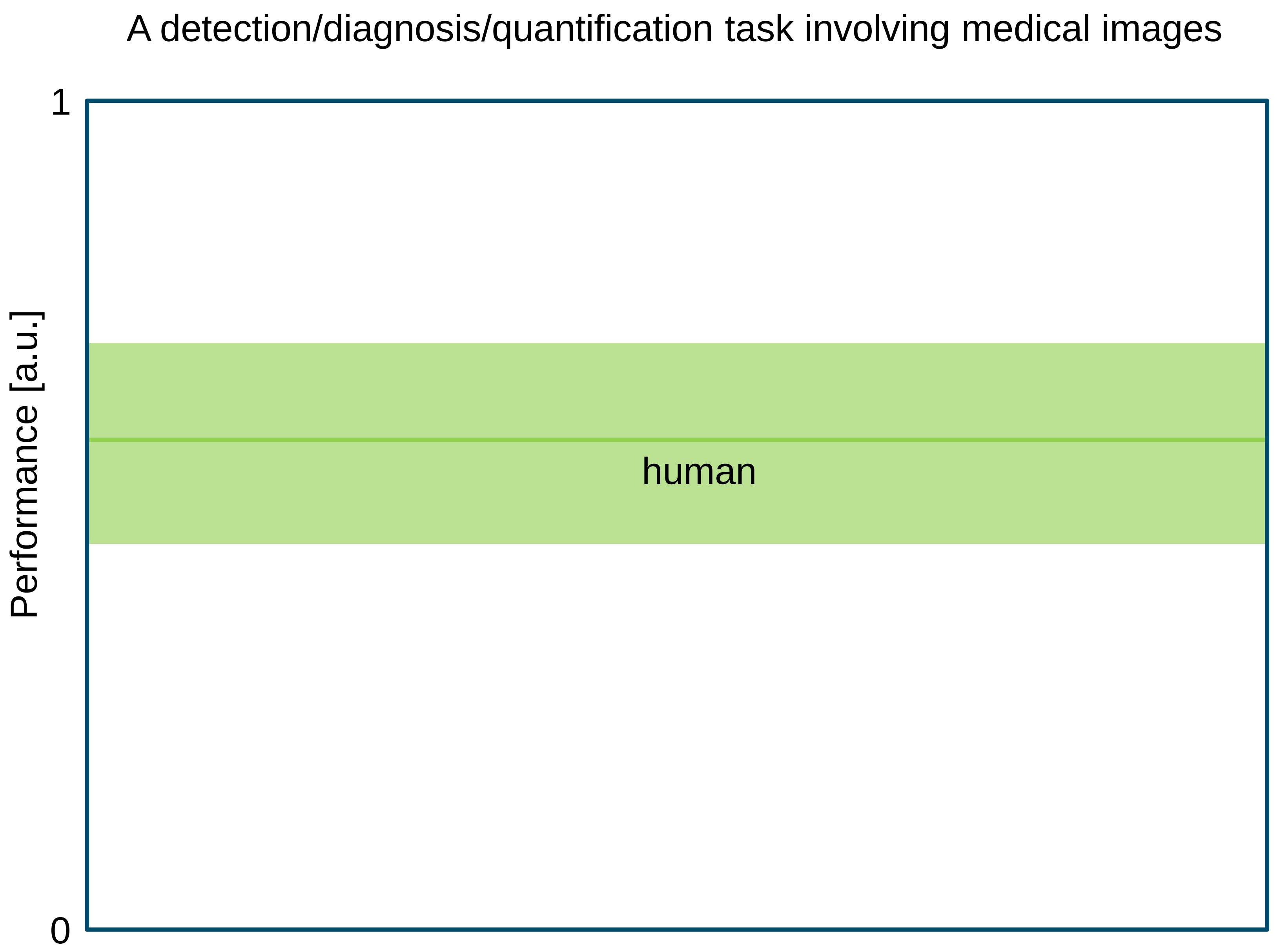
Radboudumc


4
1
2
1
3


2
1
3
1

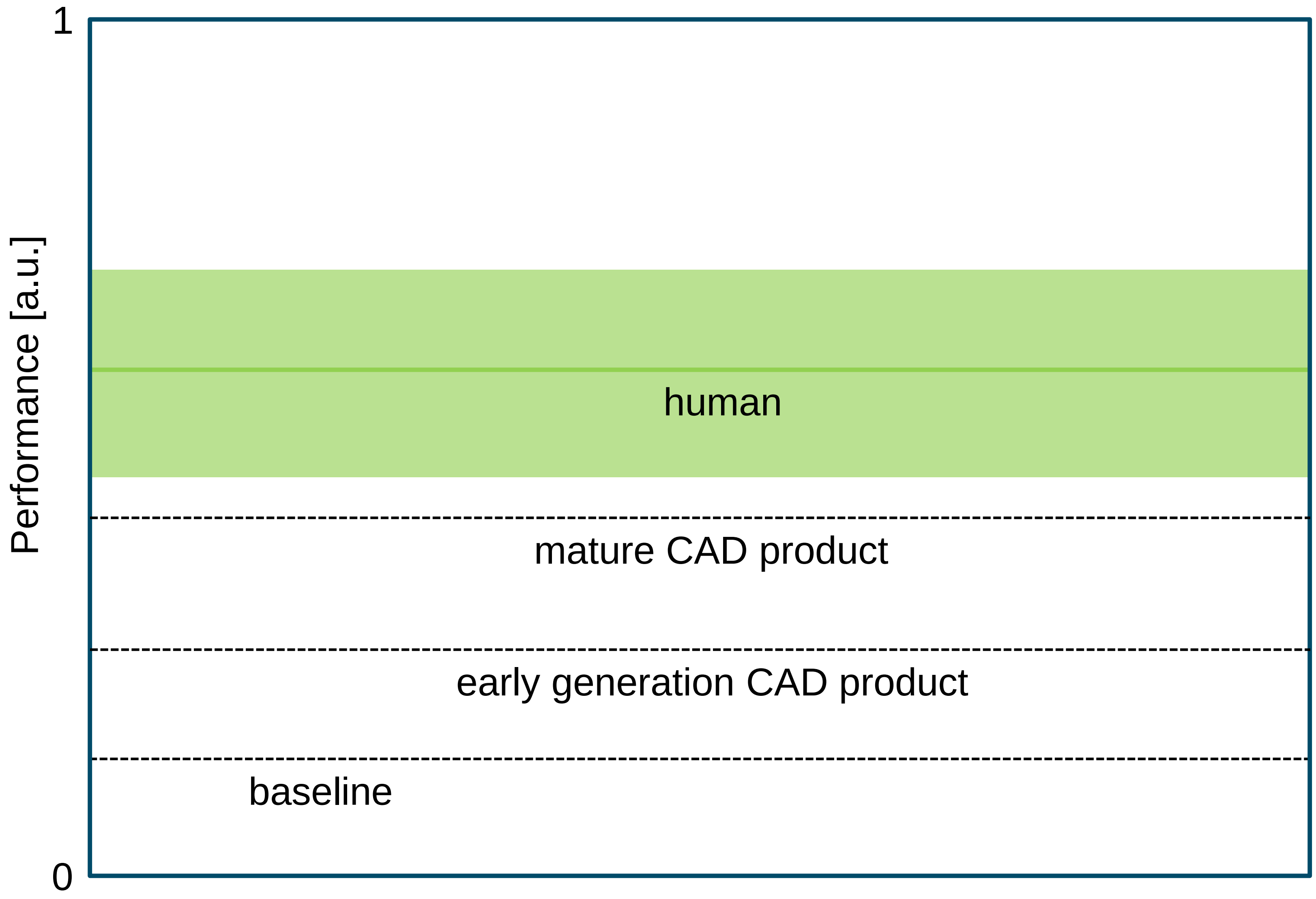


R. Hupse et al. European Radiology 2013;23:93-100


A detection/diagnosis/quantification task involving medical images
1
Performance [a.u.]
human
0

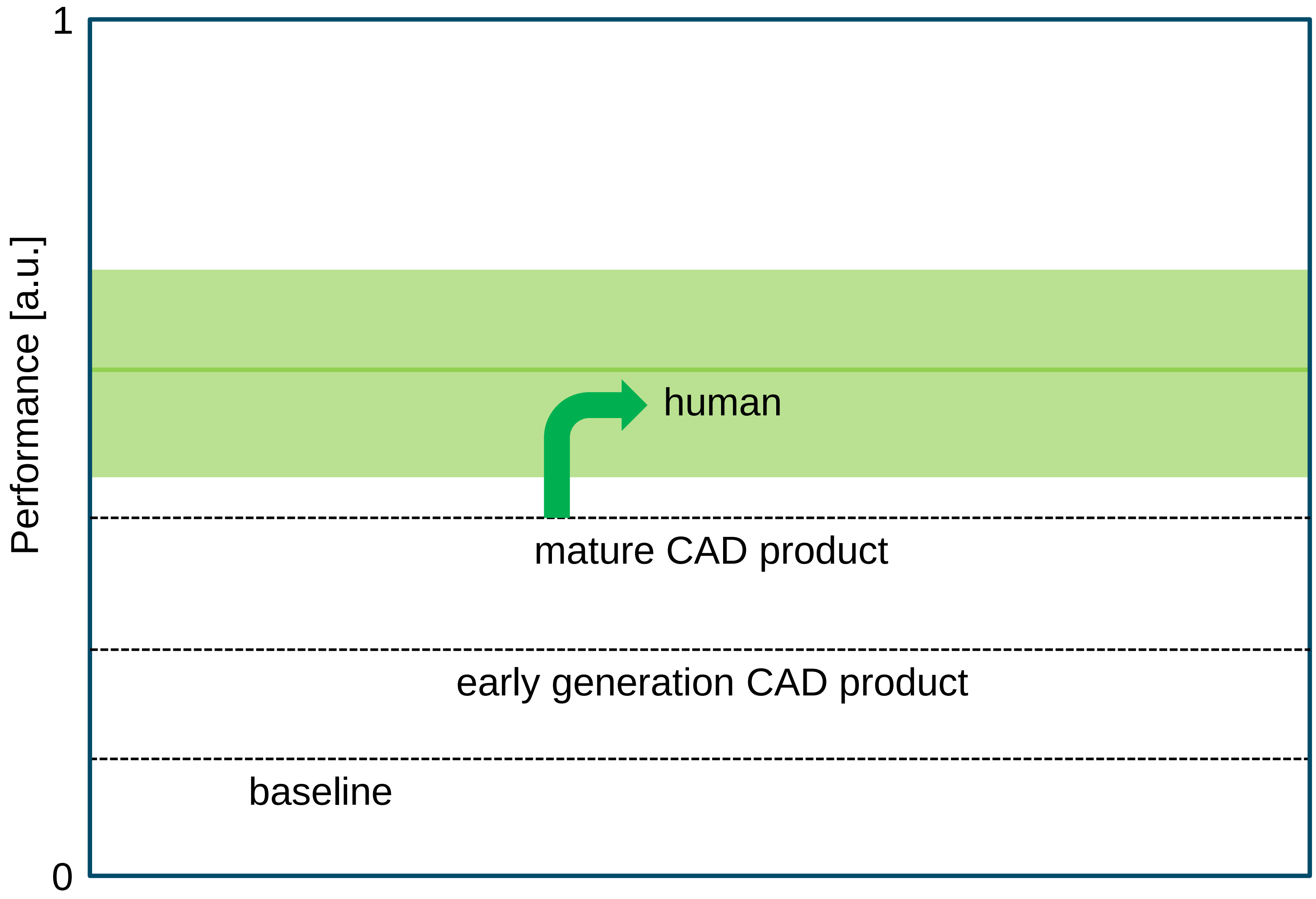
Radboudumc

A detection/diagnosis/quantification task involving medical images

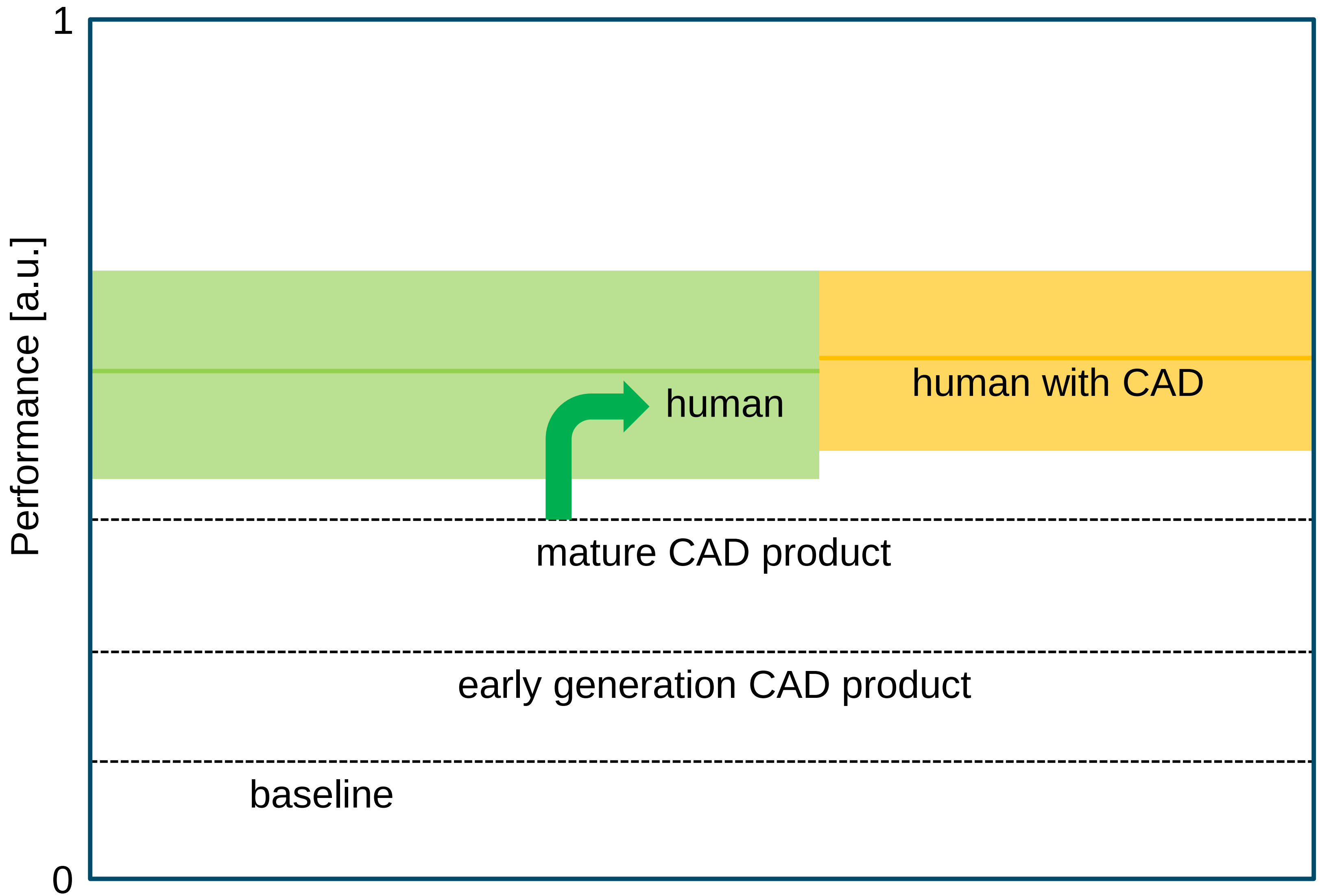


Radboudumc

A detection/diagnosis/quantification task involving medical images

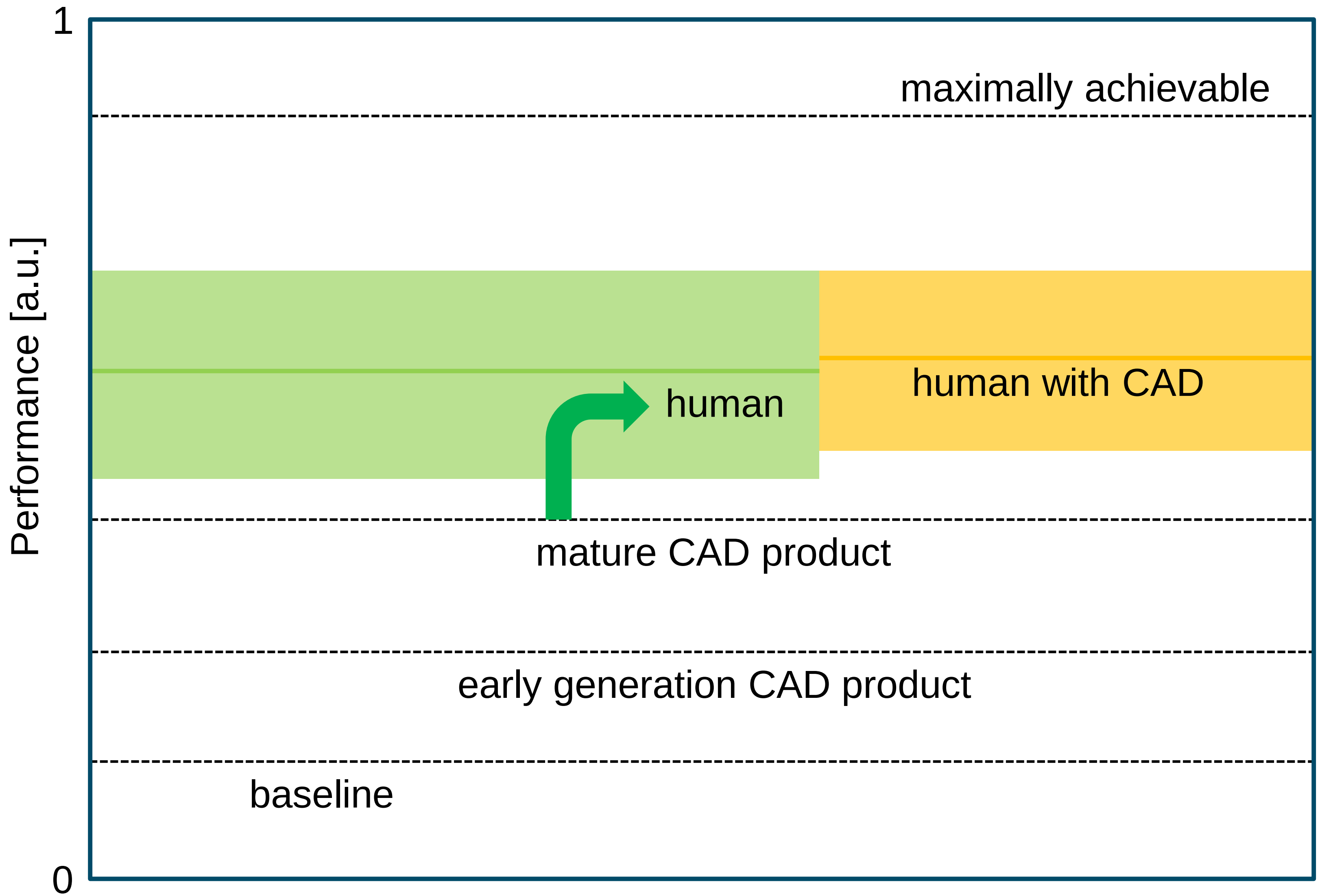


A detection/diagnosis/quantification task involving medical images



Radboudumc

A detection/diagnosis/quantification task involving medical images



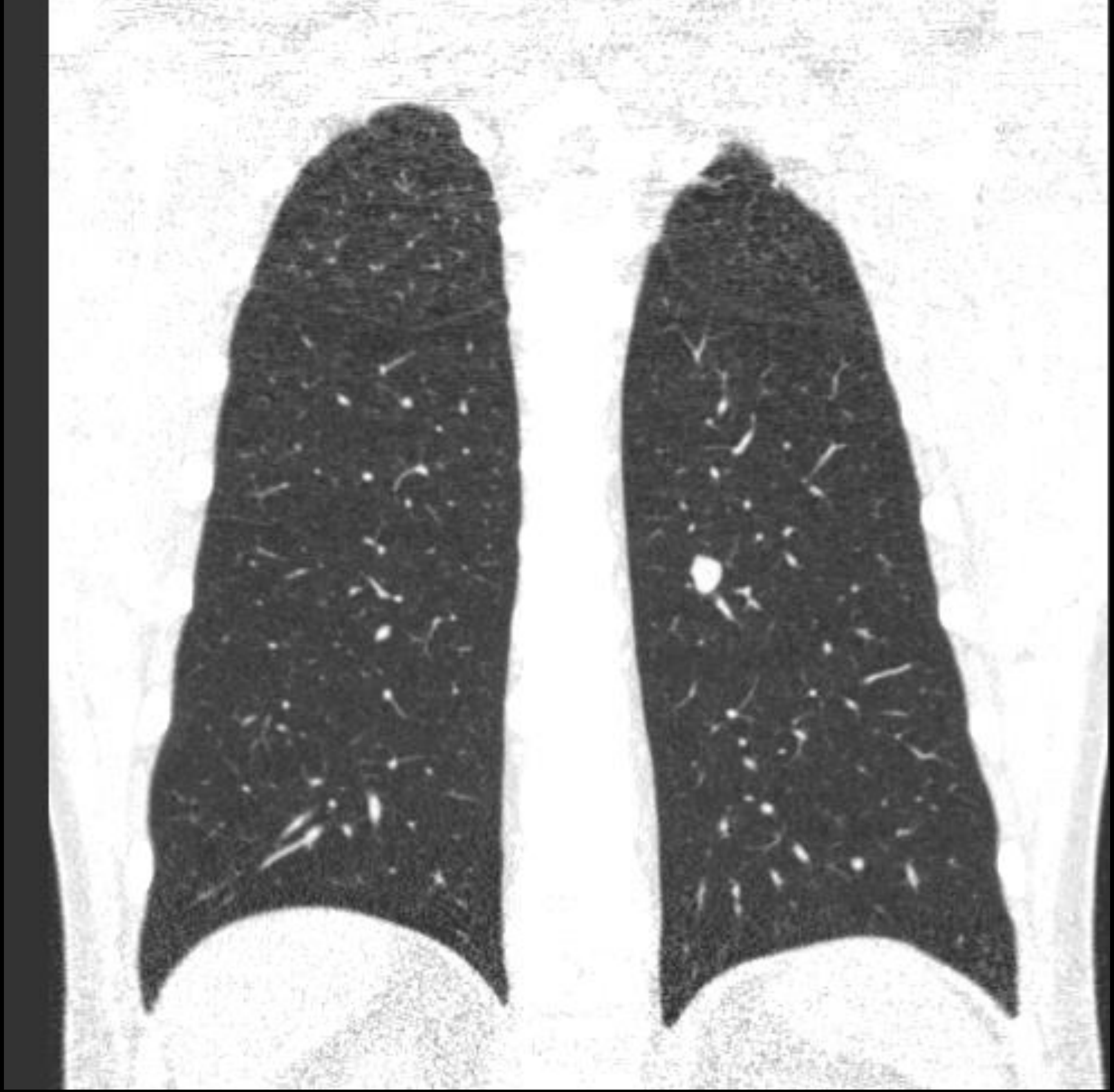
Radboudumc

Simulated
X-ray

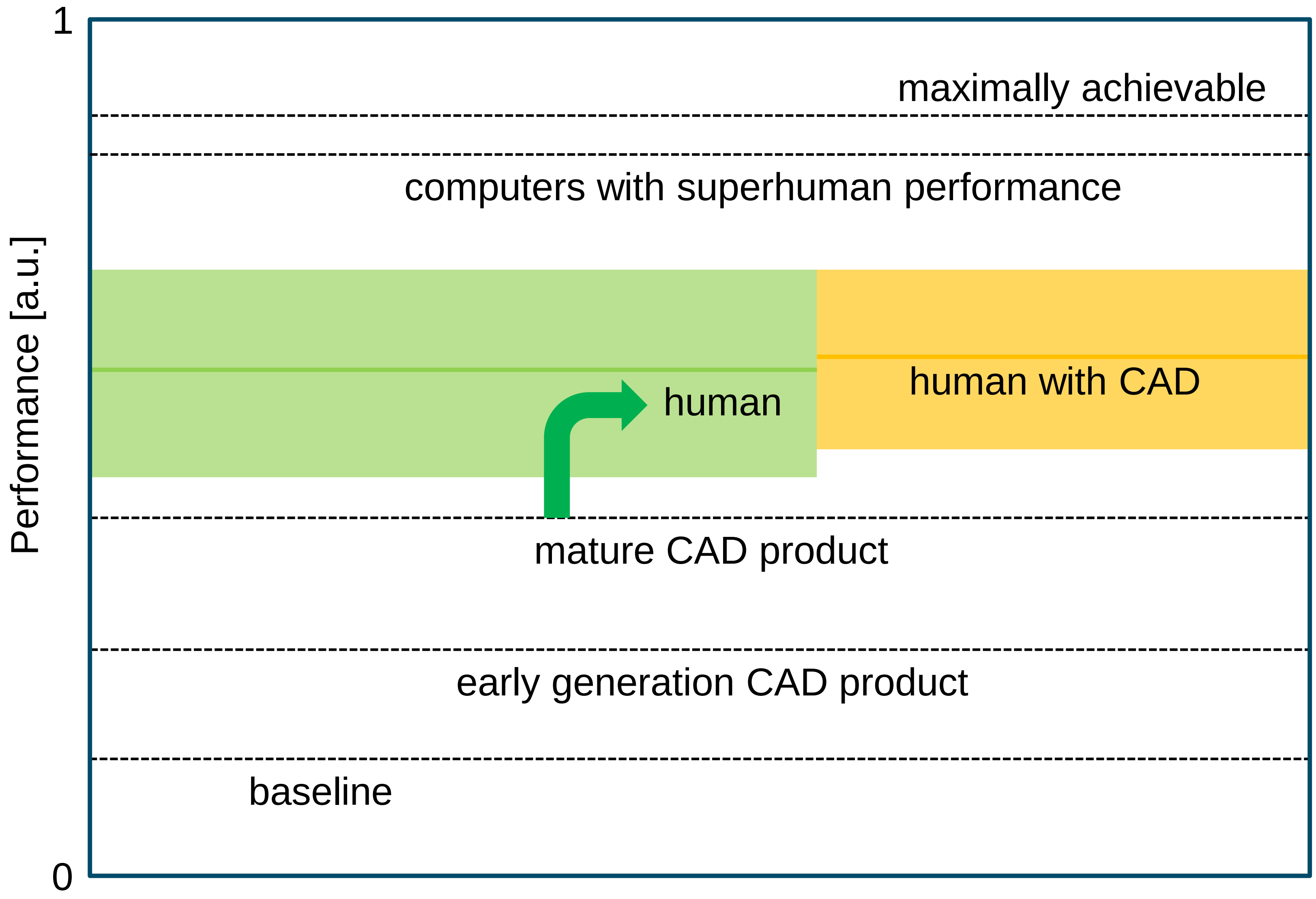
105
mm

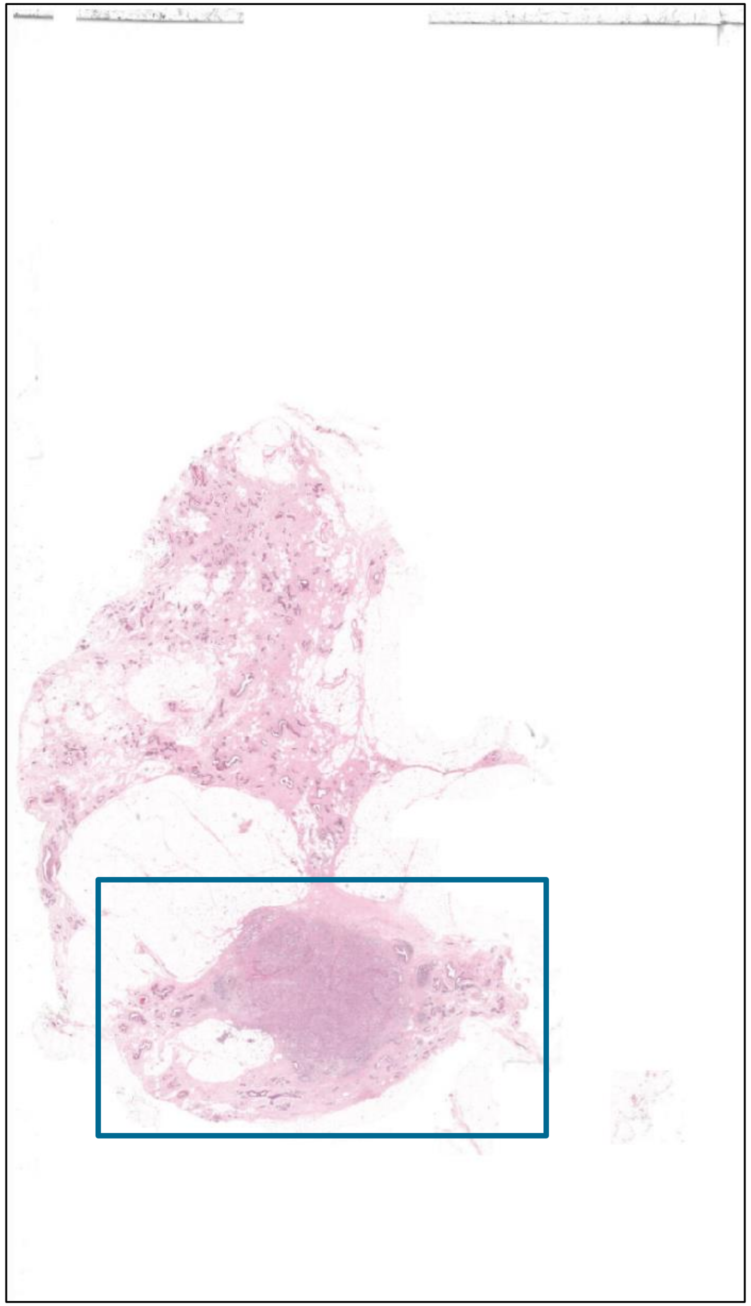
70
mm

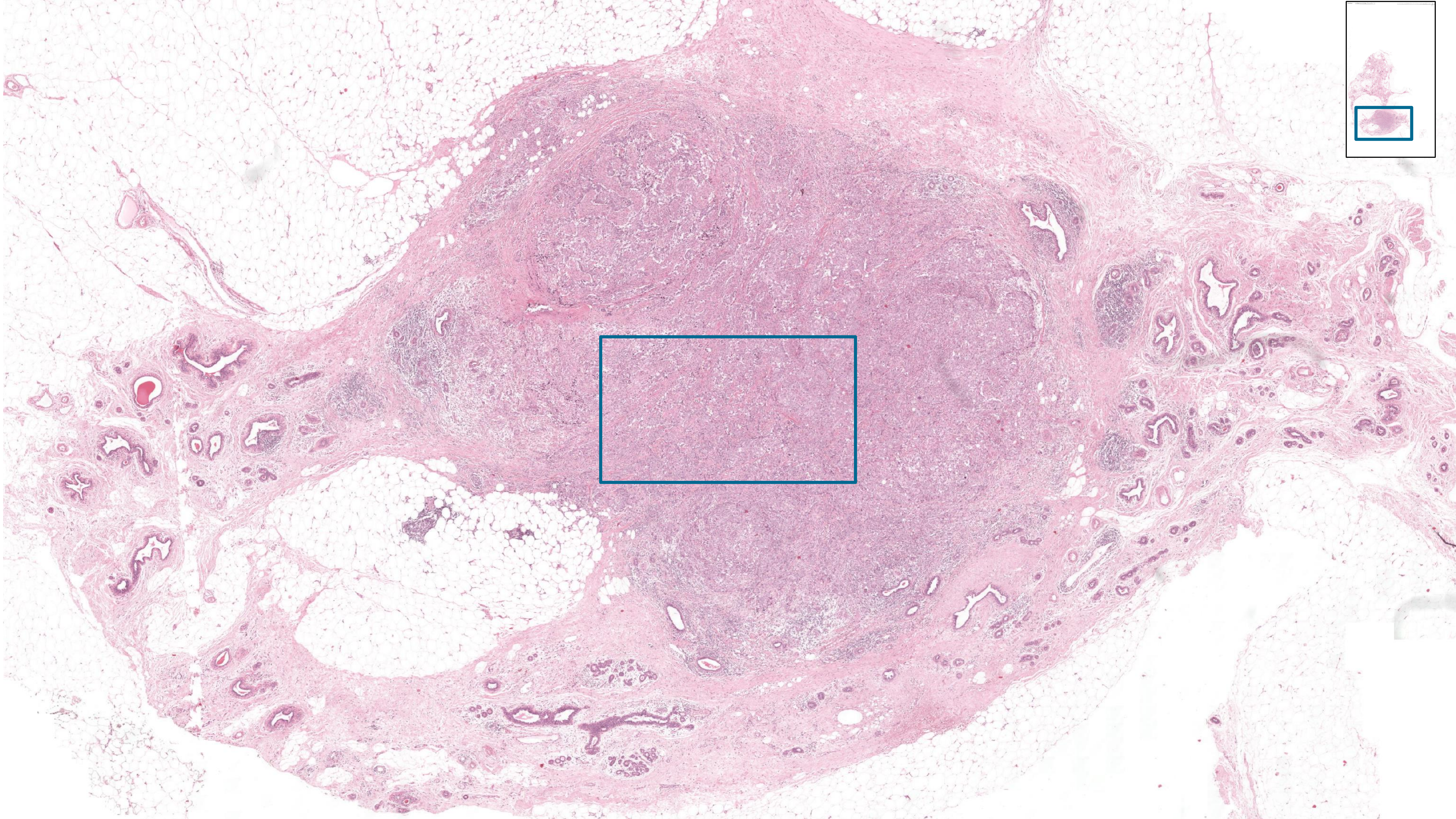
35
mm

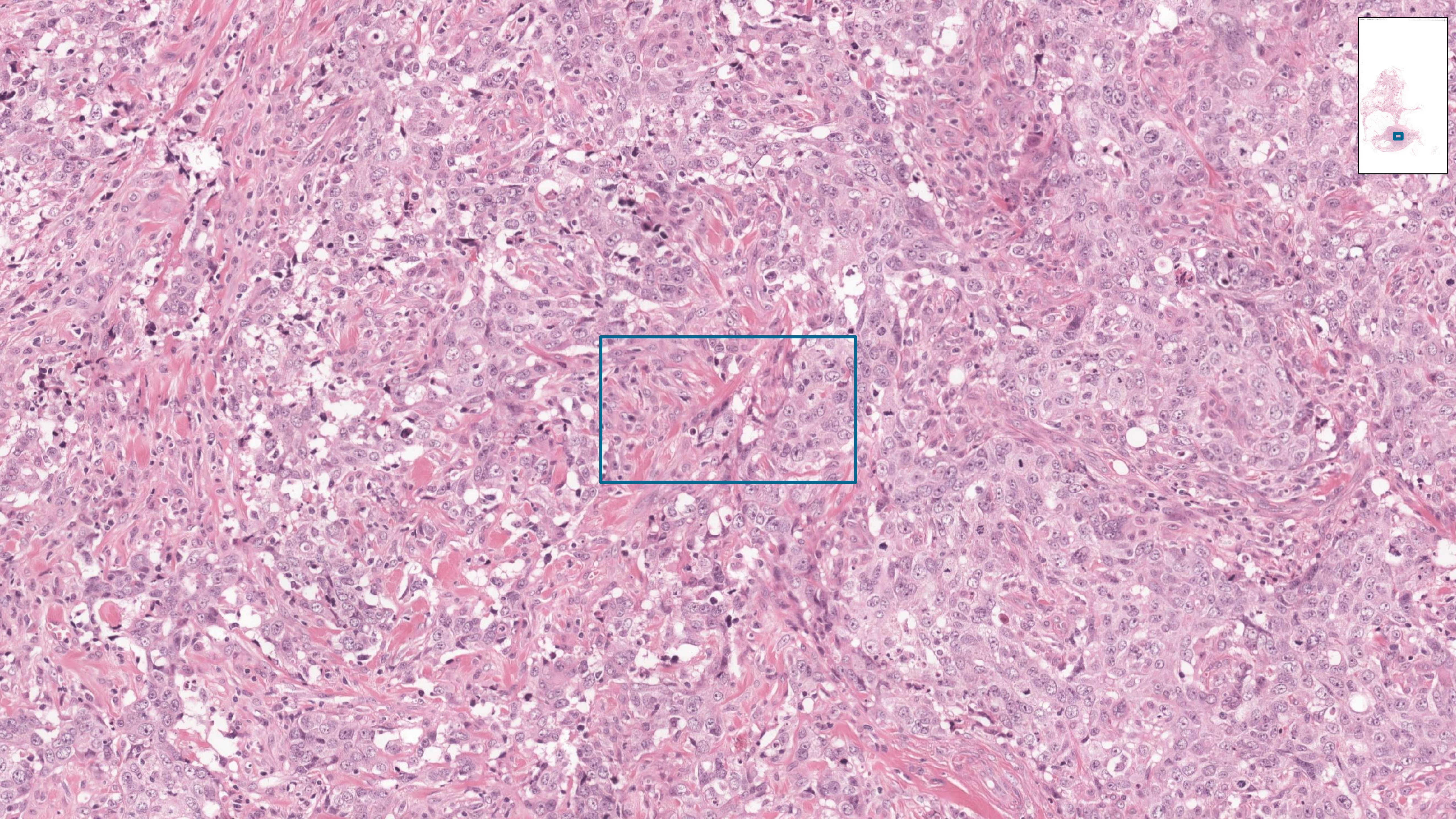


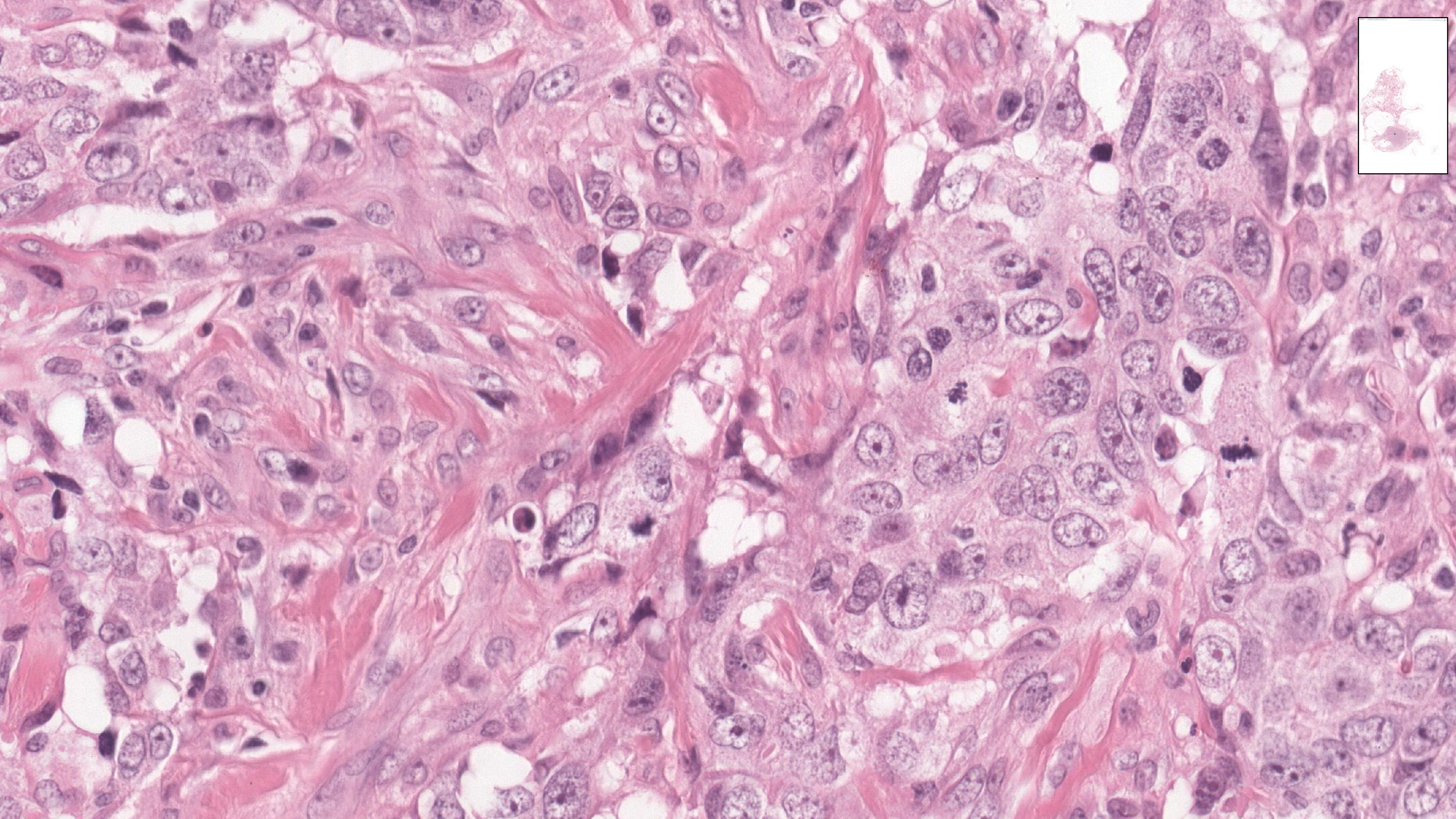
A detection/diagnosis/quantification task involving medical images

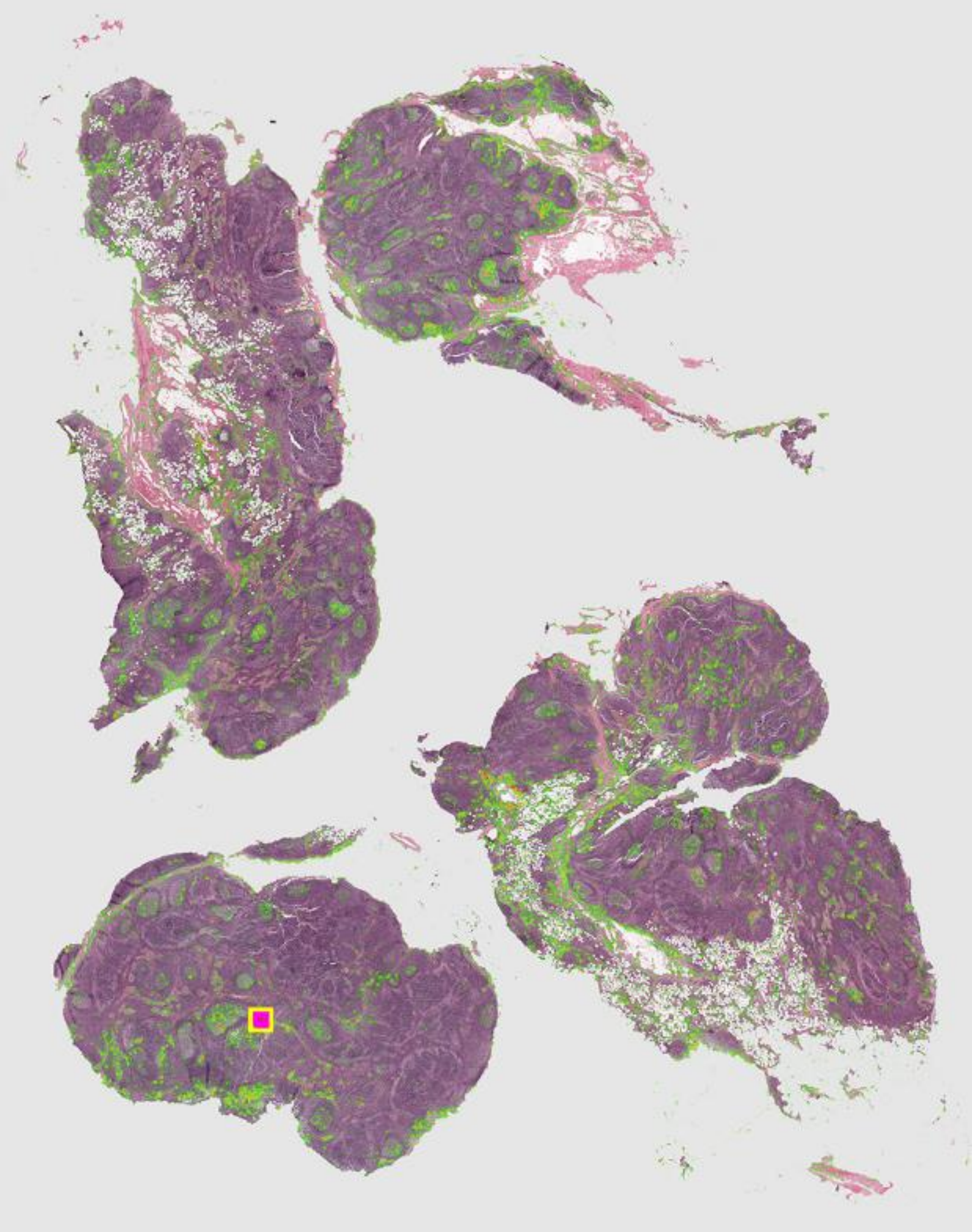


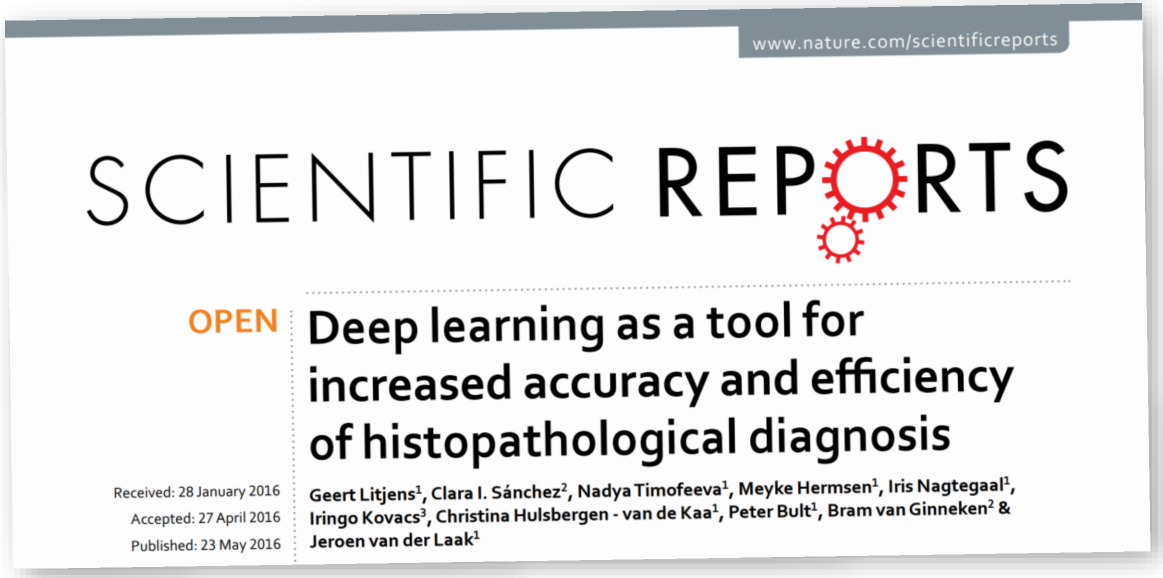












www.nature.com/scientificreports

SCIENTIFIC REPORTS

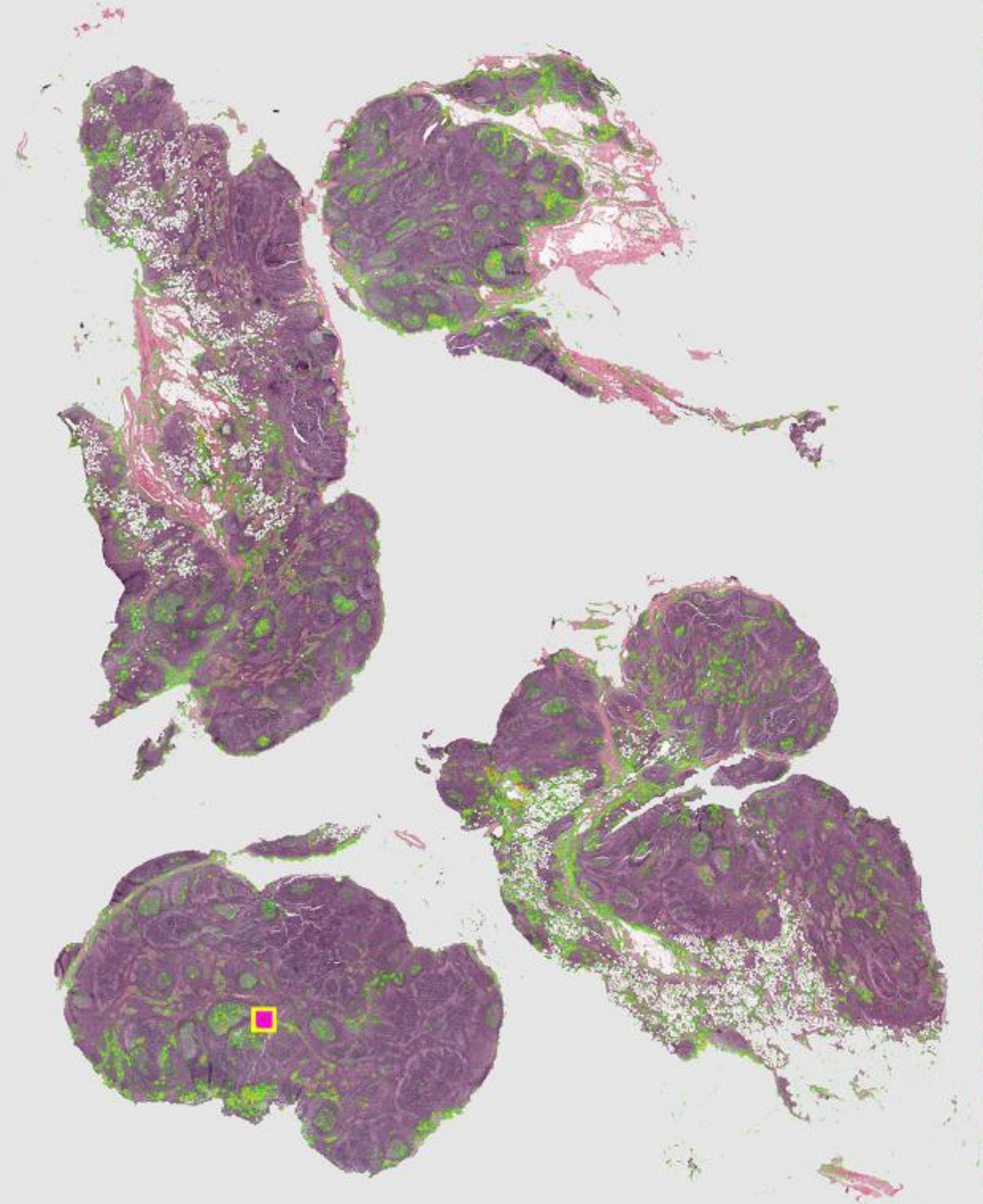
OPEN

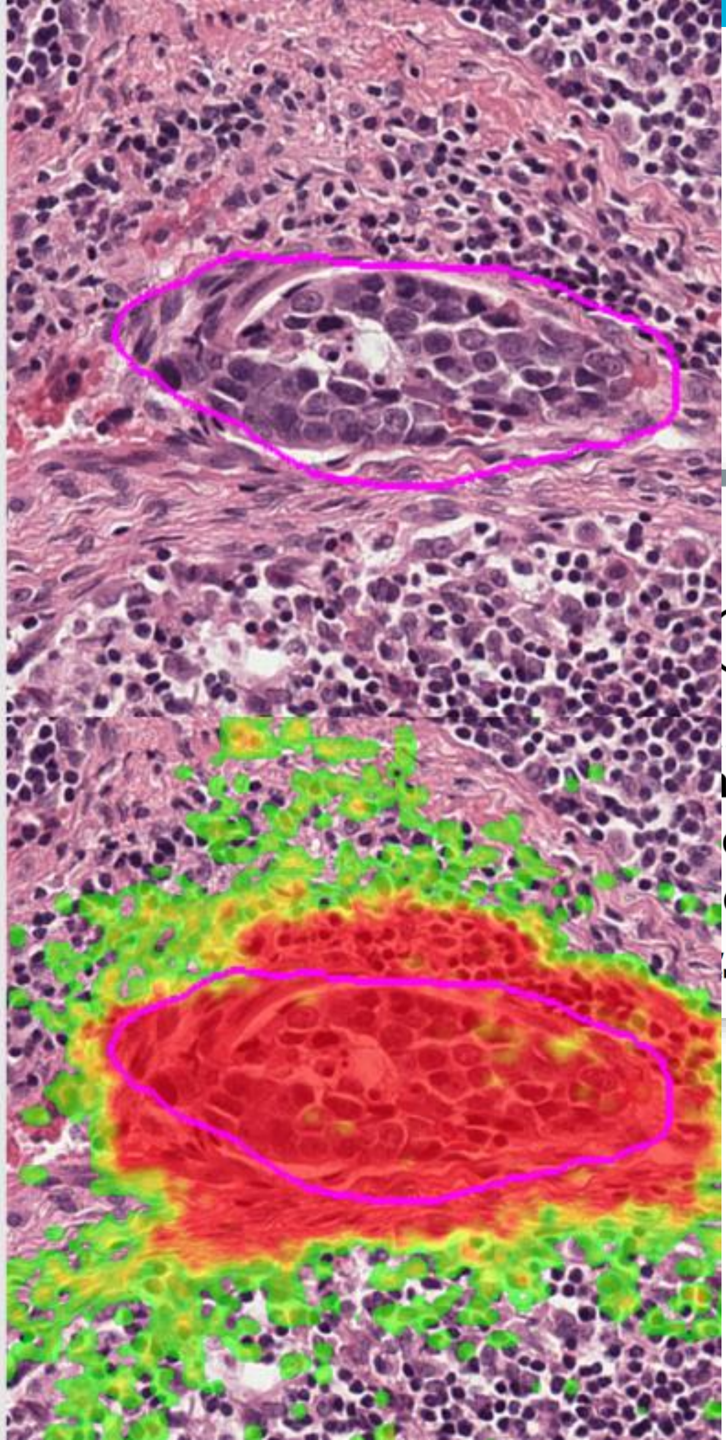
Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis

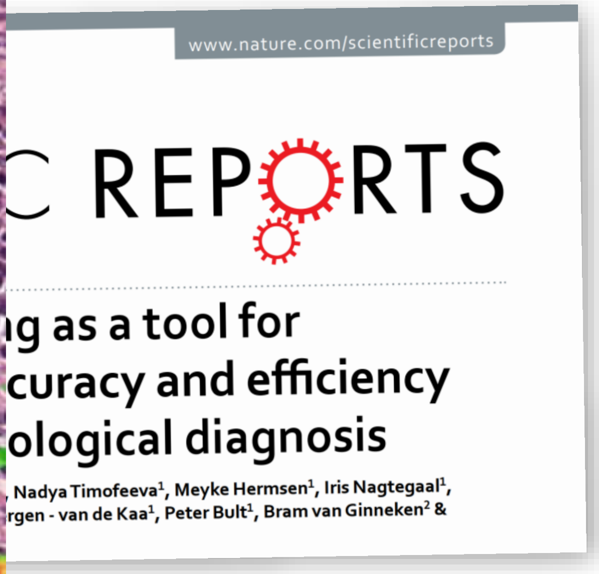
Received: 28 January 2016
Accepted: 27 April 2016
Published: 23 May 2016

Geert Litjens[1], Clara I. Sánchez[2], Nadya Timofeeva[1], Meyke Hermsen[1], Iris Nagtegaal[1], Iringo Kovacs[3], Christina Hulsbergen - van de Kaa[1], Peter Bult[1], Bram van Ginneken[2] & Jeroen van der Laak[1]

Radboudumc





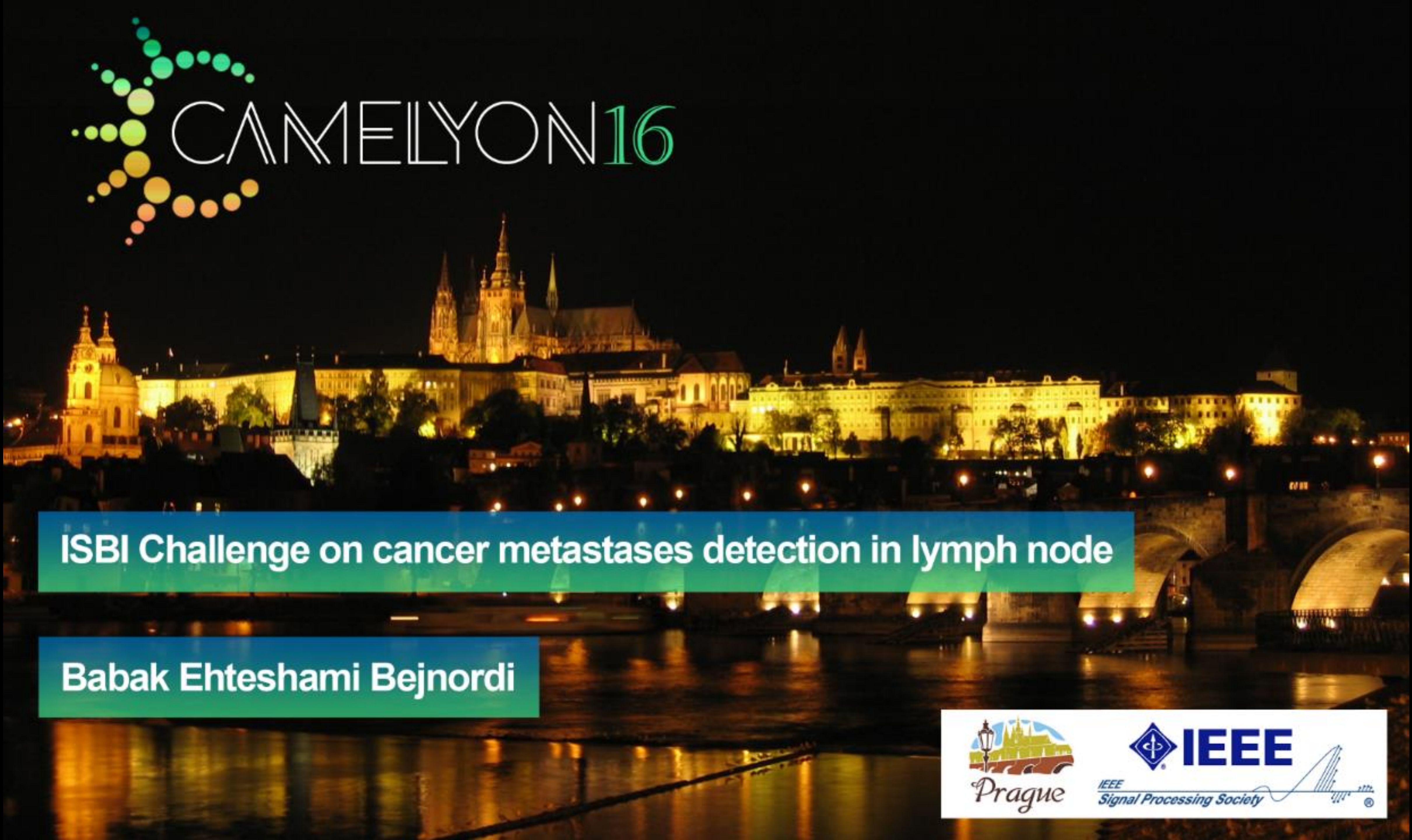

www.nature.com/scientificreports

C REPORTS

ıg as a tool for
curacy and efficiency
ological diagnosis

, Nadya Timofeeva[1], Meyke Hermsen[1], Iris Nagtegaal[1],
rgen - van de Kaa[1], Peter Bult[1], Bram van Ginneken[2] &

Radboudumc


CAMELYON16
ISBI Challenge on cancer metastases detection in lymph node
Babak Ehteshami Bejnordi
Prague
IEEE
IEEE Signal Processing Society

THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN

National Science and Technology Council

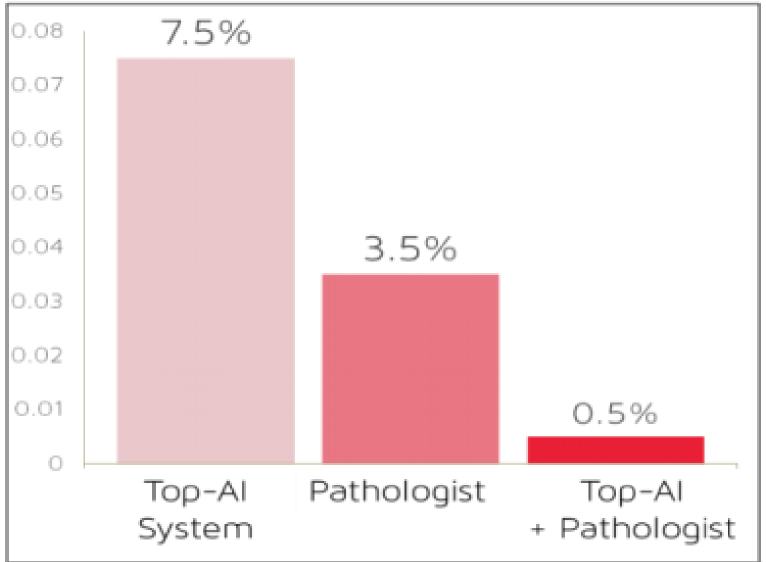
Networking and Information Technology Research and Development Subcommittee

October 2016

National Institutes of Health (NIH) grants-supported research

ARTIFICIAL INTELLIGENCE FOR COMPUTATIONAL PATHOLOGY

Image interpretation plays a central role in the pathologic diagnosis of cancer. Since the late 19th century, the primary tool used by pathologists to make definitive cancer diagnoses is the microscope. Pathologists diagnose cancer by manually examining stained sections of cancer tissues to determine the cancer subtype. Pathologic diagnosis using conventional methods is labor-intensive with poor reproducibility and quality concerns. New approaches use fundamental AI research to build tools to make pathologic analysis more efficient, accurate, and predictive. In the 2016 Camelyon Grand Challenge for metastatic cancer detection,[69] the top-performing entry in the competition was an AI-based computational system that achieved an error rate of 7.5%.[70] A pathologist reviewing the same set of evaluation images achieved an error rate of 3.5%. Combining the predictions of the AI system with the pathologist lowered the error rate to down to 0.5%, representing an 85% reduction in error (see image).[71] This example illustrates how fundamental research in AI can drive the development of high performing computational systems that offer great potential for making pathological diagnoses more efficient and more accurate.



AI significantly reduces pathologist error rate in the identification of metastatic breast cancer from sentinel lymph node biopsies.

Radboudumc

JAMA | Original Investigation

Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer

Babak Ehteshami Bejnordi, MS; Mitko Veta, PhD; Paul Johannes van Diest, MD, PhD; Bram van Ginneken, PhD; Nico Karssemeijer, PhD; Geert Litjens, PhD; Jeroen A. W. M. van der Laak, PhD; and the CAMELYON16 Consortium

IMPORTANCE Application of deep learning algorithms to whole-slide pathology images can potentially improve diagnostic accuracy and efficiency.

OBJECTIVE Assess the performance of automated deep learning algorithms at detecting metastases in hematoxylin and eosin–stained tissue sections of lymph nodes of women with breast cancer and compare it with pathologists' diagnoses in a diagnostic setting.

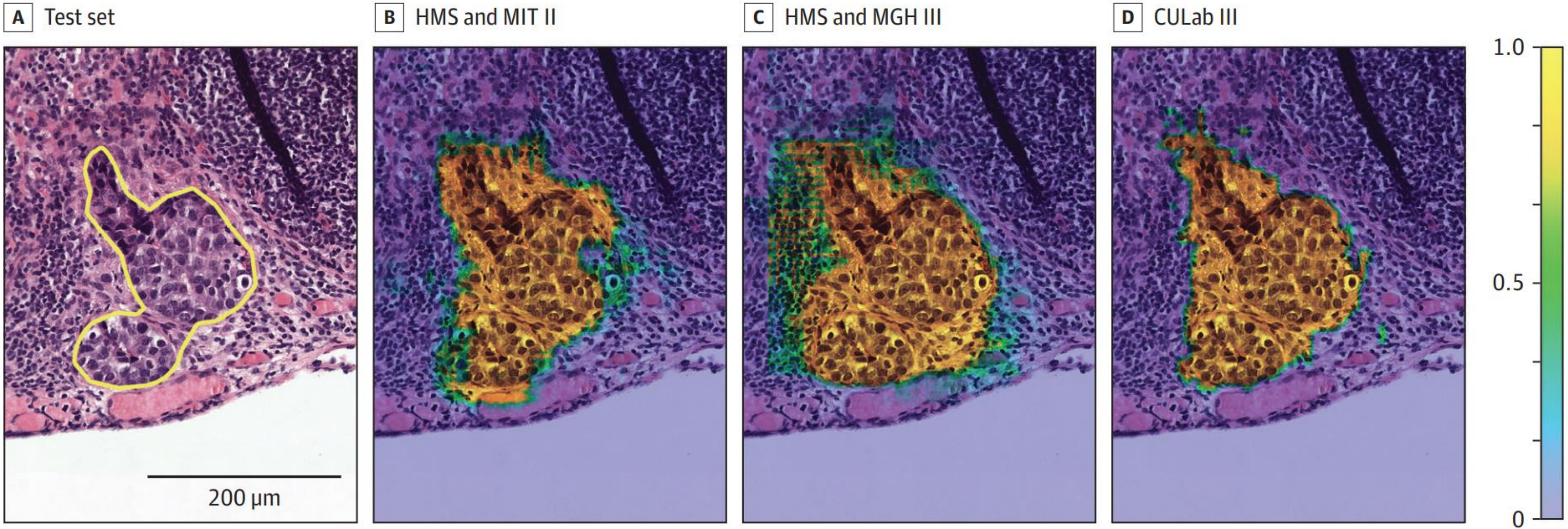
Editorial page 2184

Related articles page 2211 and page 2250

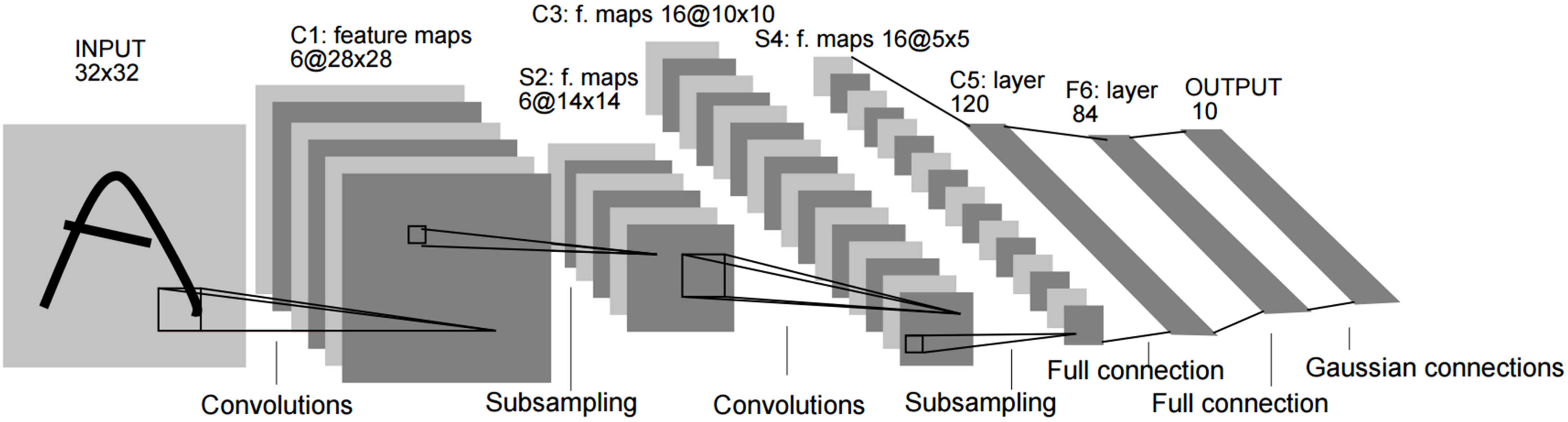
Supplemental content

CME Quiz at jamanetwork.com/learning and **CME Questions** page 2252

CONCLUSIONS AND RELEVANCE In the setting of a challenge competition, some deep learning algorithms achieved better diagnostic performance than a panel of 11 pathologists participating in a simulation exercise designed to mimic routine pathology workflow; algorithm performance was comparable with an expert pathologist interpreting whole-slide images without time constraints. Whether this approach has clinical utility will require evaluation in a clinical setting.



LeNet, Yann LeCun, 1998



- Secret: end-to-end training, all weights are trained with backpropagation
- This networks contains 60,000 weights

AlexNet, Alex Krizhevsky, December 2012, NIPS

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

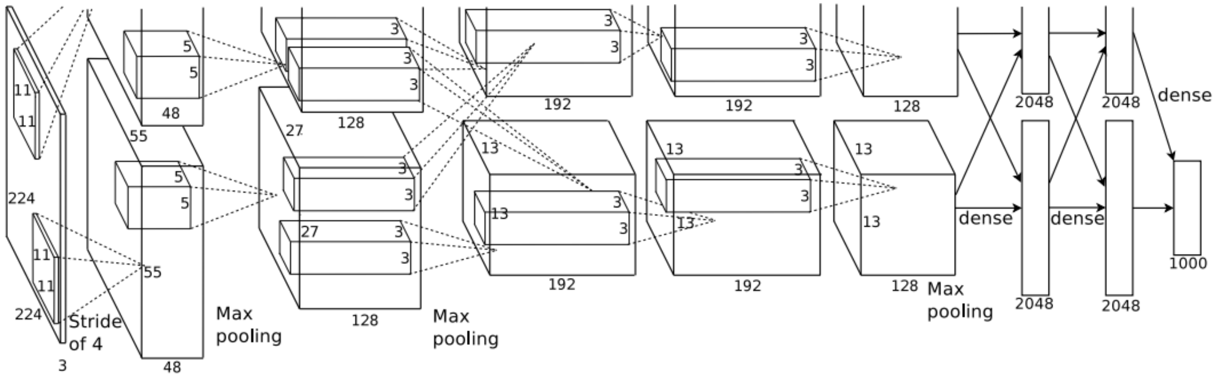


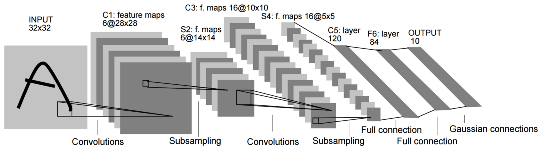
Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input 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.

- Souped-up version of LeNet
- Input ~100 times larger
- 1000 output classes instead 10
- 60M weights instead of 60k

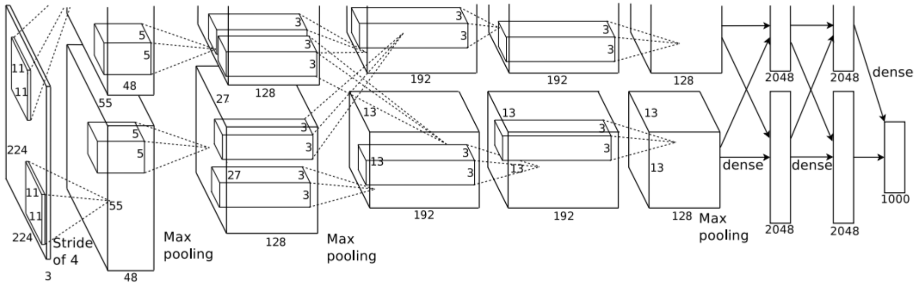
- Novelty: trained on 2 GPUs

Input: image. Output: number(s). Classifier/discriminator

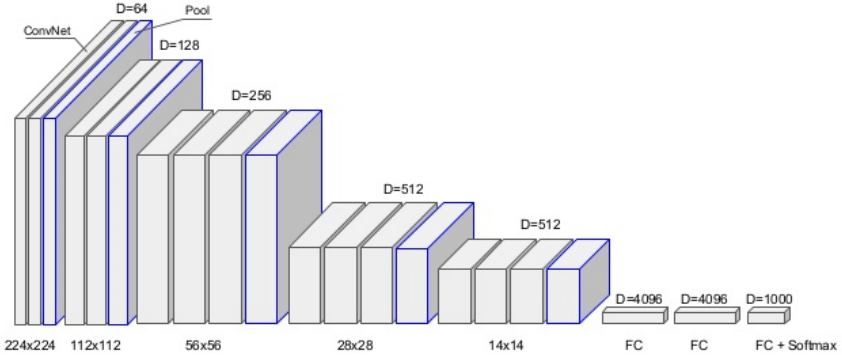
LeNet (1998)



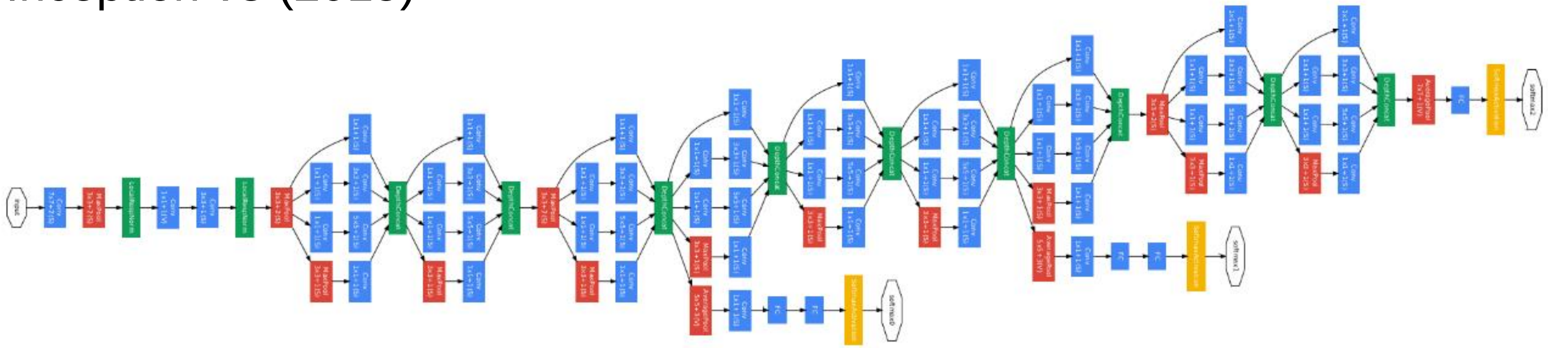
AlexNet (2012)



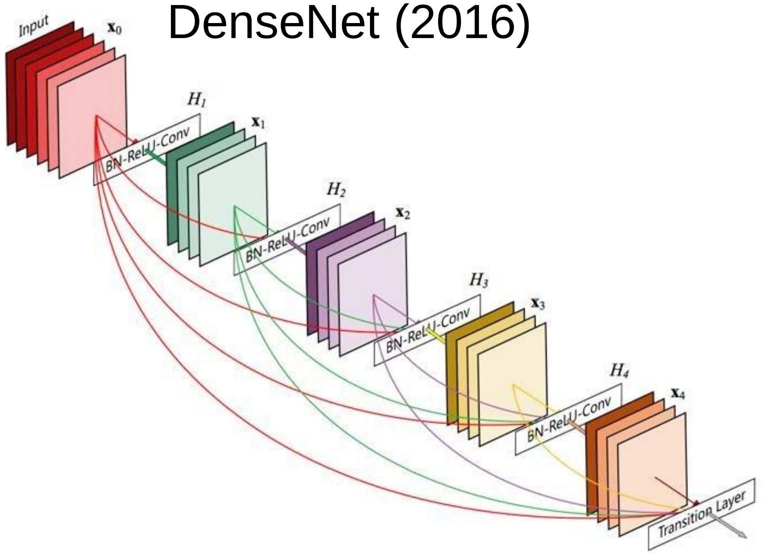
Classical CNN topology - VGGNet (2013)



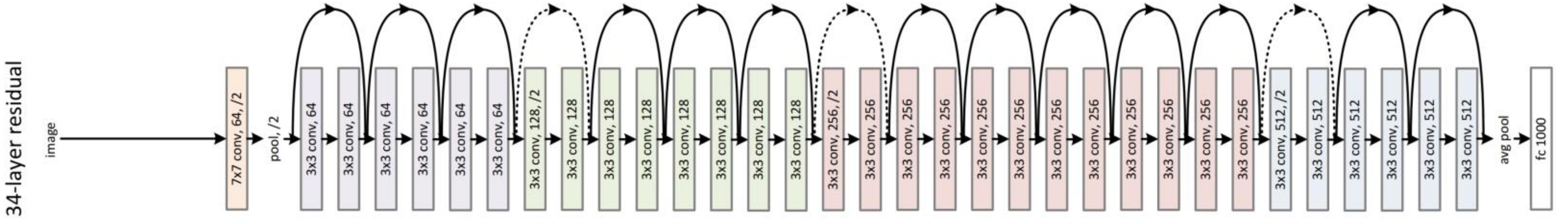
Inception v3 (2015)



DenseNet (2016)

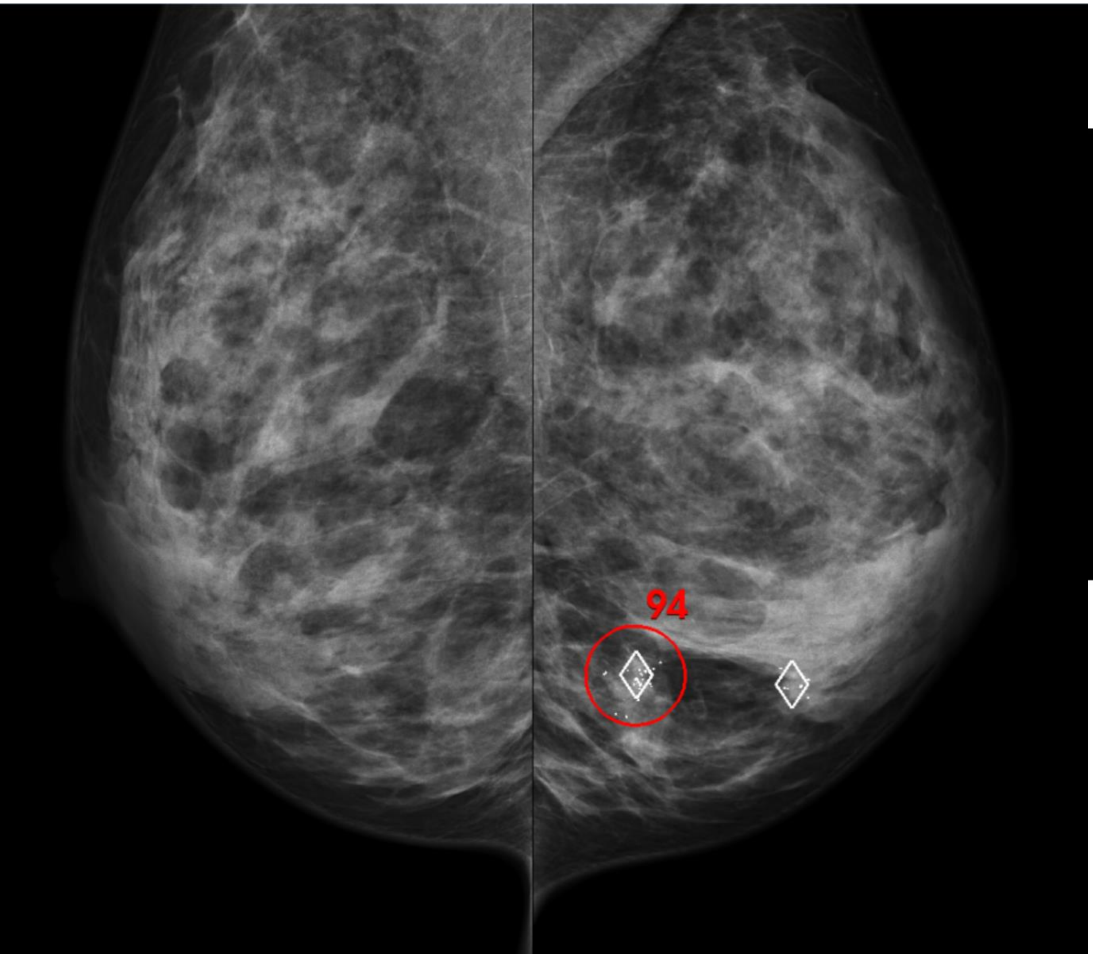


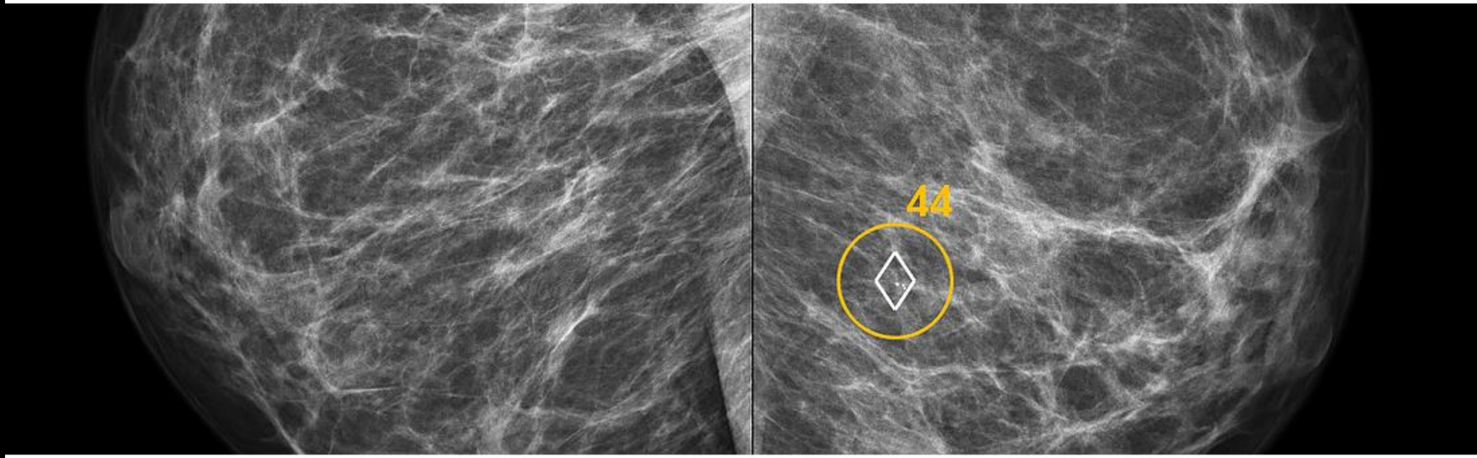
ResNet (2015)



Radboudumc

Transpara





OXFORD

JNCI J Natl Cancer Inst (2019) 111(9): djy222

doi: 10.1093/jnci/djy222
Article

ARTICLE

Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists

Alejandro Rodriguez-Ruiz, Kristina Lång, Albert Gubern-Merida, Mireille Broeders, Gisella Gennaro, Paola Clauser, Thomas H. Helbich, Margarita Chevalier, Tao Tan, Thomas Mertelmeier, Matthew G. Wallis, Ingvar Andersson, Sophia Zackrisson, Ritse M. Mann, Ioannis Sechopoulos

See the Notes section for the author's affiliations.
Correspondence to: Ioannis Sechopoulos, PhD, Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Post 766, Nijmegen, the Netherlands (e-mail: Ioannis.sechopoulos@radboudumc.nl).

Radboudumc

Methods: Nine multi-reader, multi-case study datasets previously used for different research purposes in seven countries were collected. Each dataset consisted of DM exams acquired with systems from four different vendors, multiple radiologists' assessments per exam, and ground truth verified by histopathological analysis or follow-up, yielding a total of 2652 exams (653 malignant) and interpretations by 101 radiologists (28 296 independent interpretations). An AI system analyzed these exams yielding a level of suspicion of cancer present between 1 and 10. The detection performance between the radiologists and the AI system was compared using a noninferiority null hypothesis at a margin of 0.05.

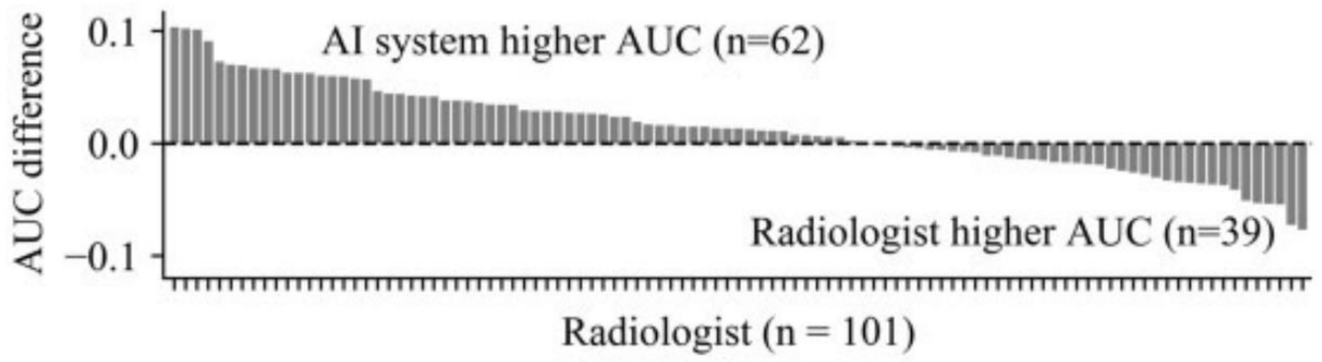


Figure 2. Differences in area under the receiver operating characteristic curve (AUC) between the artificial intelligence (AI) system and each radiologist.

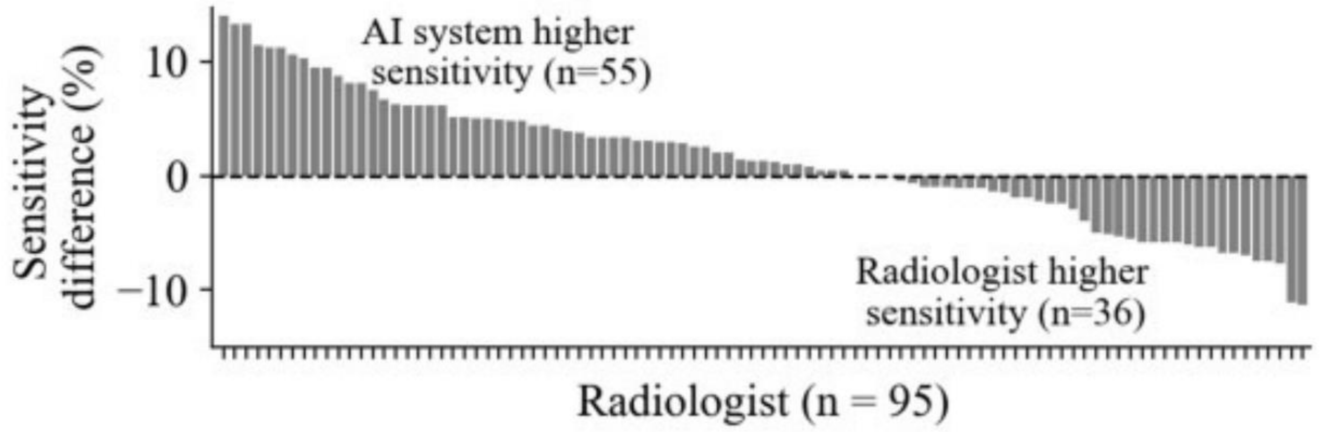
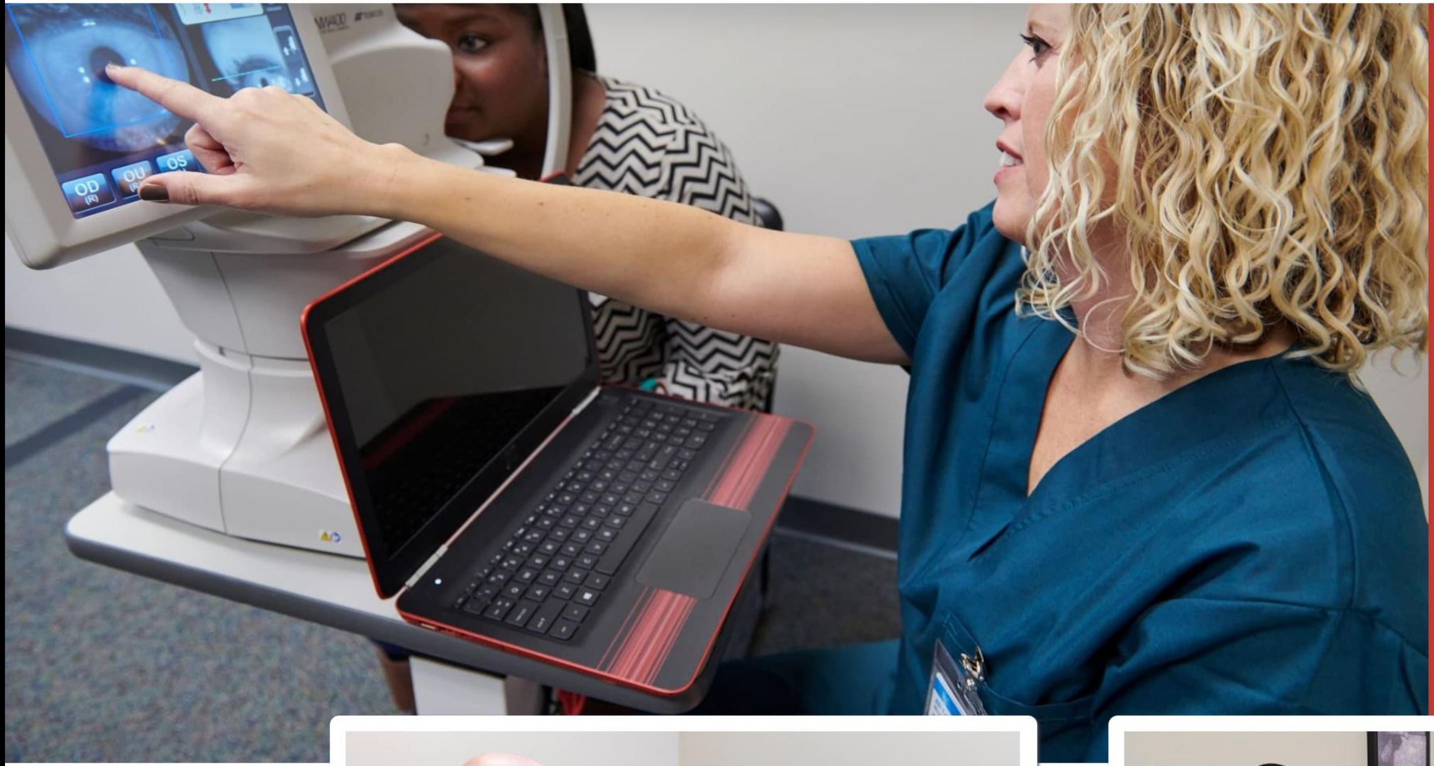


Figure 3. Differences (%) in sensitivity between the artificial intelligence (AI) system and each radiologist at the specificity of each radiologist considering BI-RADS three and over as positive recall. BI-RADS = Breast Imaging Reporting and Data System.



Autonomous AI that instantly detects disease

Increase patient access to early disease detection.

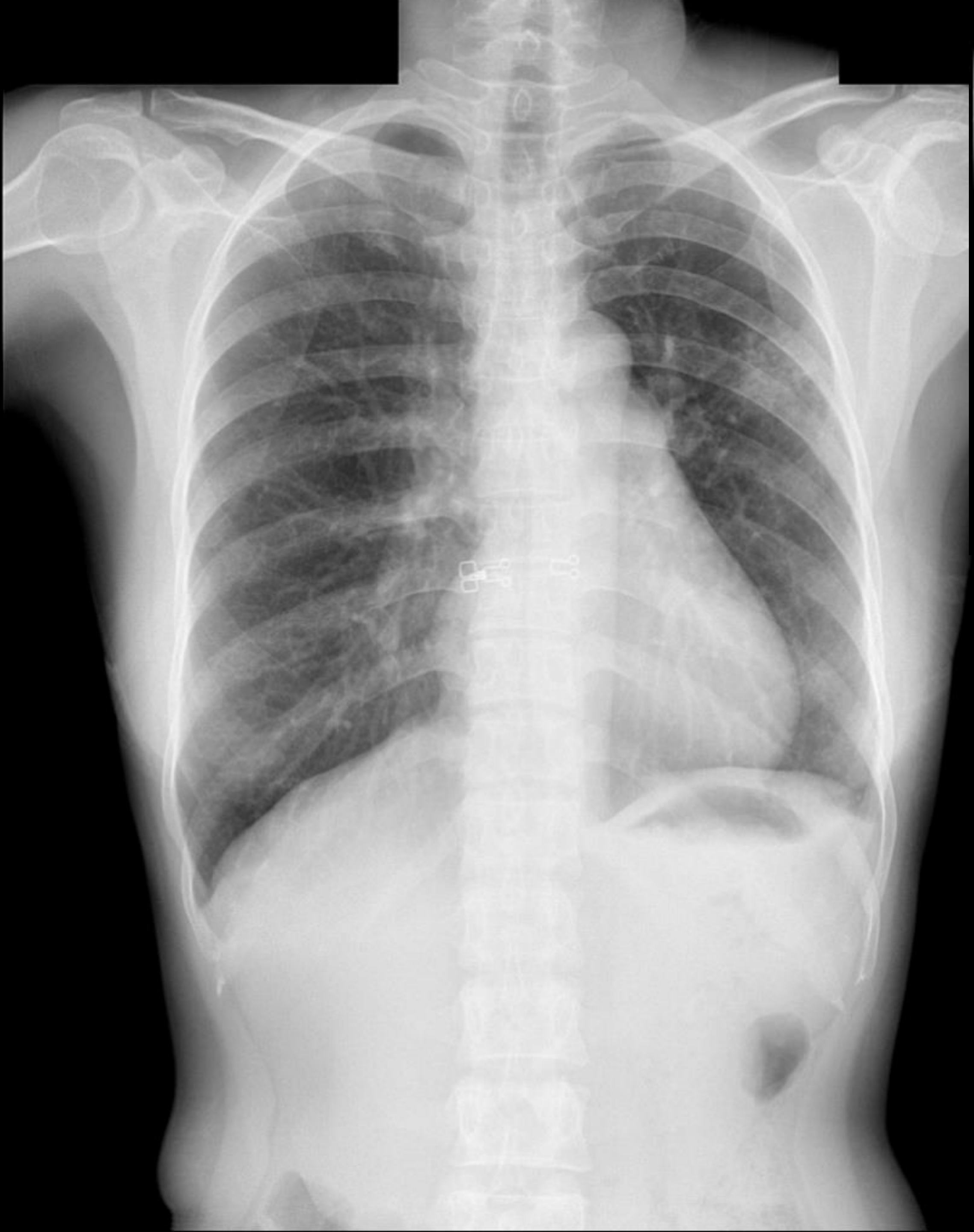
Learn More

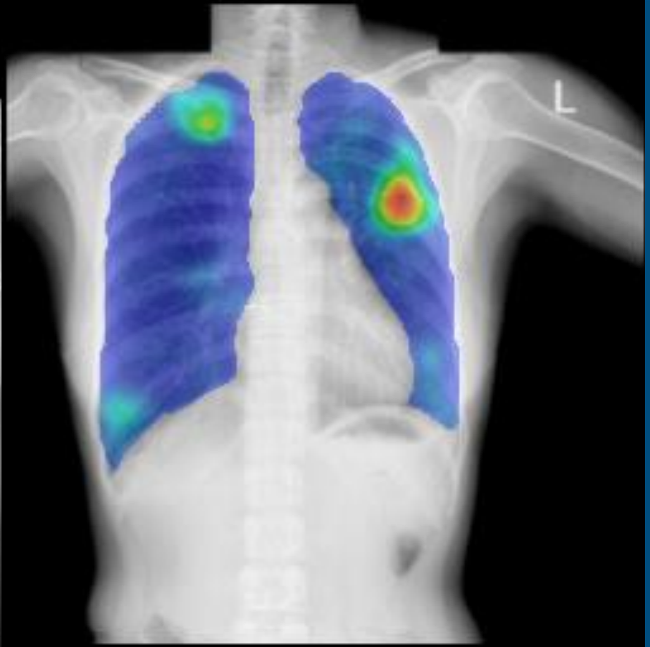
IDx-DR

The first ever autonomous AI system cleared by the FDA to provide a diagnostic decision

How to test for DR in primary care

Download our white paper on the pros and cons of telmedicine and AI solutions





CAD4TB

Artificial intelligence for the detection of tuberculosis

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


مفت
TB

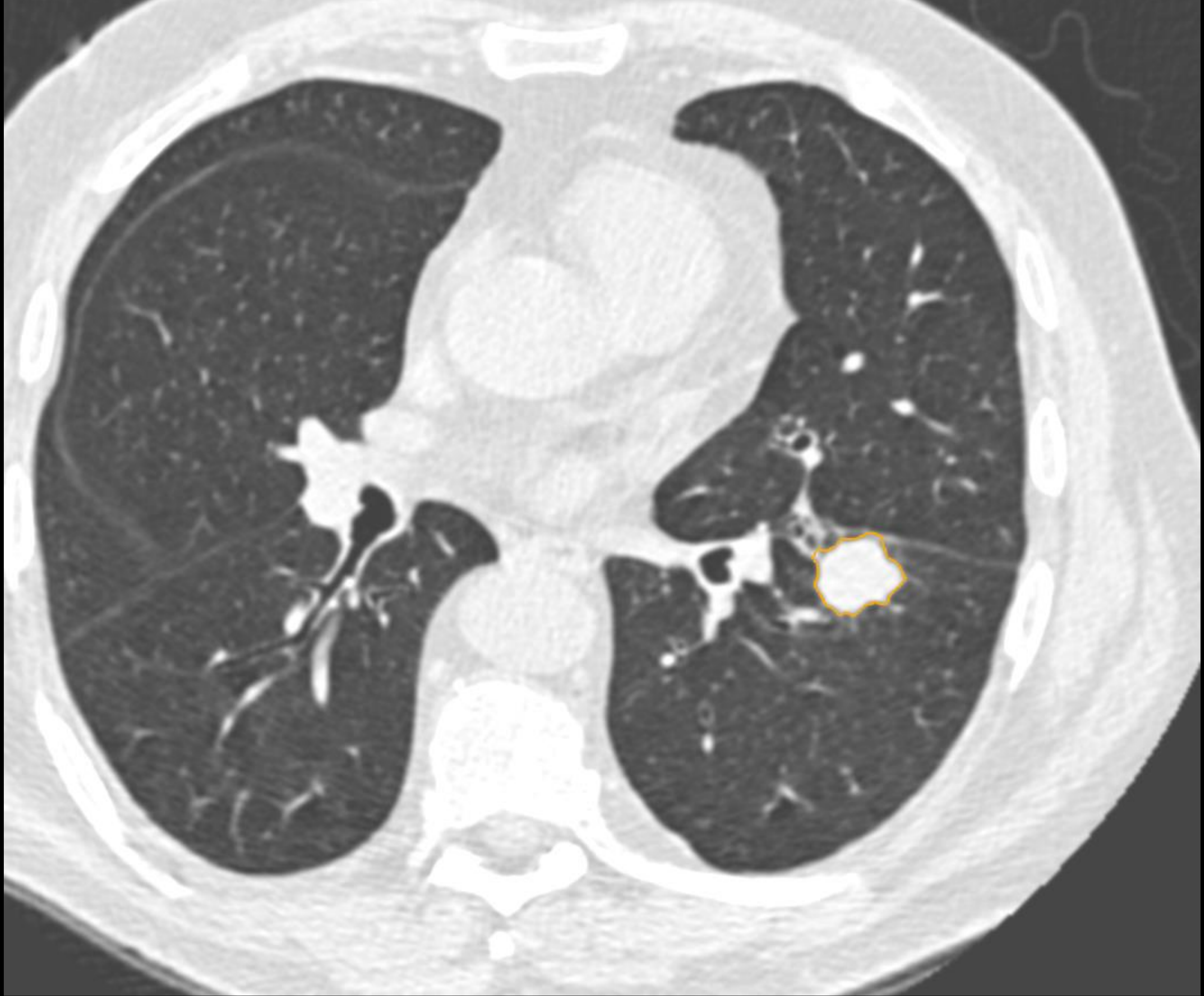

TB

Screening with CT

- Hold your breath, slide through scanner in <10s
- High resolution (<1mm) 3D image of the lungs
- Very low dose possible
- No contrast material needed, very user friendly
- No high-end scanner needed: cheap (~200 k$)

Baseline

After 1 year
6mm nodule



After 2 years
2cm nodule

Squamous cell lung cancer

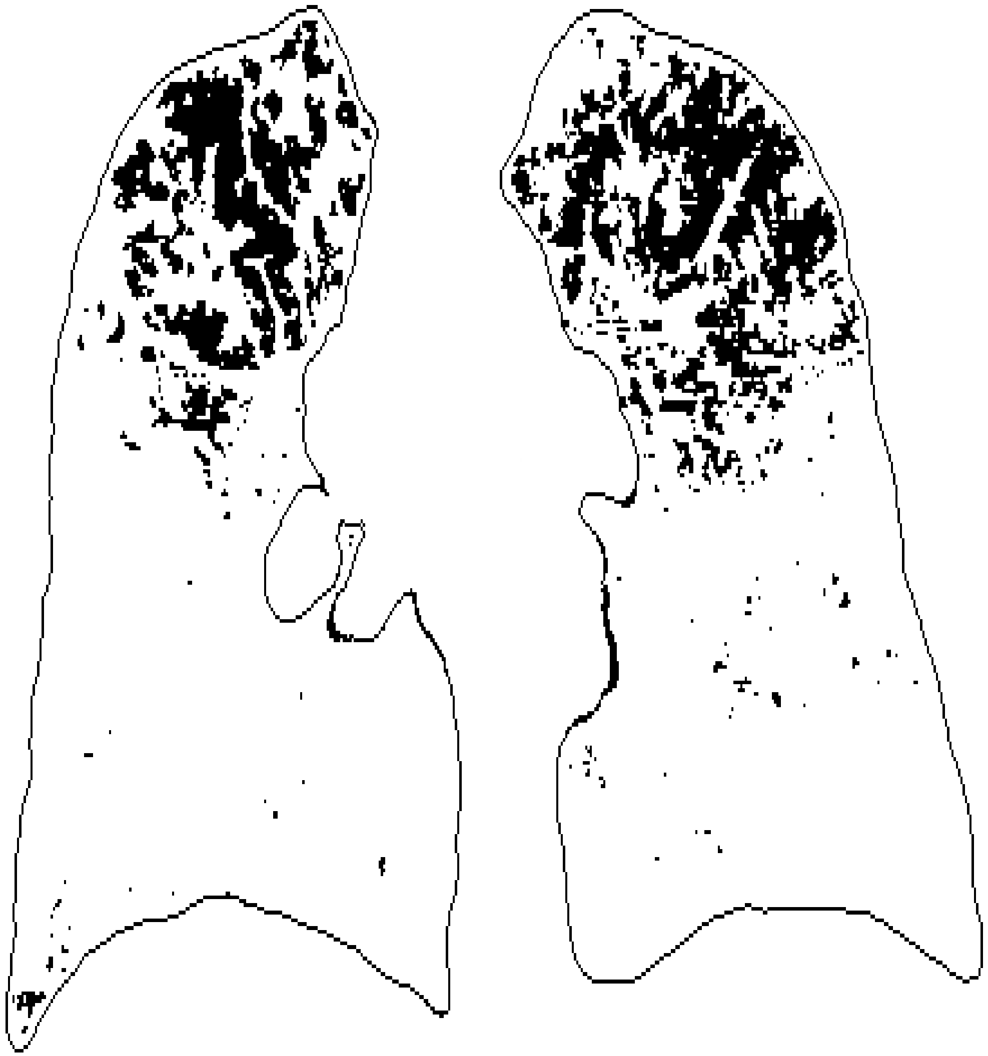
Still early stage!

Artificial Intelligence in medicine development - without the hot air

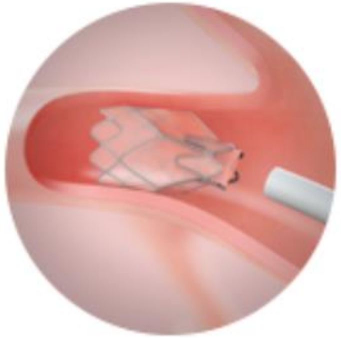
- Developments in Artificial Intelligence (AI) are (…) drivers for a fourth industrial revolution, which will impact all human activities, including medicine development and health care.

- There are obvious benefits for patients
  - More data about the patient that can be assessed
  - Freeing up time for health care professionals for human interaction
  - Including the full body of medical knowledge to identify the optimal treatment for a patient

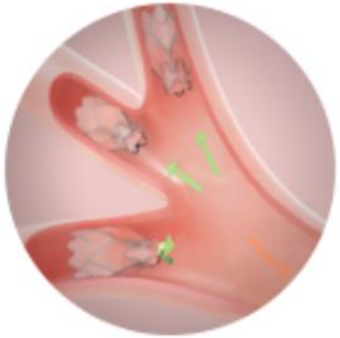
- Furthermore, AI creates new opportunities in clinical development, such as defining inclusion/exclusion criteria, the assessment of clinical efficacy of new drugs and evaluation of (new) clinical endpoints using innovative data points



StratX



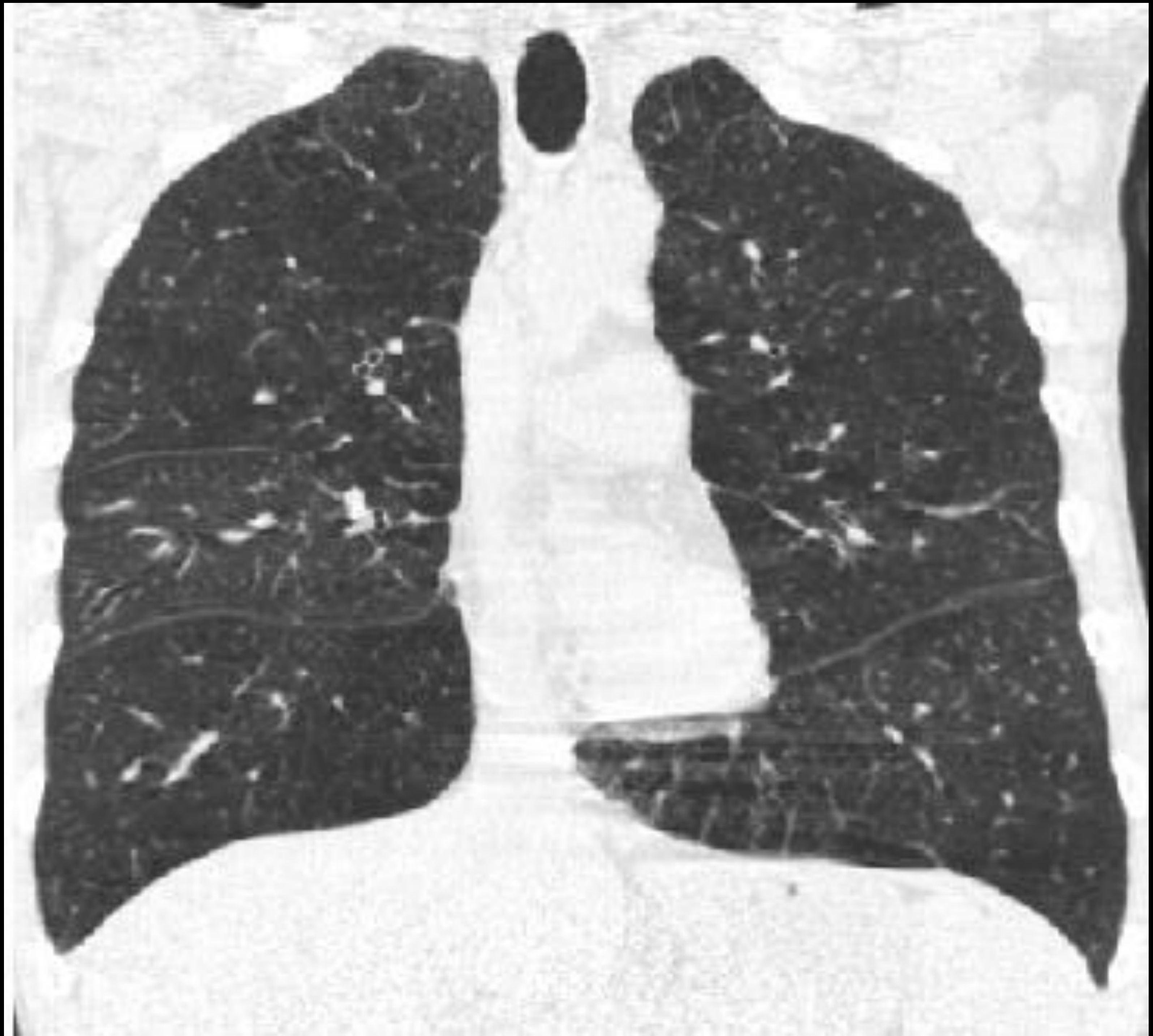
Using a standard bronchoscope, Zephyr valves are delivered to target airways using a flexible delivery catheter

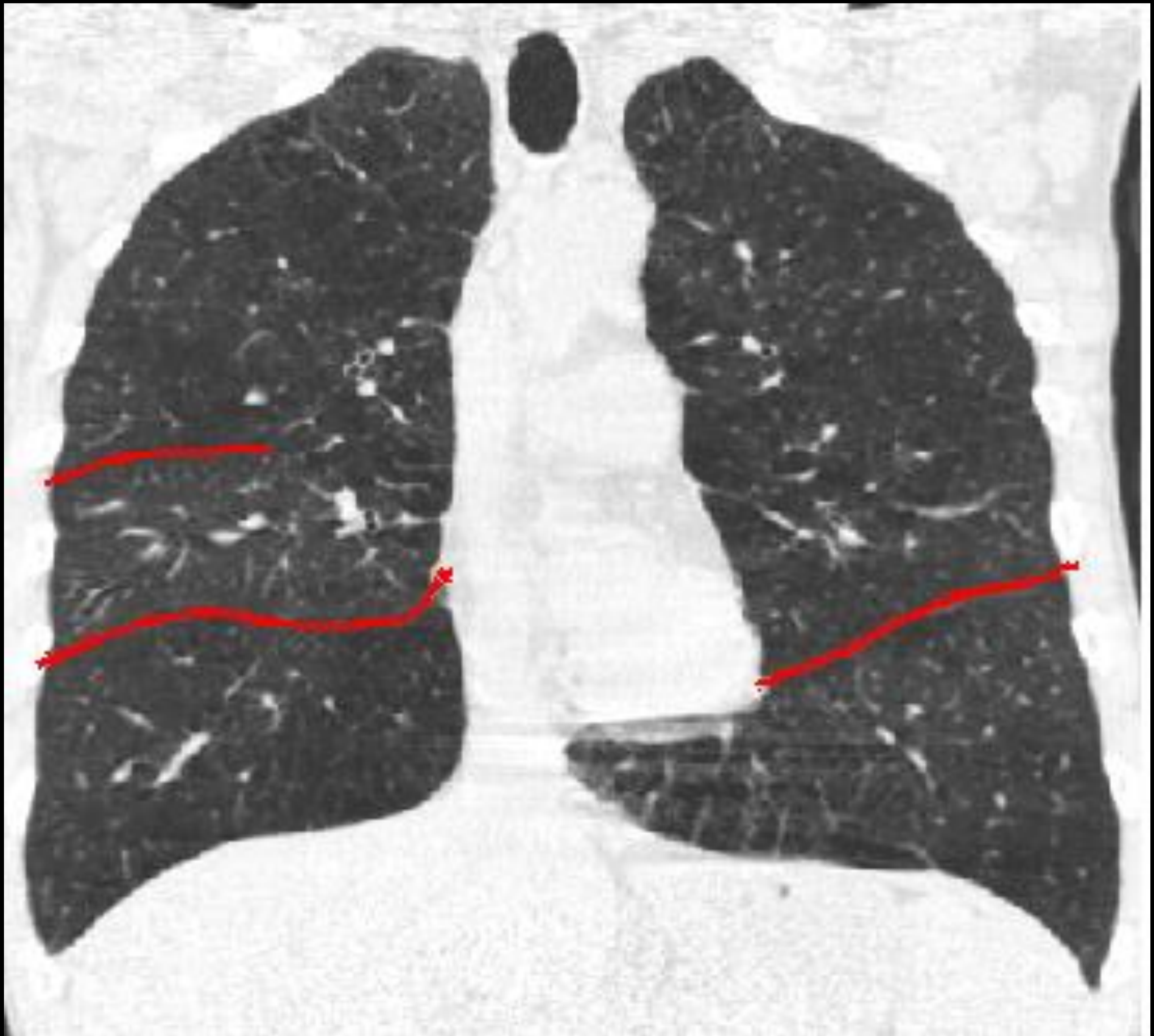


Once implanted, the one-way valve prevents airflow into the diseased region, while allowing trapped air and fluids to escape

Reducing the volume of the diseased region may allow healthier regions to expand and function more efficiently


Right upper lobe
Left upper lobe
Right middle lobe
Right lower lobe
Left lower lobe





StratX

thirona

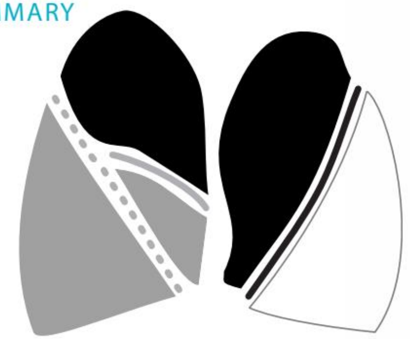
StratX™ Lung Report

pulmonX

Patient ID 35-127-AJJ
Scan ID 0026671
CT Scan Date 16 February 2011

Upload Date 27 July 2016
Report Date 27 July 2016
Scan Comments

SUMMARY



KEY

- ≥70% Voxel Density Less Than -910 HU
- 60-70% Voxel Density Less Than -910 HU
- 50-60% Voxel Density Less Than -910 HU
- <50% Voxel Density Less Than -910 HU
- >95% Fissure Completeness
- 80-95% Fissure Completeness
- <80% Fissure Completeness

RESULTS

	RIGHT LUNG				LEFT LUNG	
	RUL	RUL+RML	RML	RLL	LUL	LLL
% Fissure Completeness	64.2	65.0	84.9	65.0	100	100
% Voxel Density Less Than -910 HU	84	79	63	66	83	43
% Voxel Density Less Than -950 HU	67	59	36	40	69	16
Inspiratory Volume(ml)	1120	1492	372	1901	1717	1520

CE XXXXX. Powered by Thirona LungQ v1.0.0

SUMMARY

Illustration summarizes key information

RESULTS

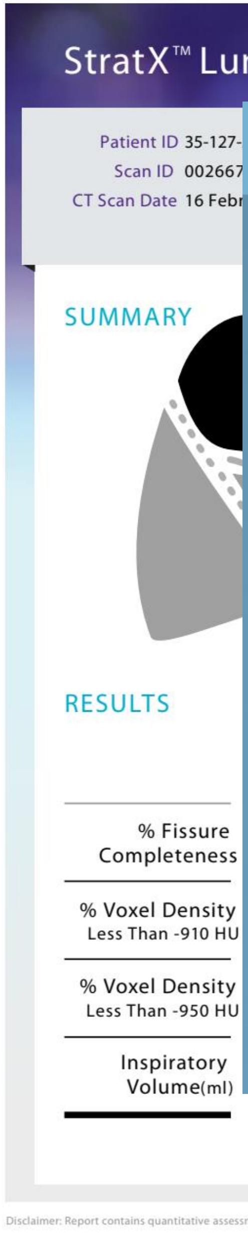
Table lists validated measurements by lobe:

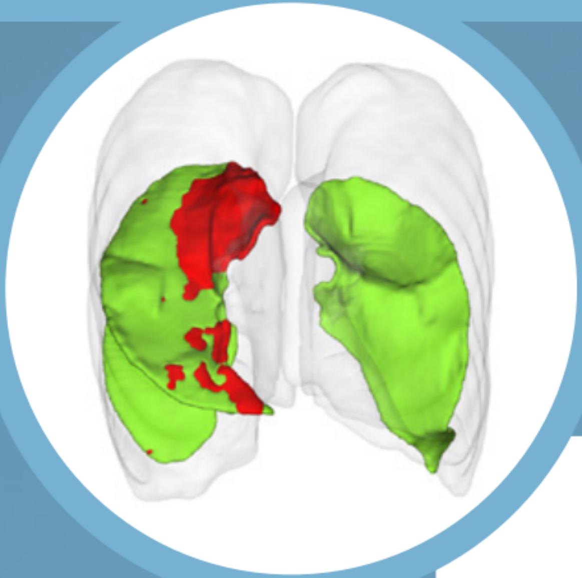
- Fissure completeness
- Emphysema density (based on voxel density less than -910 HU)
- Inspiratory volume

Radboudumc

StratX

thirona





The StratX analysis quantifies the completeness of each fissure using an algorithm that has been validated in a retrospective study of over 200 EBV patients, the largest such analysis performed to date.[2] Fissure completeness is a proven predictor for volume reduction resulting from EBV therapy.[3]

	RIGHT LUNG			
	RUL	RUL+RML	RML	RLL
% Fissure Completeness	64.2	65.0	84.9	65.0

Radboudumc

StratX

thirona

WORKFLOW WITH RAPID TURNAROUND TIME

Capture CT Scan

Capture a high resolution chest CT scan according to the StratX CT parameters.

Upload CT Scan

Use web browser to upload CT scan to the secure, cloud-based StratX platform.

- Automatically anonymized data with no patient health information transfer
- Secure 256 bit SSL socket level encryption

Analyze Data + Generate Report

Data is analyzed by validated algorithms and the StratX report is uploaded to the StratX platform within 2-3 working days.

Review Report

Access pulmonxstratx.com to review the report in a .pdf (2D) or .html (3D) format from any clinical setting.

Confidently Determine Treatment Options

Determine the most suitable treatment option for your patient using the quantitative StratX information and clinical judgment.

In summary

- AI and deep learning have been successfully applied to many medical image analysis tasks
  - Lots of data and good engineering → Human level performance
  - Possibility for automation and saving physician's time
- For tasks for which clinical guidelines do not exist, such as predicting response to treatment results are far less convincing
- The key question is not if AI can get the signal out of the data, but if the signal is in (all) the data