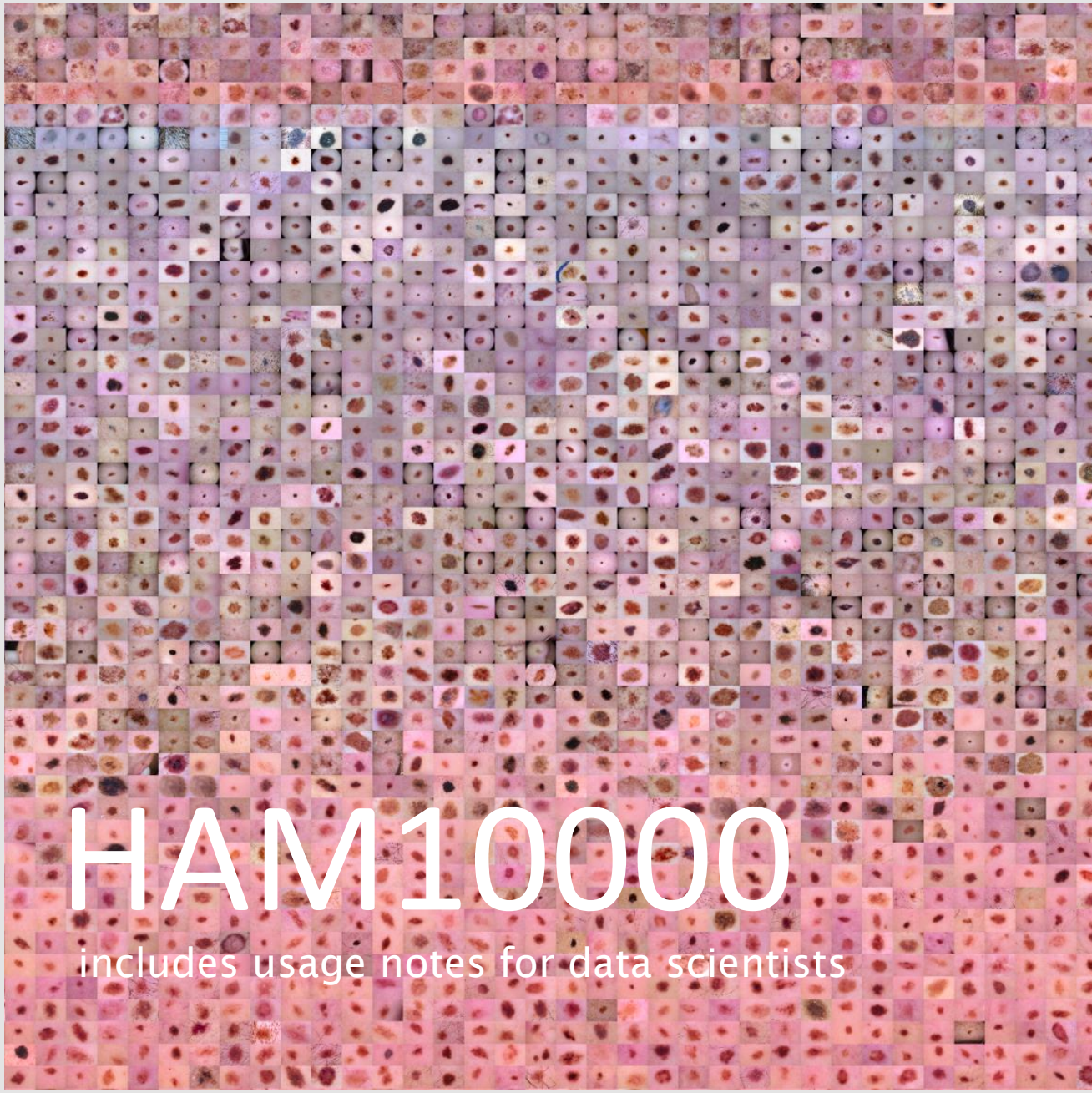
The importance of domain-specific knowledge

Harald Kittler
Vienna Dermatologic Imaging Research (ViDIR)
Department of Dermatology
Medical University of Vienna, Austria

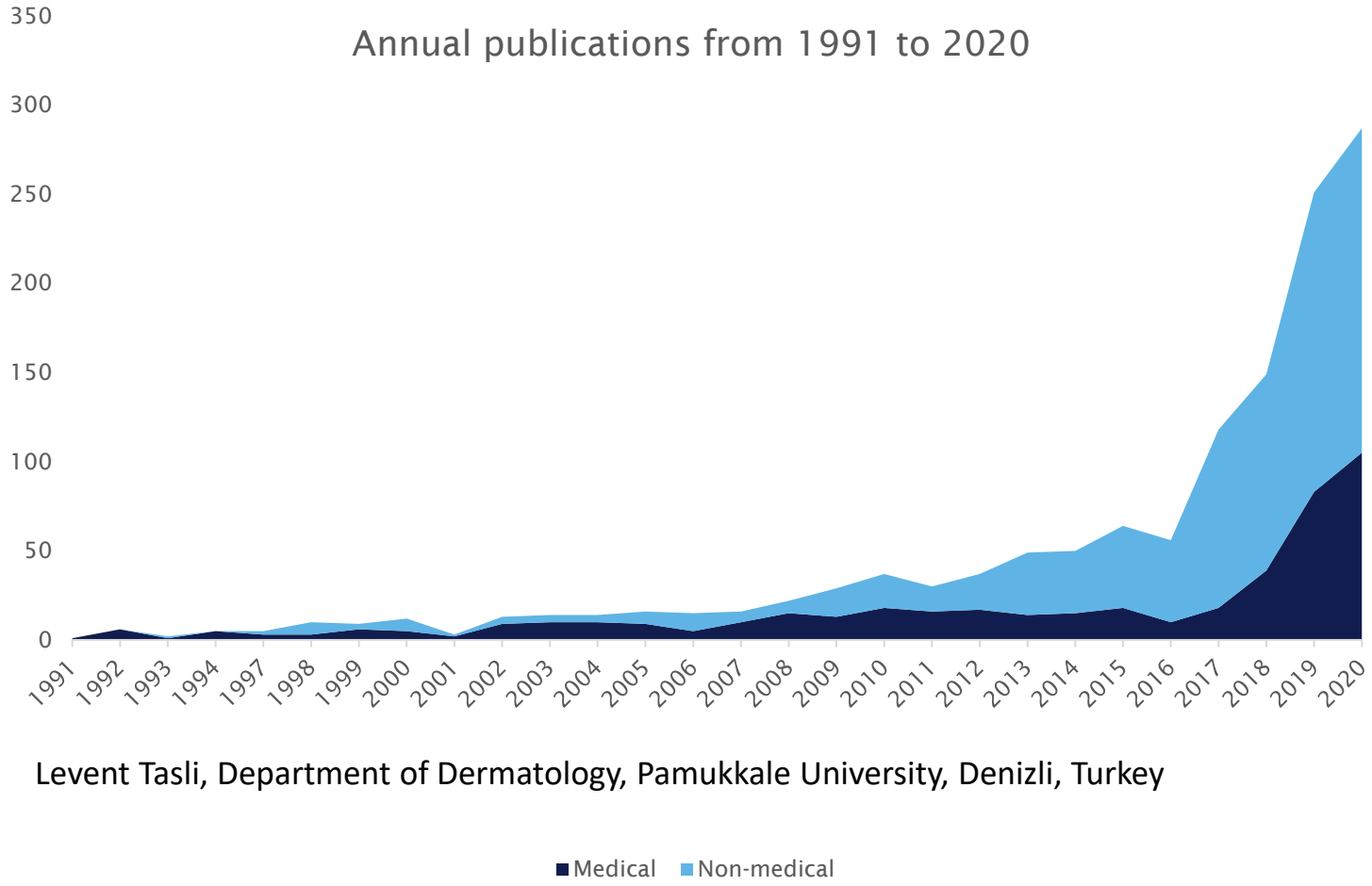
Who am I? Who are we?

ViDIR - Vienna Dermatologic Imaging Research Group

Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5,** 180161 (2018). https://doi.org/10.1038/sdata.2018.161



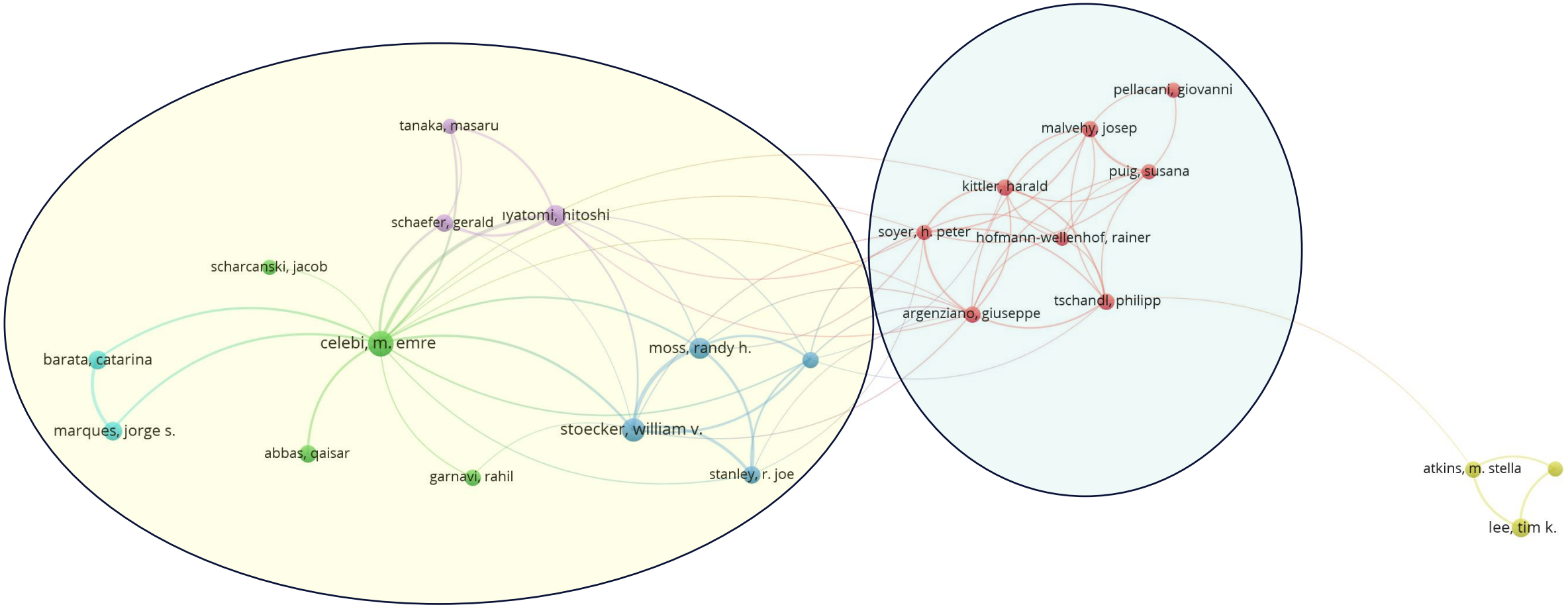
Skin cancer diagnosis: The importance of domain-specific knowledge



Levent Tasli, Department of Dermatology, Pamukkale University, Denizli, Turkey

Co-authorship network of a subset of the most productive authors in the field

Levent Tasli, Department of Dermatology, Pamukkale University, Denizli, Turkey



House divided: Computer scientists versus physician scientists

Skin cancer diagnosis: The importance of domain-specific knowledge

	Publications	Citations	Average citations per item	H index
Medical	463	14261	30.80	58
Dermatology	53	1090	20.57	16
Interdisciplinary	93	2261	24.31	26
Multidisciplinary	317	10910	34.42	46
Non-medical	867	8117	9.43	43

Levent Tasli, Department of Dermatology, Pamukkale University, Denizli, Turkey

Multidisciplinary studies are most influential

The importance of domain-specific knowledge

“Skin disease, the most overwhelming kind of malignancy, manifests when skin cells lose control”

“Skin cancer disease can be classified into two groups: benign and malignant”

“Benign cancers are classified as basal cell carcinoma and squamous cell carcinoma”

“Computer science, the most important type of machine learning, manifests when computer scientists loose control”

“Artificial learning can be classified into two groups: supervised and unsupervised”

“Supervised learning can be classified into reinforcement learning and machine learning”.

The importance of domain-specific knowledge:

Problem formulation

Data collection

Data preprocessing

Modeling

Interpretation of results

Understand and predict limitations and exceptions

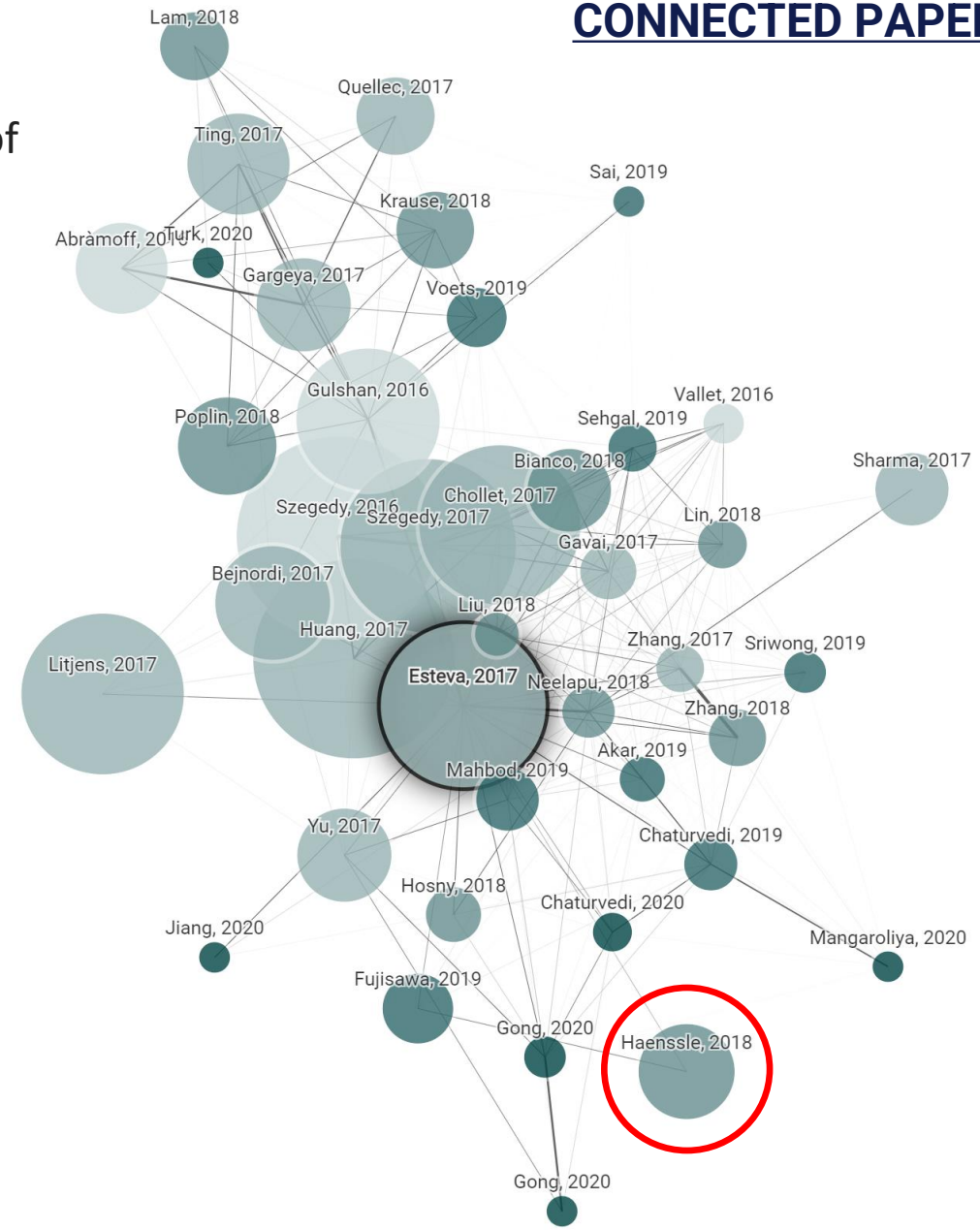
Make the algorithm fit for intended use

Applied computer science

CONNECTED PAPERS

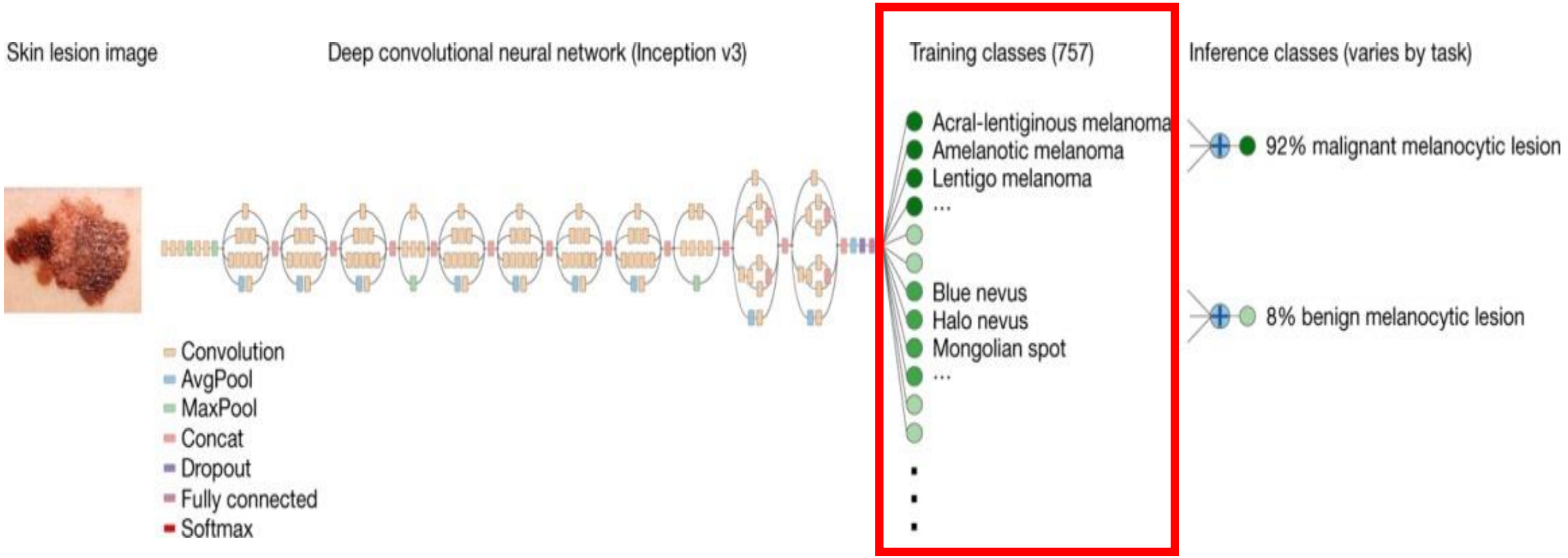
nature

Esteva, A., Kuprel, B., Novoa, R. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542,** 115–118 (2017)

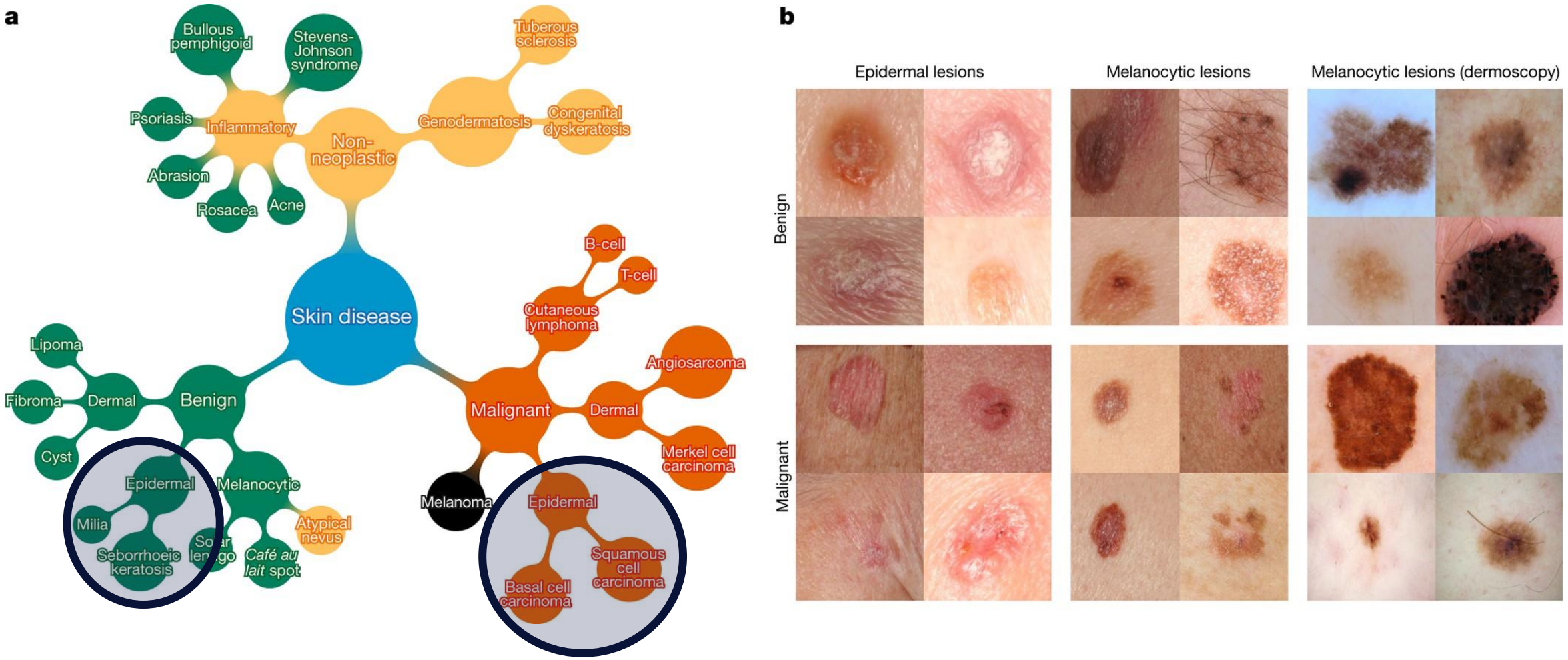


nature

Esteva, A., Kuprel, B., Novoa, R. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542,** 115–118 (2017)



Framing of the problem

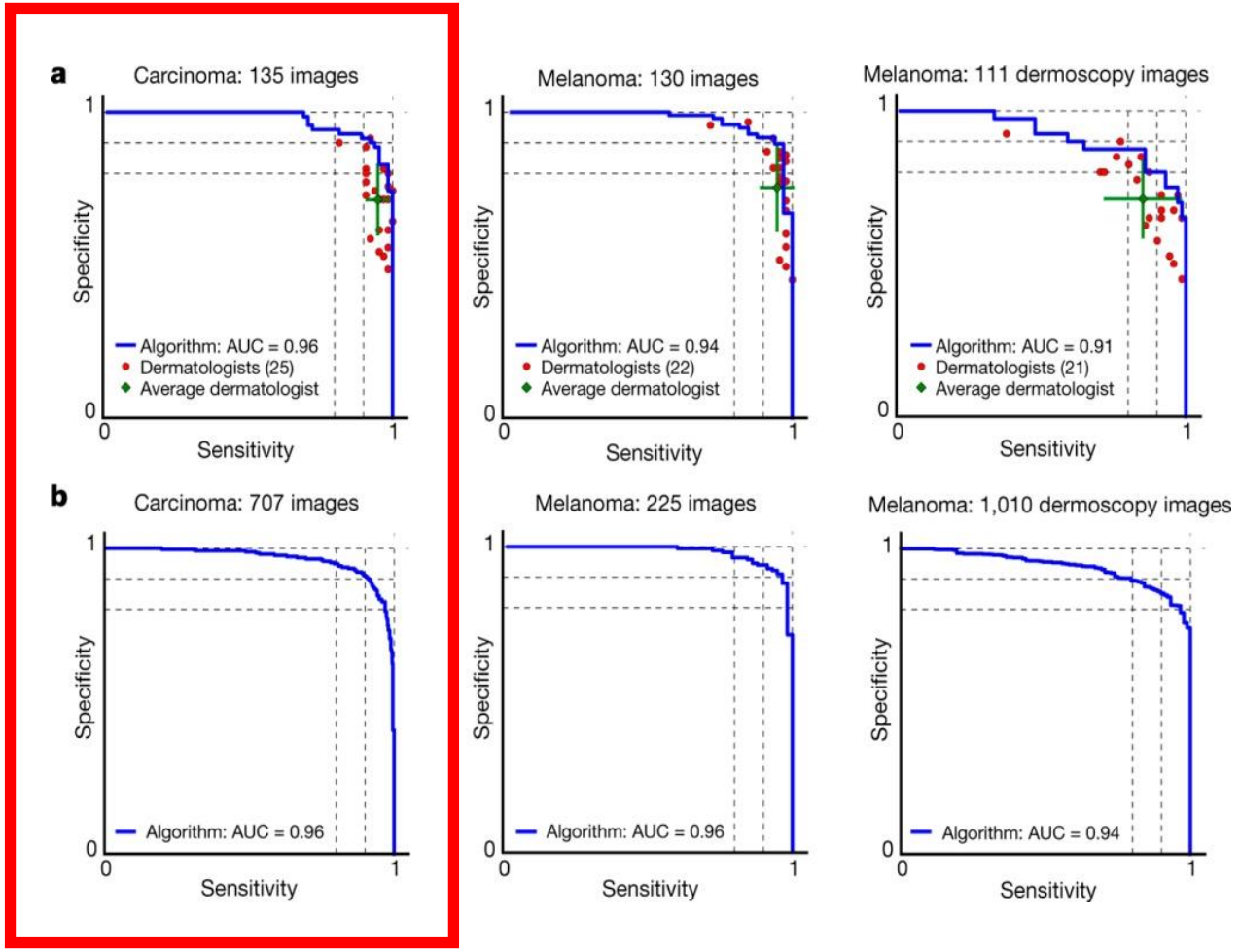


A Esteva *et al. Nature* 1–4 (2017) doi:10.1038/nature21056

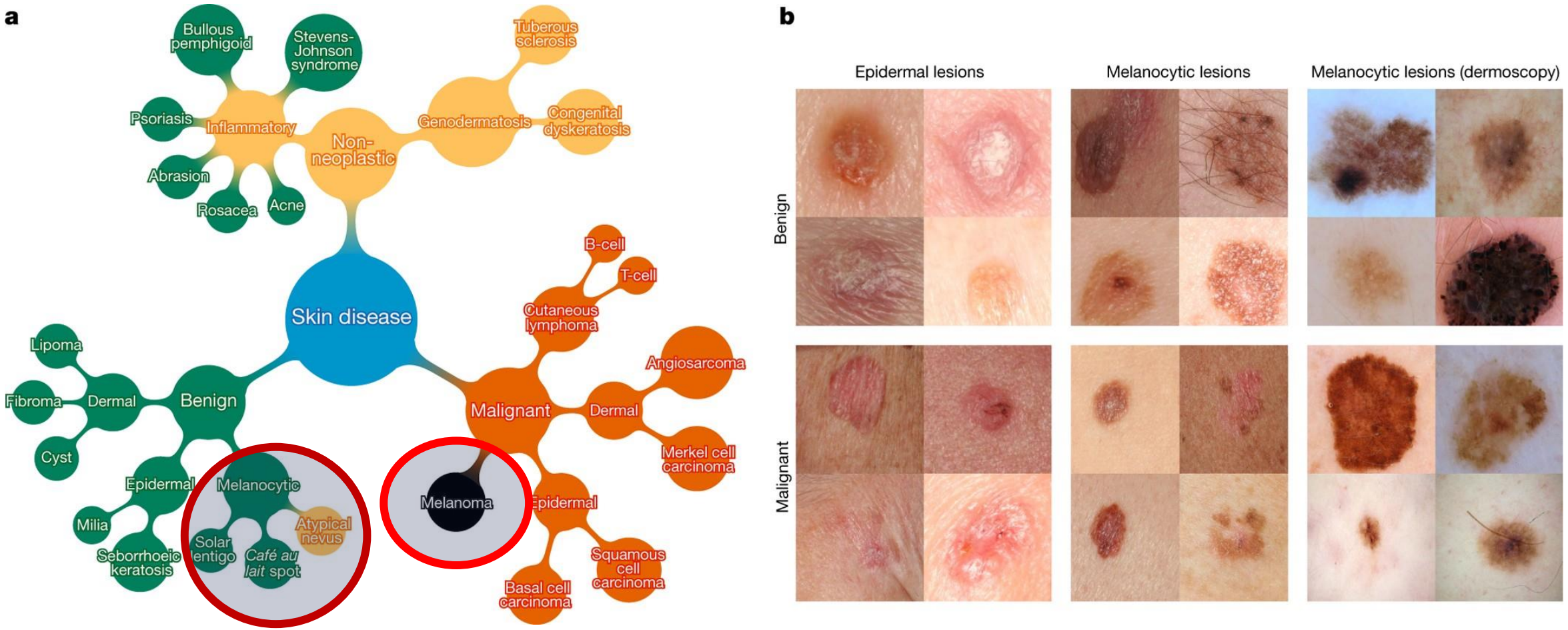
nature

nature

Esteva, A., Kuprel, B., Novoa, R. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542,** 115–118 (2017)



Framing of the problem



A Esteva *et al. Nature* 1–4 (2017) doi:10.1038/nature21056

nature

“…….when people say Kaggle, I have done this competition, I kind of roll my eyes, because they’re focusing on sort of 10% of the problem”

Ian Scott, Chief Science Officer at Omnia AI

The importance of domain-specific knowledge:

Problem formulation

Data collection

Data preprocessing

Modeling

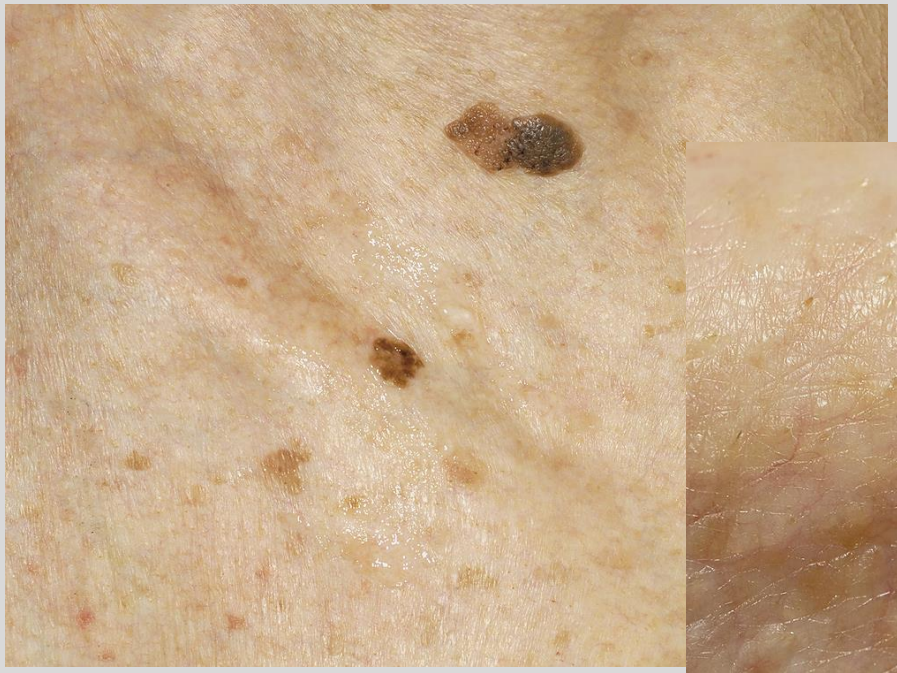
Interpretation of results

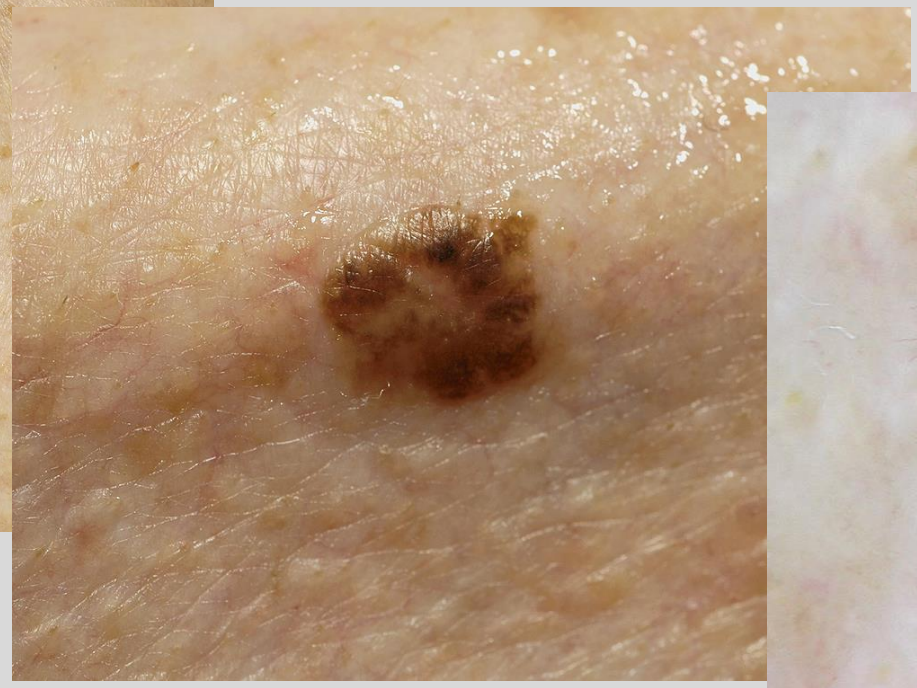
Understand and predict limitations and exceptions

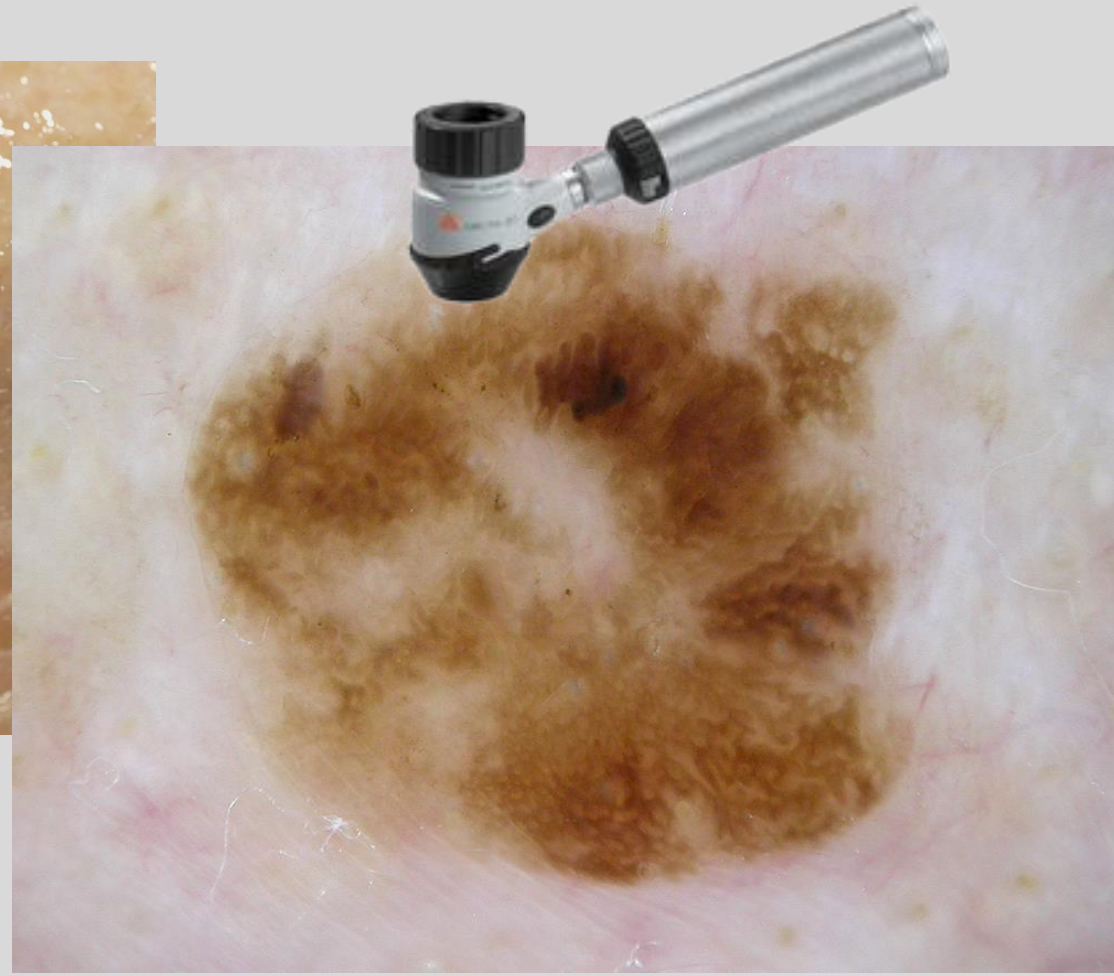
Make the algorithm fit for intended use

Applied computer science

The importance of domain-specific knowledge: Data collection







Search by filters Search by name

Search images

APPLIED FILTERS Clear applied filters

DIAGNOSTIC ATTRIBUTES

- BENIGN OR MALIGNANT
- LESION DIAGNOSIS

CLINICAL ATTRIBUTES

- APPROXIMATE AGE
- GENERAL ANATOMIC SITE
- CLINICAL SIZE - LONGEST DIAMETER (MM)
- TYPE OF DIAGNOSIS
- FAMILY HISTORY OF MELANOMA
- MELANOMA CLASS
- MELANOMA MITOTIC INDEX
- MELANOMA THICKNESS (MM)
- MELANOMA TYPE
- MELANOMA ULCERATION
- MELANOCYTIC
- NEVUS TYPE
- PERSONAL HISTORY OF MELANOMA
- SEX

TECHNOLOGICAL ATTRIBUTES

- DERMOSCOPIC TYPE

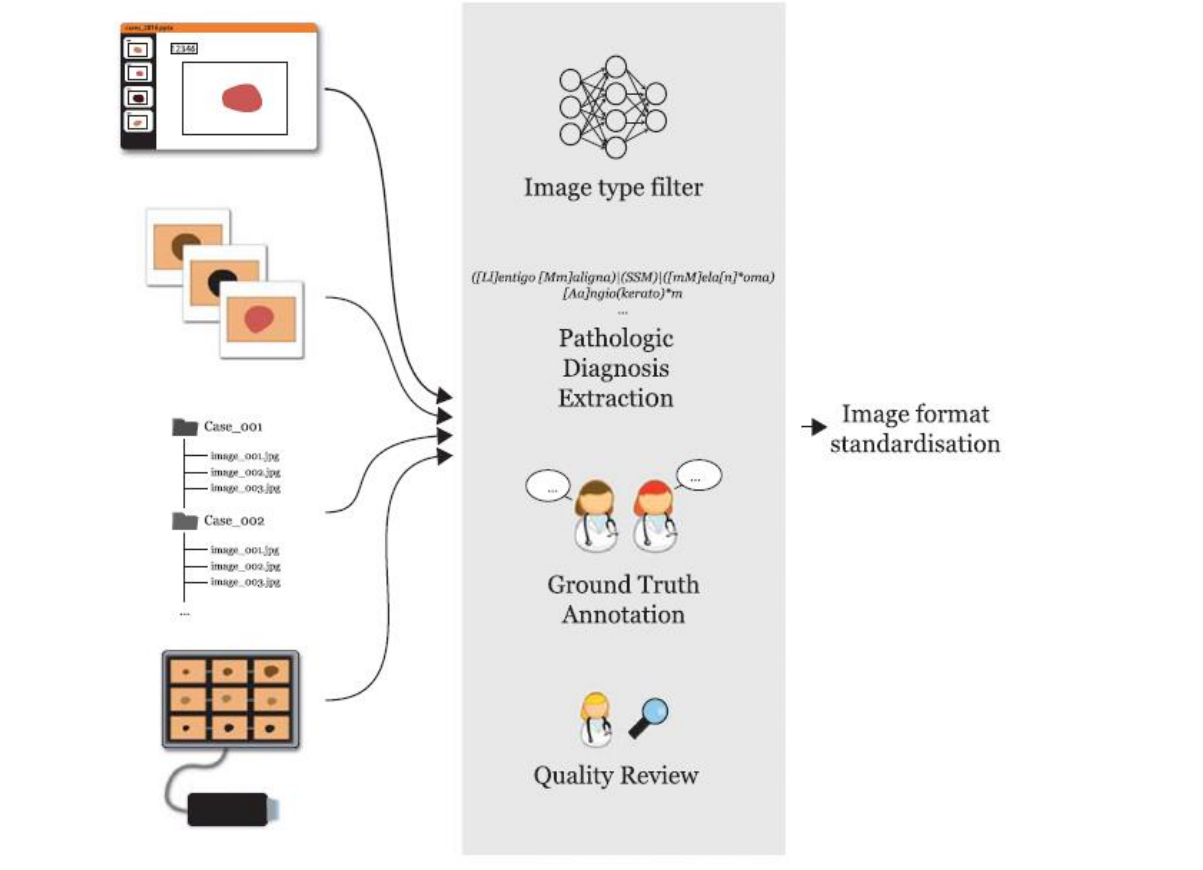
Shown images: **1-80**. Total amount of images: **69445**.

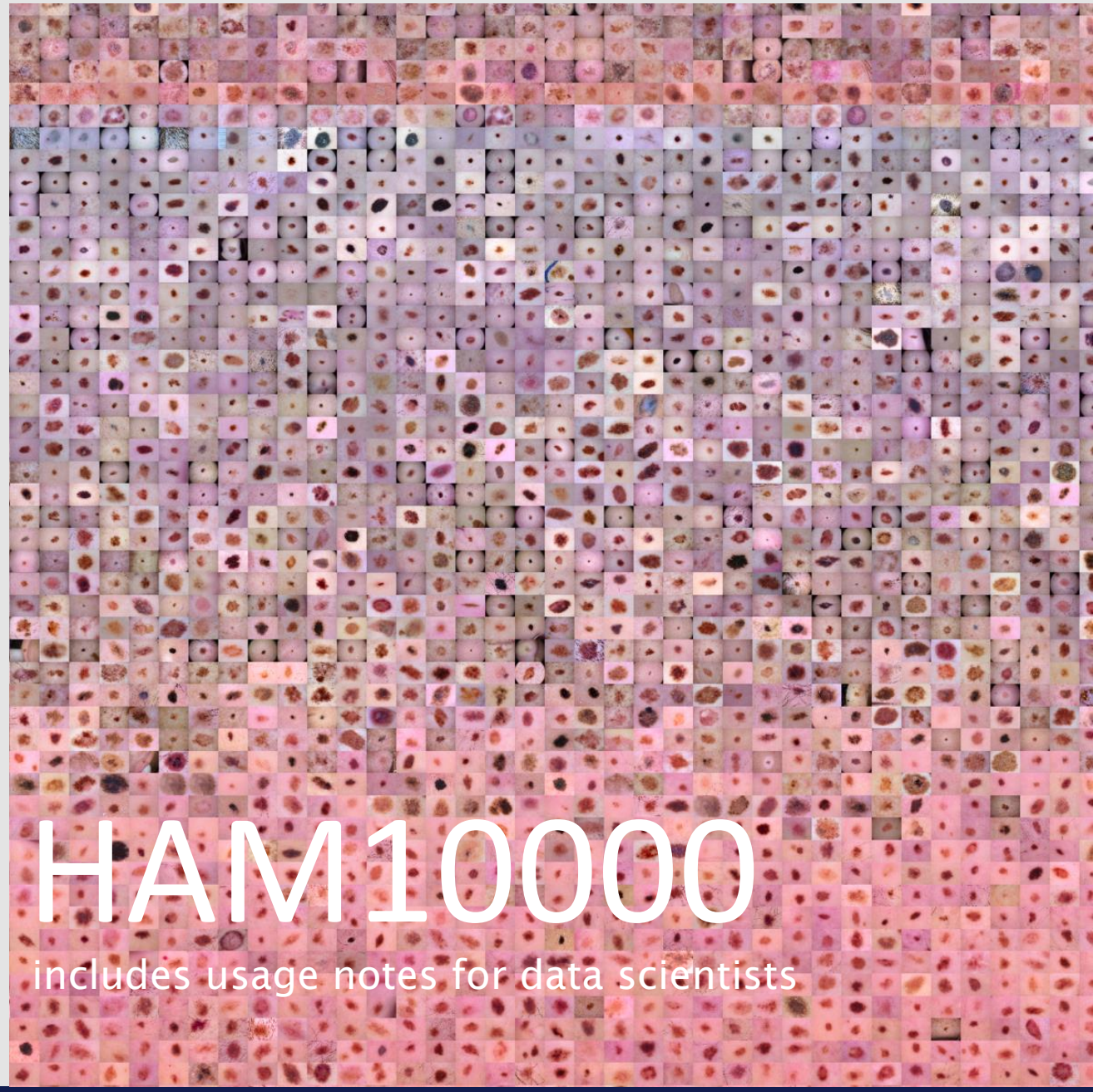
Select All on the Page for Download

« ‹ 1 of 869 › »

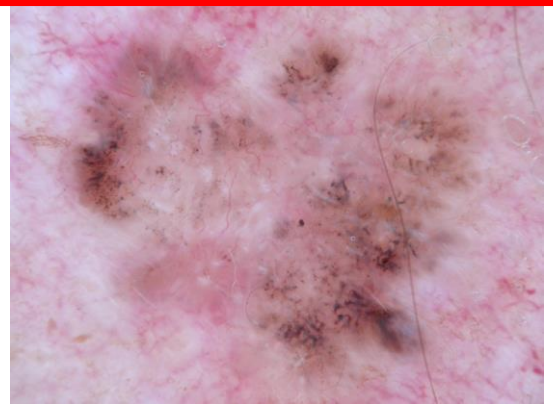
ISIC_7216877

Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5,** 180161 (2018). https://doi.org/10.1038/sdata.2018.161

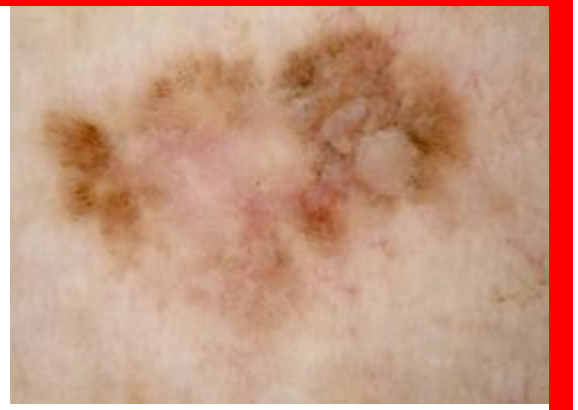




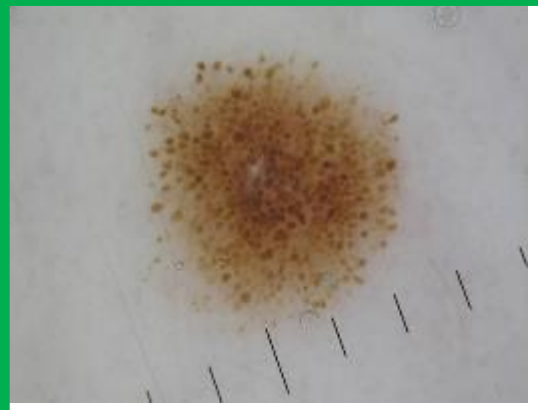
Melanoma



Basal cell carcinoma



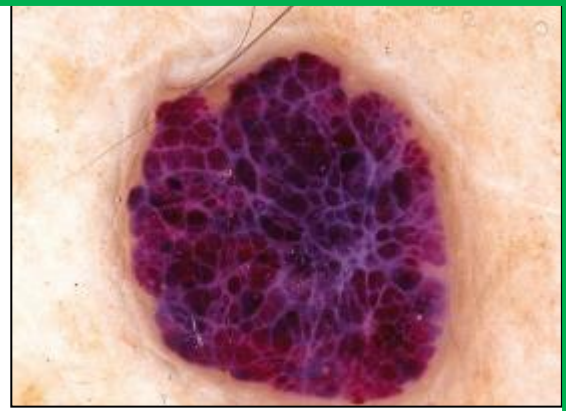
Squamous cell carcinoma



Nevus

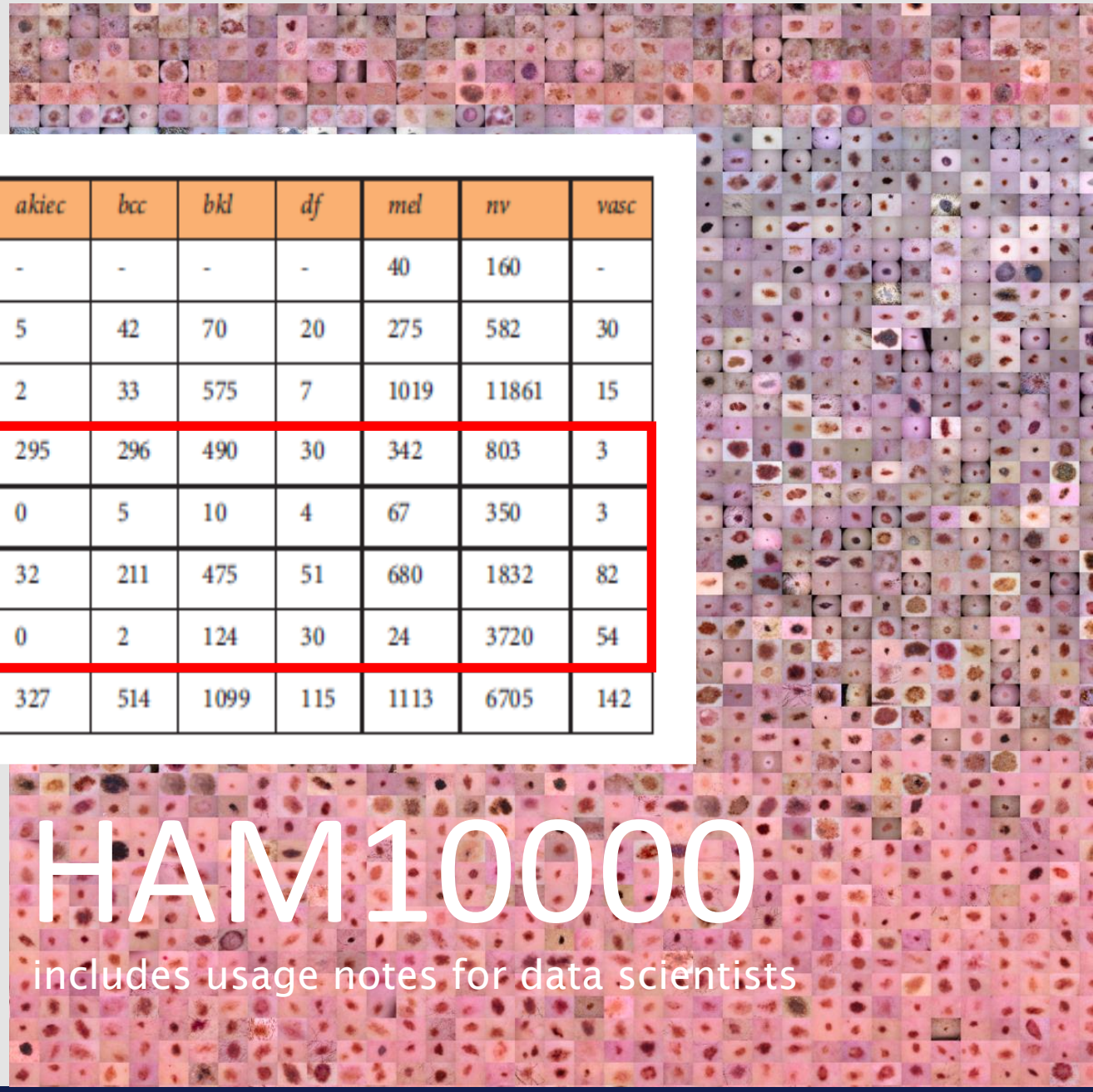
Benign keratoses

Dermatofibroma



Angioma

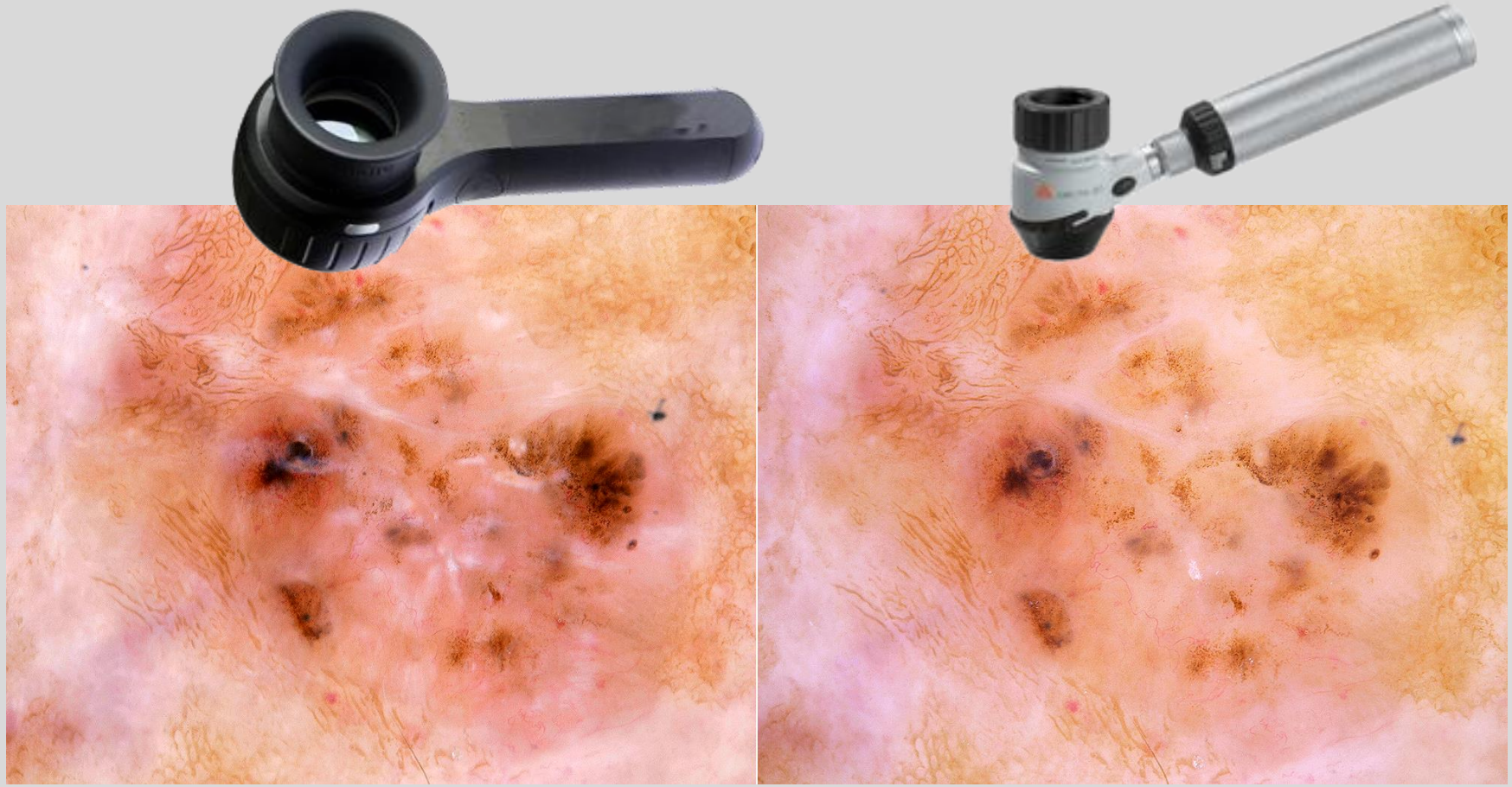
Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5,** 180161 (2018). https://doi.org/10.1038/sdata.2018.161



Dataset	*License*	*Total images*	*Pathologic verification (%)*	*akiec*	*bcc*	*bkl*	*df*	*mel*	*nv*	*vasc*
PH2	Research&Education[a]	200	20.5%	-	-	-	-	40	160	-
Atlas	No license	1024	unknown	5	42	70	20	275	582	30
ISIC 2017[b]	CC-0	13786	26.3%	2	33	575	7	1019	11861	15
Rosendahl	CC BY-NC 4.0	2259	100%	295	296	490	30	342	803	3
ViDIR Legacy	CC BY-NC 4.0	439	100%	0	5	10	4	67	350	3
ViDIR Current	CC BY-NC 4.0	3363	77.1%	32	211	475	51	680	1832	82
ViDIR MoleMax	CC BY-NC 4.0	3954	1.2%	0	2	124	30	24	3720	54
HAM10000	CC BY-NC 4.0	**10015**	53.3%	327	514	1099	115	1113	6705	142

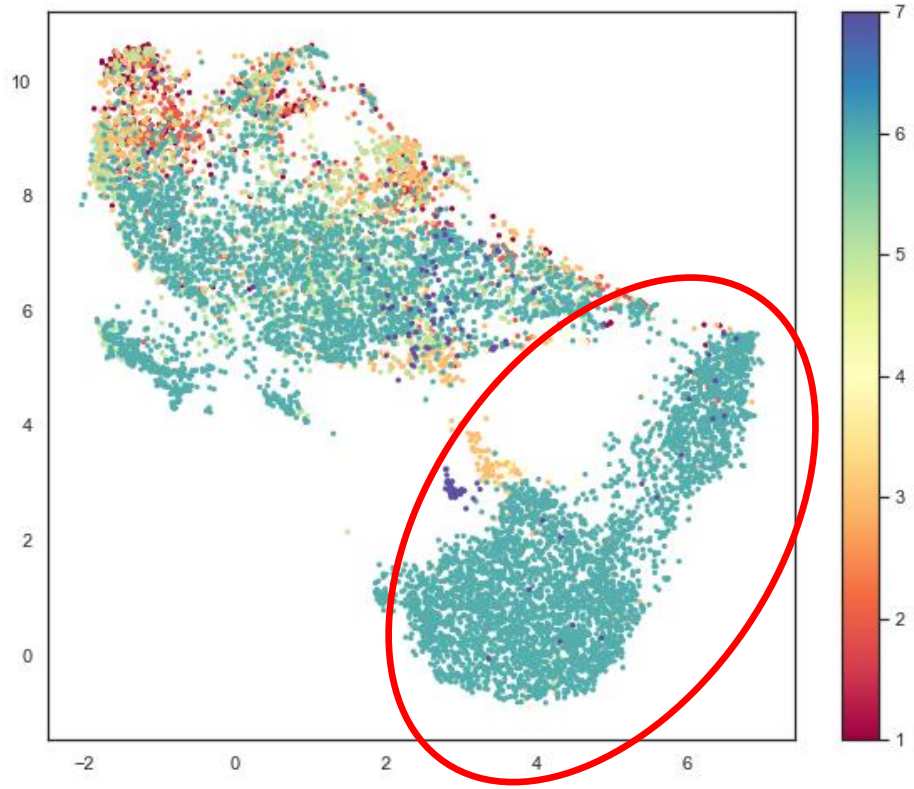
HAM10000

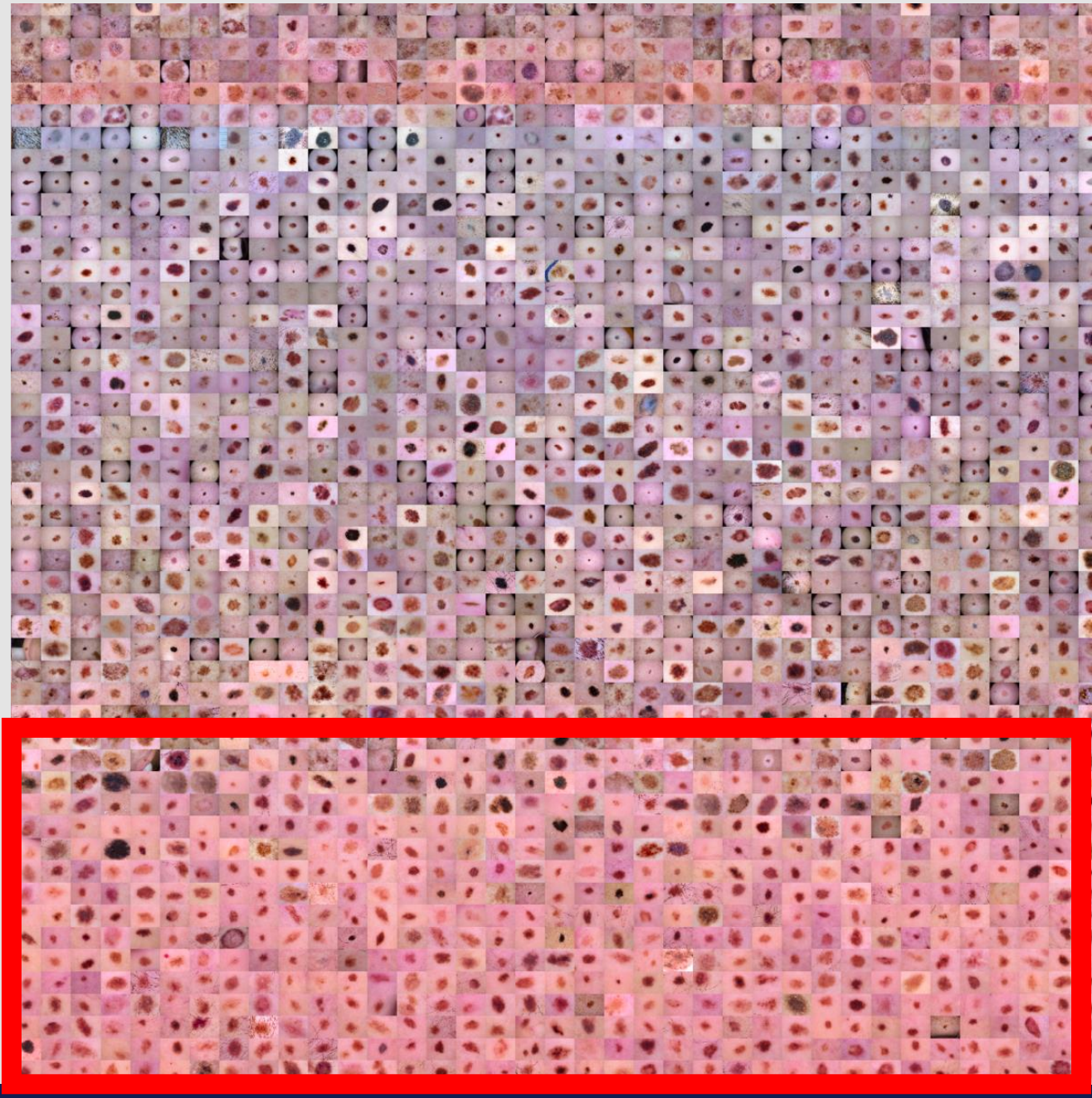
includes usage notes for data scientists



Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5,** 180161 (2018). https://doi.org/10.1038/sdata.2018.161

Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5,** 180161 (2018). https://doi.org/10.1038/sdata.2018.161





The importance of domain-specific knowledge: Collider bias

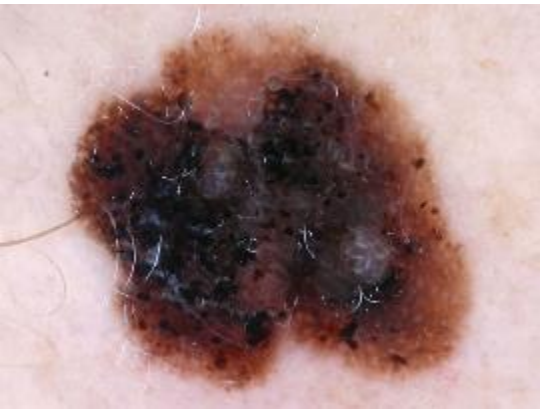
Many images of benign lesions

Only suspicious lesions

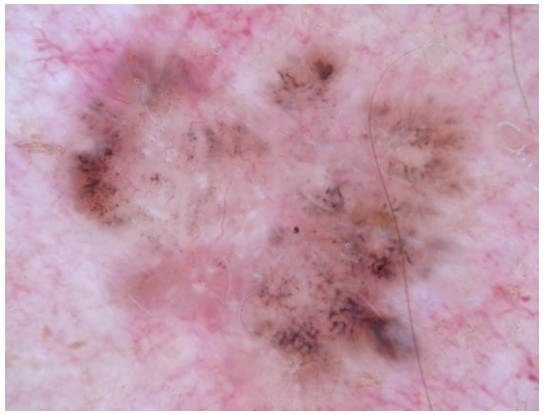
Spurious correlation: Images taken with device X are probably benign

There is no causal relation between images taken with device X and outcome

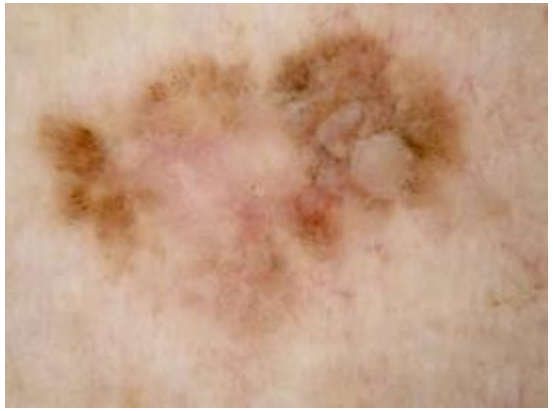
The importance of domain-specific knowledge: Class imbalance



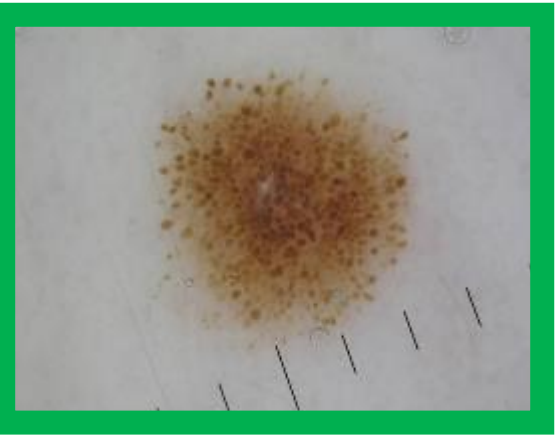
Melanoma



Basal cell carcinoma



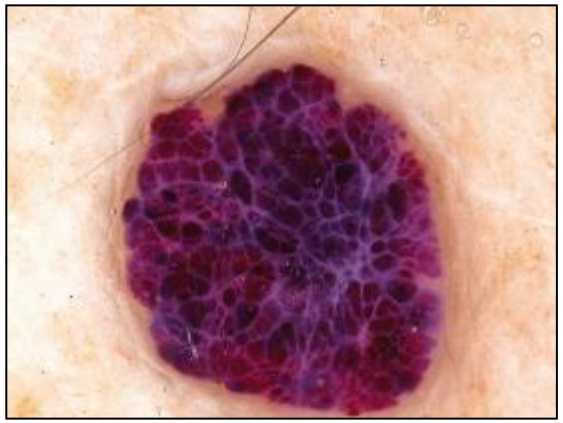
Squamous cell carcinoma



Nevus

Benign keratoses

Dermatofibroma



Angioma

The importance of domain-specific knowledge: Class imbalance

Data level methods

- Oversampling
- Undersampling

Classifier level methods

- Thresholding → **Thresholds are problem-dependent**
- Cost sensitive learning
- One-class classification
- Hybrid of methods

nature medicine

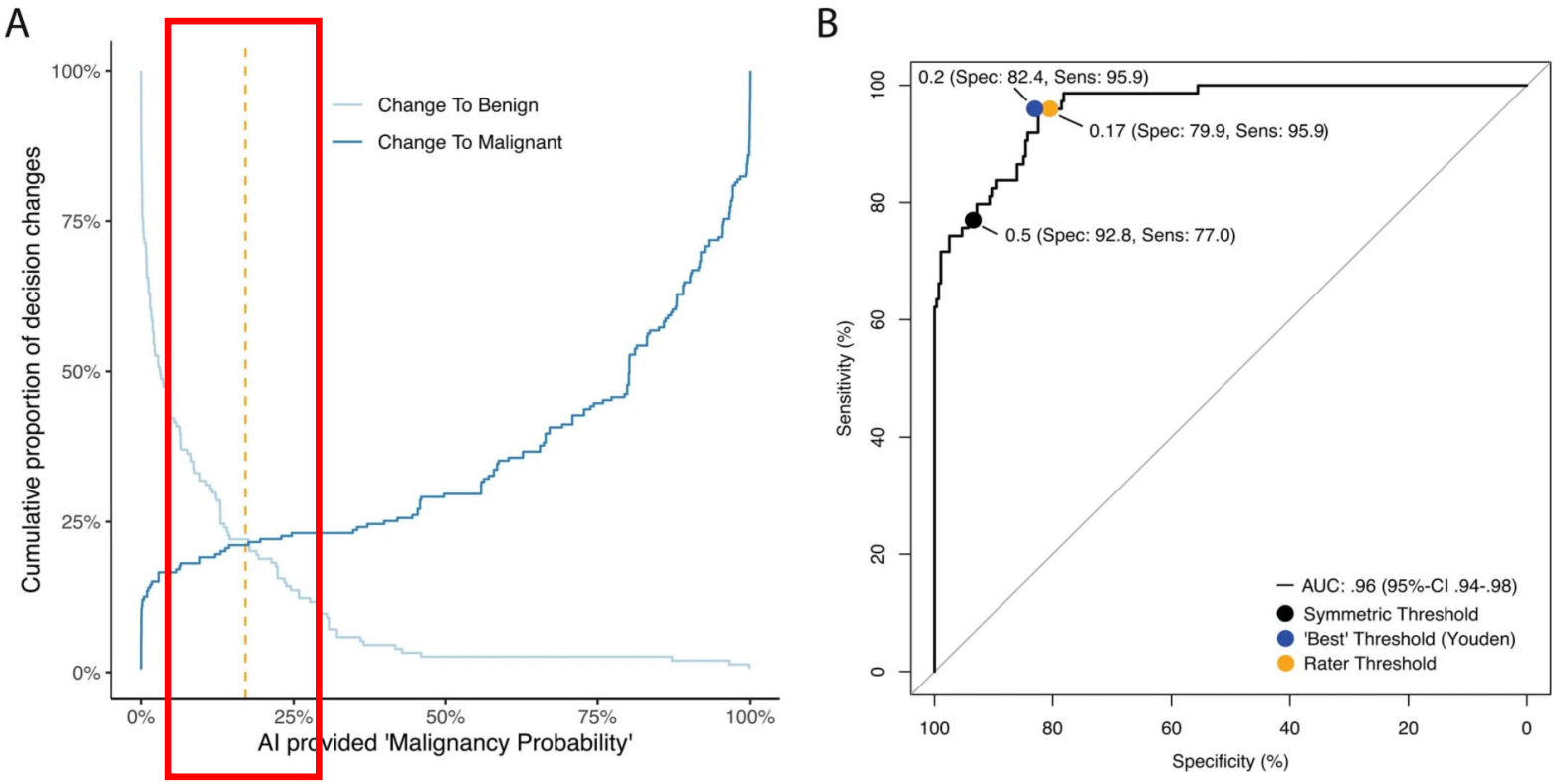
Human–computer collaboration for skin cancer recognition

Philipp Tschandl[1,17], Christoph Rinner[2,17], Zoe Apalla[3], Giuseppe Argenziano[4], Noel Codella[5], Allan Halpern[6], Monika Janda[7], Aimilios Lallas[3], Caterina Longo[8,9], Josep Malvehy[10,11], John Paoli[12,13], Susana Puig[10,11], Cliff Rosendahl[14], H. Peter Soyer[15], Iris Zalaudek[16] and Harald Kittler[1]

To find optimal thresholds you need domain specific insight

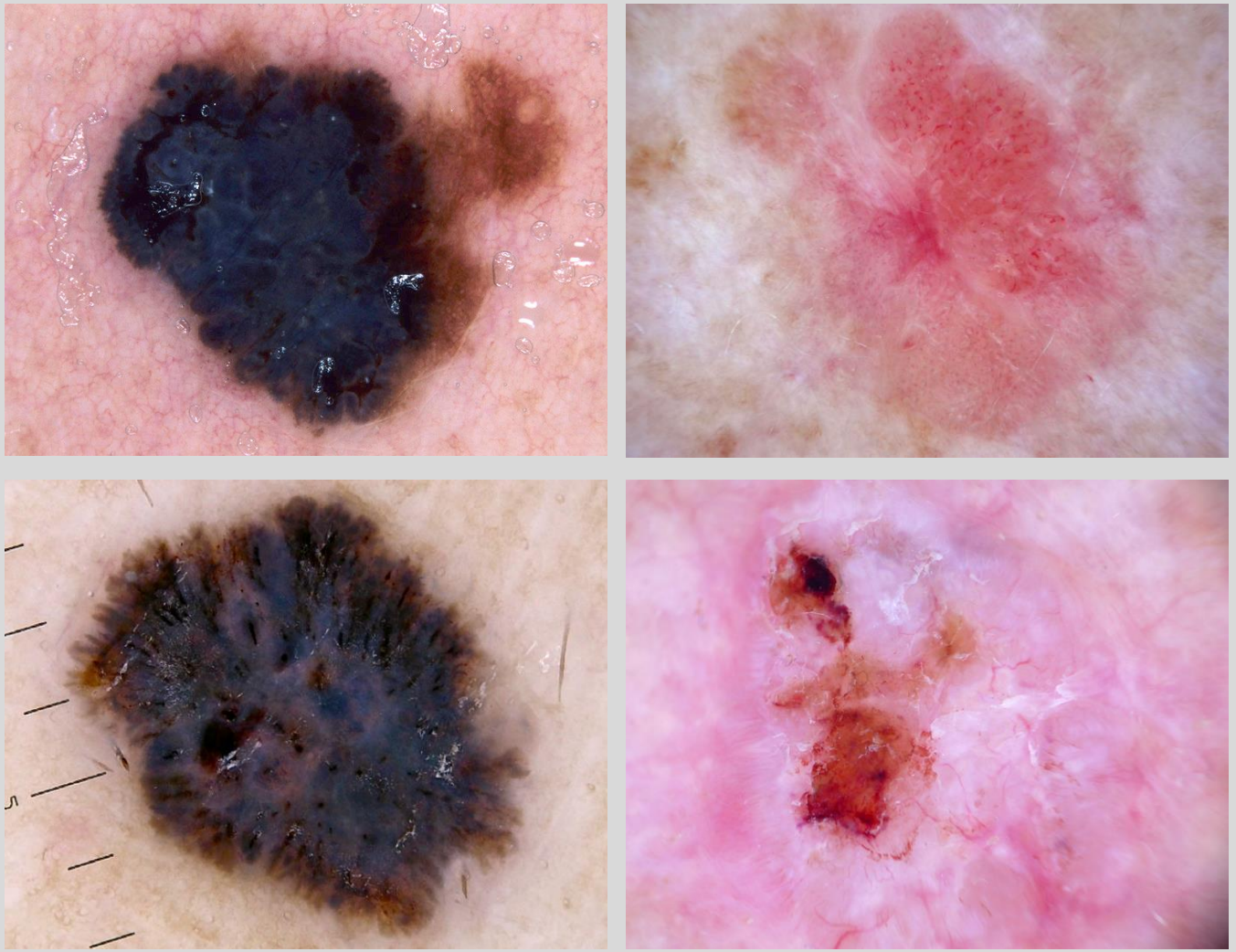
Extended Data Fig. 2: Raters choose an asymmetric decision cutoff for malignancy.

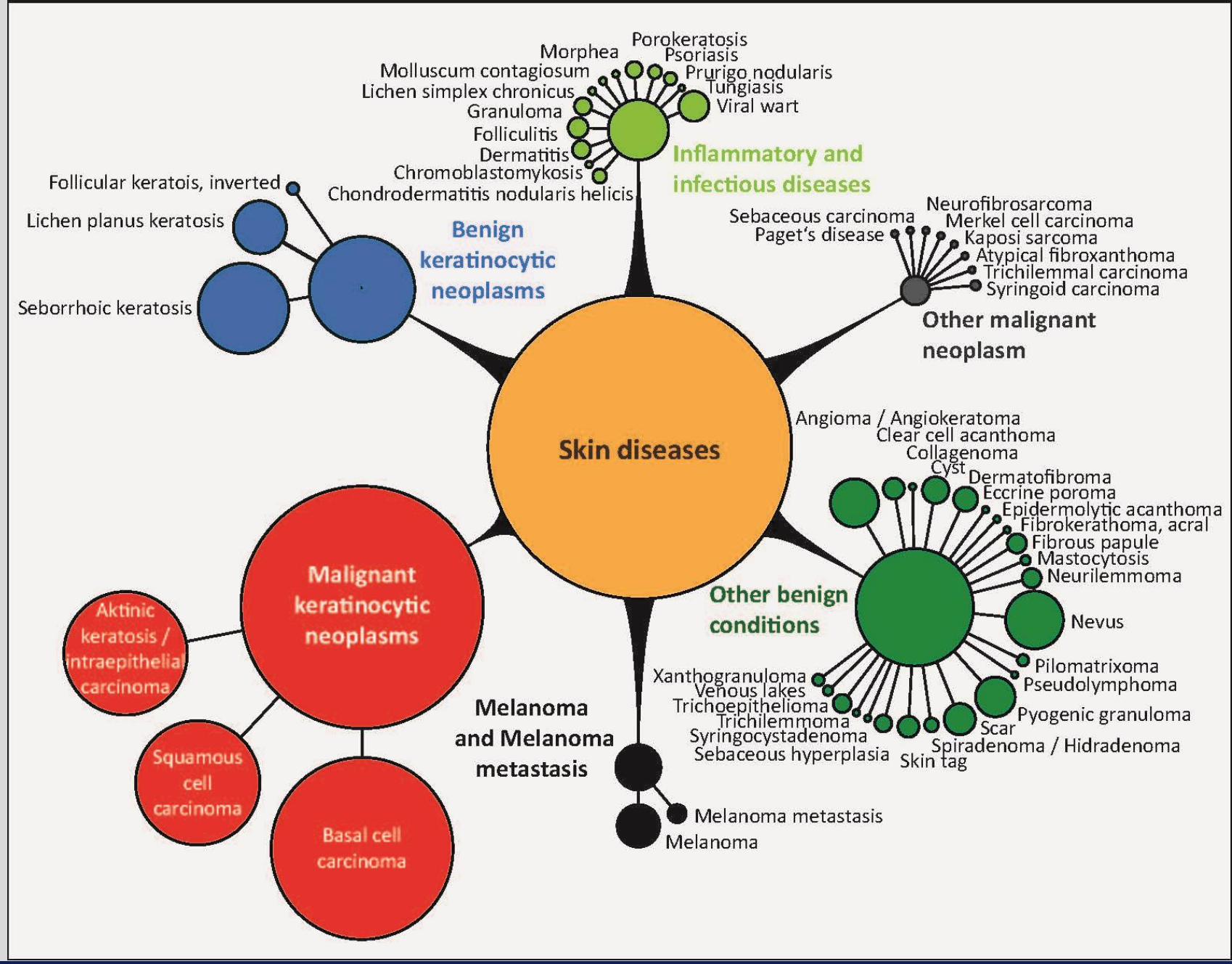
From: Human–computer collaboration for skin cancer recognition



a, When changing answers from benign to malignant (dark blue) or malignant to benign (light blue) diagnoses, the average cutoff for the AI-provided malignancy-probability was not 50% but <25% (yellow dotted line). **b**, On the ROC-curve for detecting malignant cases of the underlying AI (black line), this cutoff chosen inherently by the users (yellow dot), that is without instructions or prior knowledge about the AI accuracy, had a higher sensitivity and was closer to the ideal cutoff (blue dot), as measured by Youden's index, than the 'symmetric' 50% cutoff (black dot).

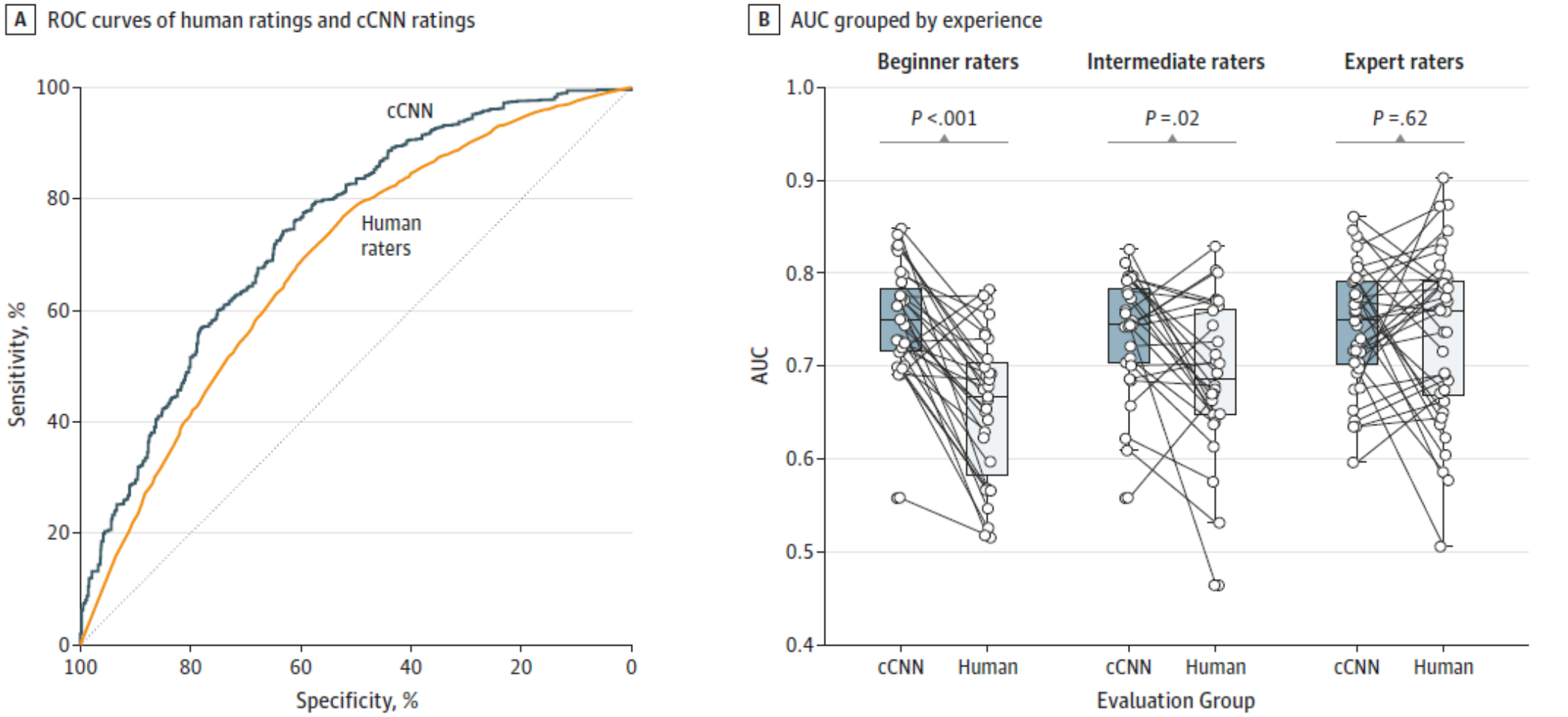
The importance of domain-specific knowledge: Selection bias

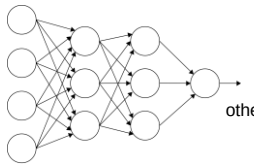


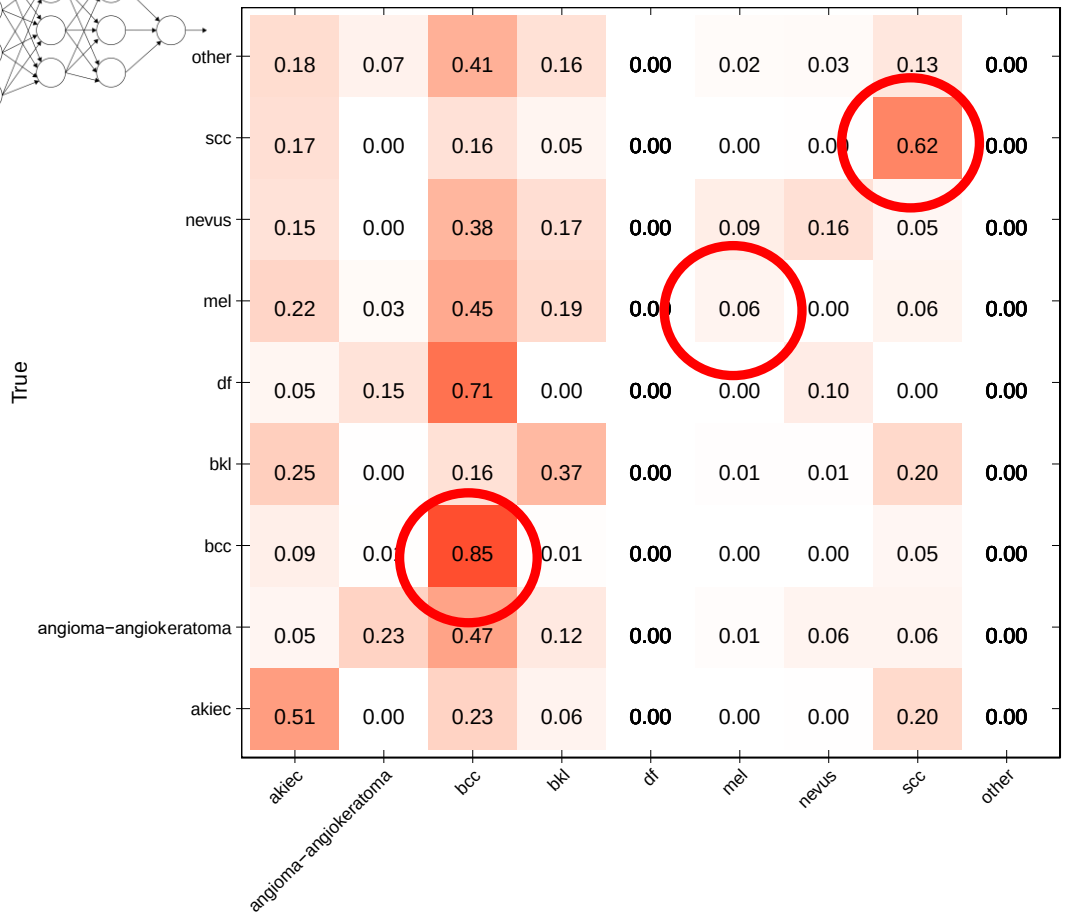

Skin diseases
Inflammatory and infectious diseases
Porokeratosis
Morphea
Psoriasis
Molluscum contagiosum
Prurigo nodularis
Lichen simplex chronicus
Tungiasis
Granuloma
Viral wart
Folliculitis
Dermatitis
Chromoblastomykosis
Chondrodermatitis nodularis helicis
Benign keratinocytic neoplasms
Follicular keratois, inverted
Lichen planus keratosis
Seborrhoic keratosis
Other malignant neoplasm
Sebaceous carcinoma
Paget's disease
Neurofibrosarcoma
Merkel cell carcinoma
Kaposi sarcoma
Atypical fibroxanthoma
Trichilemmal carcinoma
Syringoid carcinoma
Other benign conditions
Angioma / Angiokeratoma
Clear cell acanthoma
Collagenoma
Cyst
Dermatofibroma
Eccrine poroma
Epidermolytic acanthoma
Fibrokerathoma, acral
Fibrous papule
Mastocytosis
Neurilemmoma
Nevus
Pilomatrixoma
Pseudolymphoma
Pyogenic granuloma
Scar
Spiradenoma / Hidradenoma
Skin tag
Xanthogranuloma
Venous lakes
Trichoepithelioma
Trichilemmoma
Syringocystadenoma
Sebaceous hyperplasia
Malignant keratinocytic neoplasms
Aktinic keratosis / intraepithelial carcinoma
Squamous cell carcinoma
Basal cell carcinoma
Melanoma and Melanoma metastasis
Melanoma metastasis
Melanoma

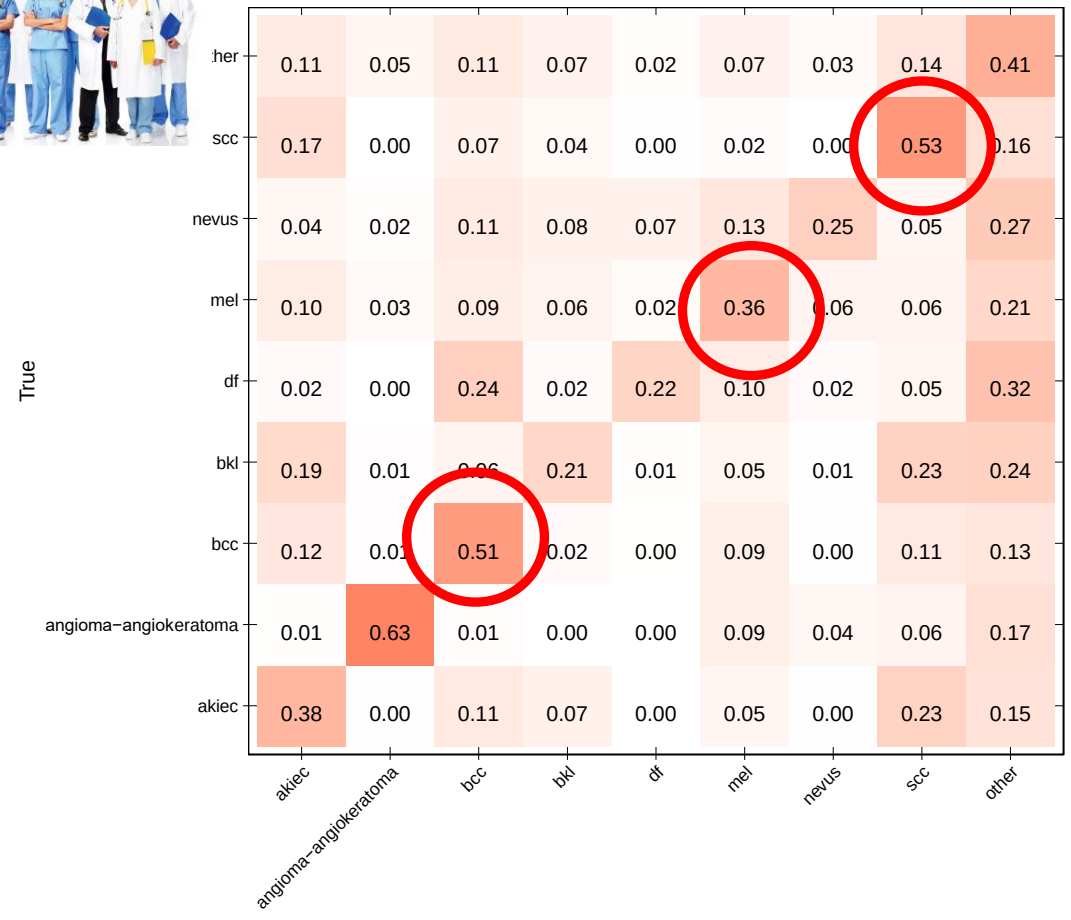
Expert-Level Diagnosis of Nonpigmented Skin Cancer by Combined Convolutional Neural Networks

Figure 1. Comparison of Skin Cancer Detection on Digital Images Between Human Readers and a Neural Network–Based Classifier






other 0.18 0.07 0.41 0.16 0.00 0.02 0.03 0.13 0.00
scc 0.17 0.00 0.16 0.05 0.00 0.00 0.00 0.62 0.00
nevus 0.15 0.00 0.38 0.17 0.00 0.09 0.16 0.05 0.00
mel 0.22 0.03 0.45 0.19 0.00 0.06 0.00 0.06 0.00
df 0.05 0.15 0.71 0.00 0.00 0.00 0.10 0.00 0.00
bkl 0.25 0.00 0.16 0.37 0.00 0.01 0.01 0.20 0.00
bcc 0.09 0.01 0.85 0.01 0.00 0.00 0.00 0.05 0.00
angioma−angiokeratoma 0.05 0.23 0.47 0.12 0.00 0.01 0.06 0.06 0.00
akiec 0.51 0.00 0.23 0.06 0.00 0.00 0.00 0.20 0.00
True
akiec
angioma−angiokeratoma
bcc
bkl
df
mel
nevus
scc
other
Predicted


her 0.11 0.05 0.11 0.07 0.02 0.07 0.03 0.14 0.41
scc 0.17 0.00 0.07 0.04 0.00 0.02 0.00 0.53 0.16
nevus 0.04 0.02 0.11 0.08 0.07 0.13 0.25 0.05 0.27
mel 0.10 0.03 0.09 0.06 0.02 0.36 0.06 0.06 0.21
df 0.02 0.00 0.24 0.02 0.22 0.10 0.02 0.05 0.32
bkl 0.19 0.01 0.06 0.21 0.01 0.05 0.01 0.23 0.24
bcc 0.12 0.01 0.51 0.02 0.00 0.09 0.00 0.11 0.13
angioma−angiokeratoma 0.01 0.63 0.01 0.00 0.00 0.09 0.04 0.06 0.17
akiec 0.38 0.00 0.11 0.07 0.00 0.05 0.00 0.23 0.15
True
akiec
angioma−angiokeratoma
bcc
bkl
df
mel
nevus
scc
other
Predicted

Model's Prediction :
  R/O Dermatofibroma
Model's Output :
  Actinic keratosis : 0.0000
  Basal cell carcinoma : 0.0005
  Intraepithelial carcinoma : 0.0001
  Dermatofibroma : 0.0884
  Hemangioma : 0.0000
  Lentigo : 0.0008
  Malignant melanoma : 0.0000
  Melanocytic nevus : 0.0024
  Pyogenic granuloma : 0.0001
  Seborrheic keratosis : 0.0005
  Squamous cell carcinoma : 0.0006
  Wart : 0.0015

The importance of domain-specific knowledge: Verfication bias and ground truth

Any algorithm will inherit and amplify the implicit values of the people who collected the lesions and who decided on the ground truth

The importance of domain-specific knowledge:

Problem formulation

Data collection

Data preprocessing

Modeling

Interpretation of results

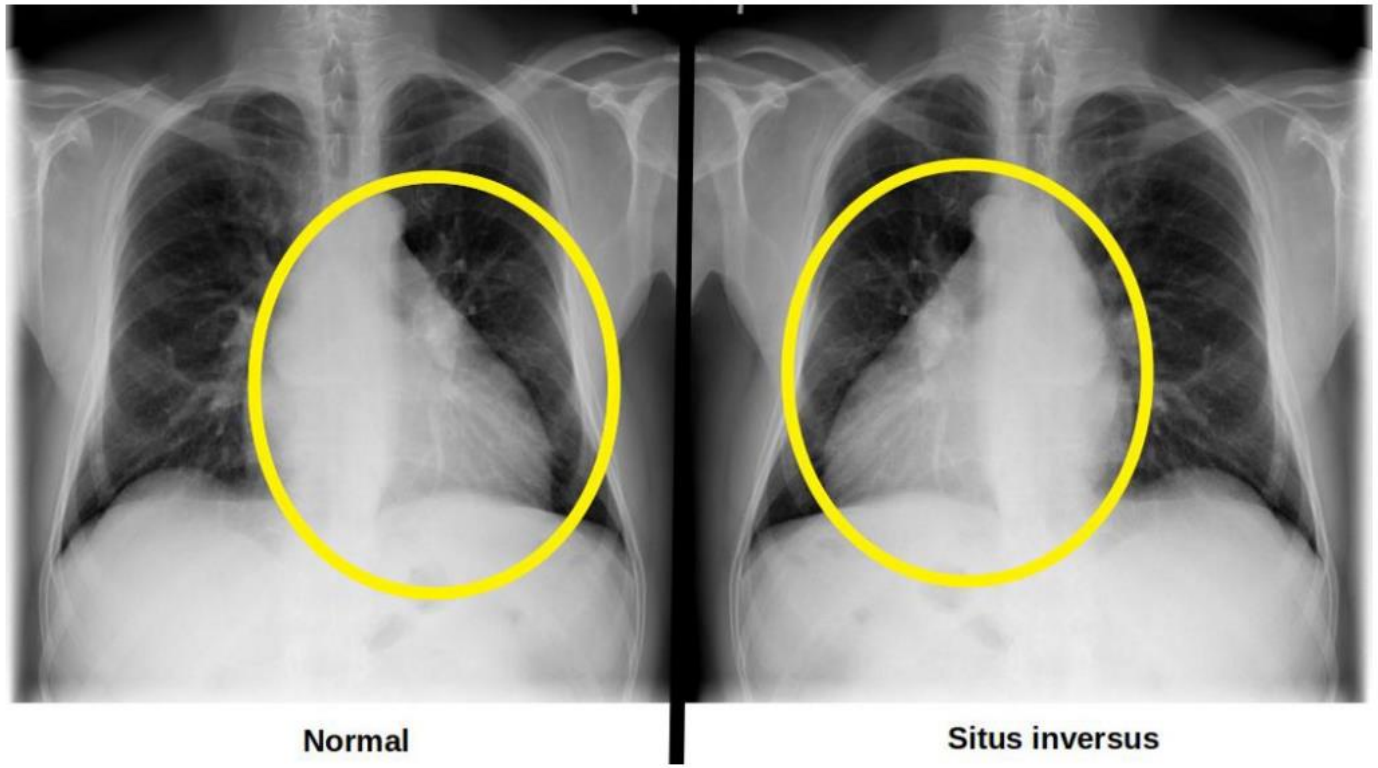
Understand and predict limitations and exceptions

Make the algorithm fit for intended use

Applied computer science

The importance of domain-specific knowledge: Image Augmentation

zooming in/out
rotating along the axis
vertical/horizontal flips
adjusting the brightness and sheer intensity etc.



Data augmentation may create issues with class labels if position is important. The yellow circle highlights the heart. This is the same image horizontally flipped. Image by author.

Source: Cody Glickman, PhD, https://towardsdatascience.com/

The importance of domain-specific knowledge: Interpretation of results

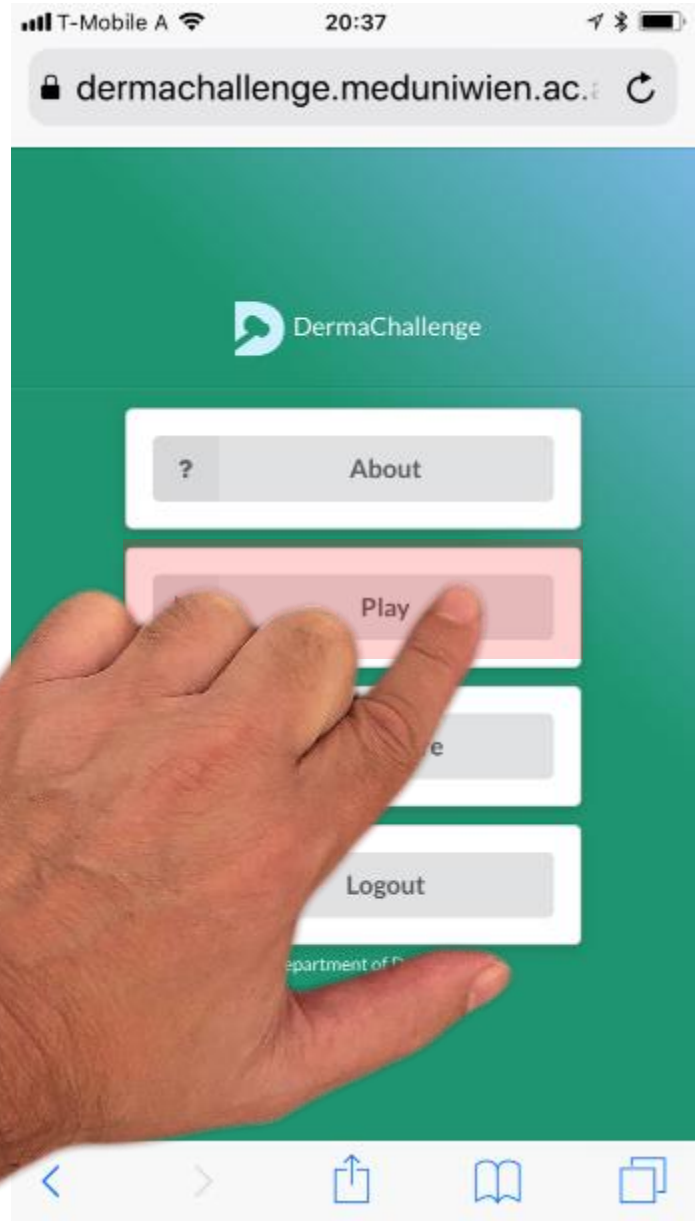
129 algorithms of
77 computer labs

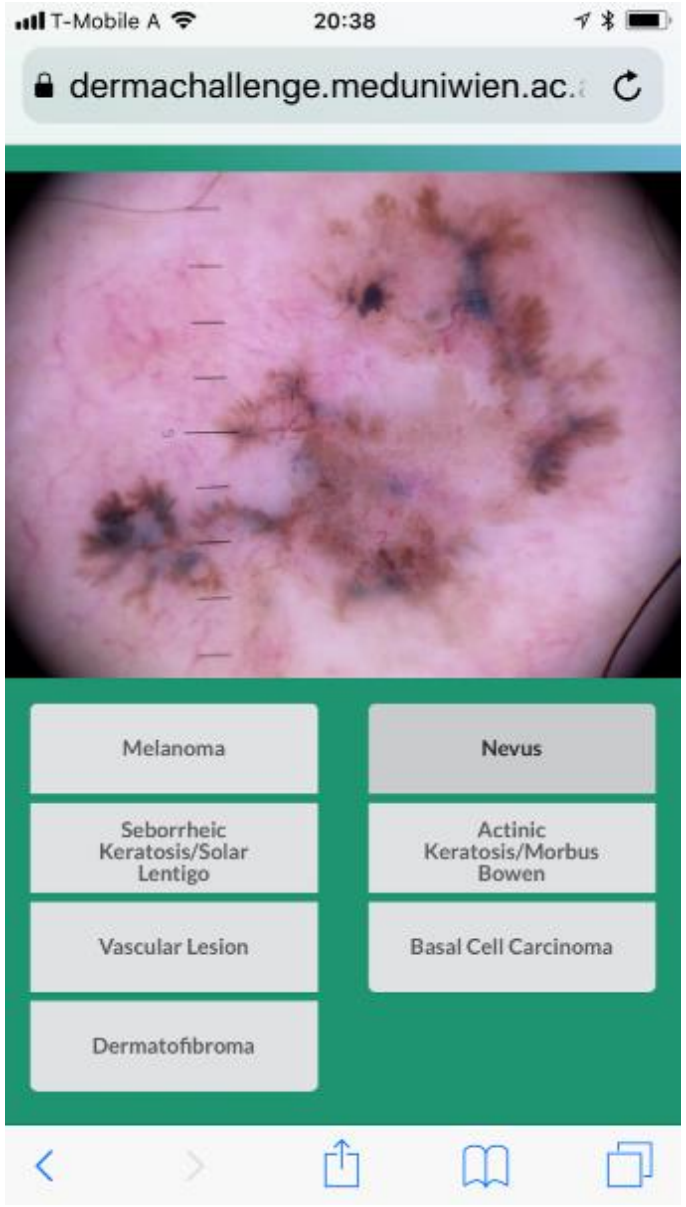
511 humans (with different level of experience)

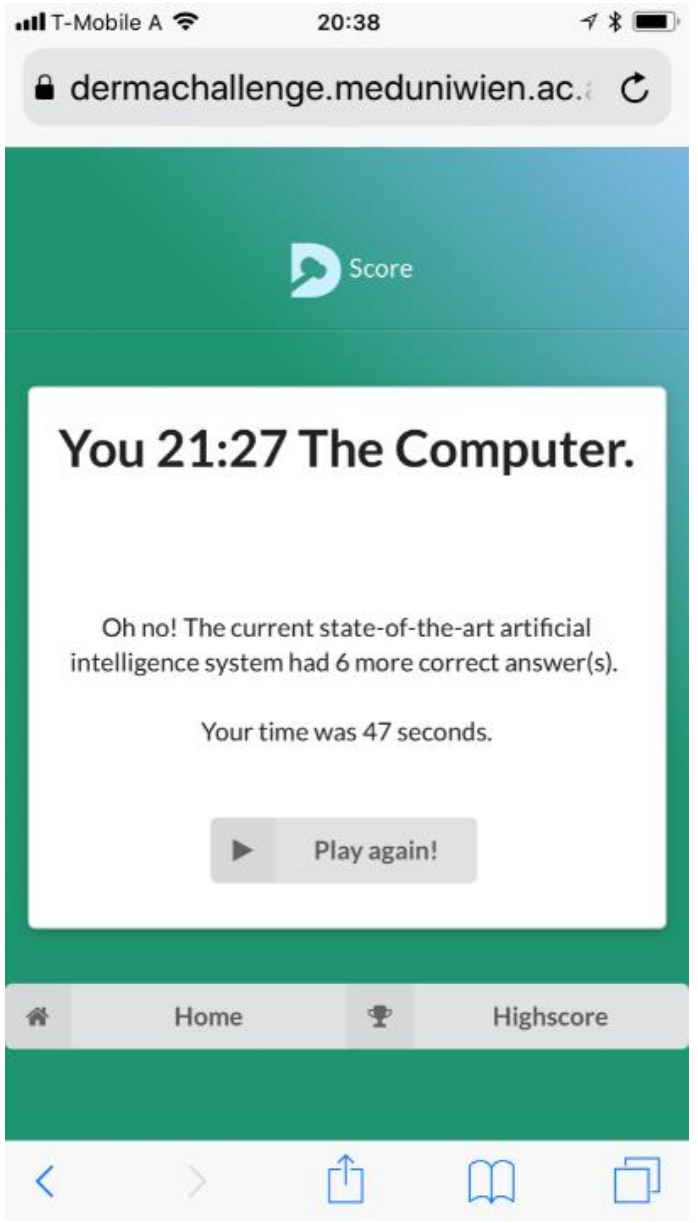
Training set: 10015 images

Test set: 1512 images

Batches of 30 lesions


T-Mobile A
20:37
dermachallenge.meduniwien.ac.
DermaChallenge
? About
Play
Logout

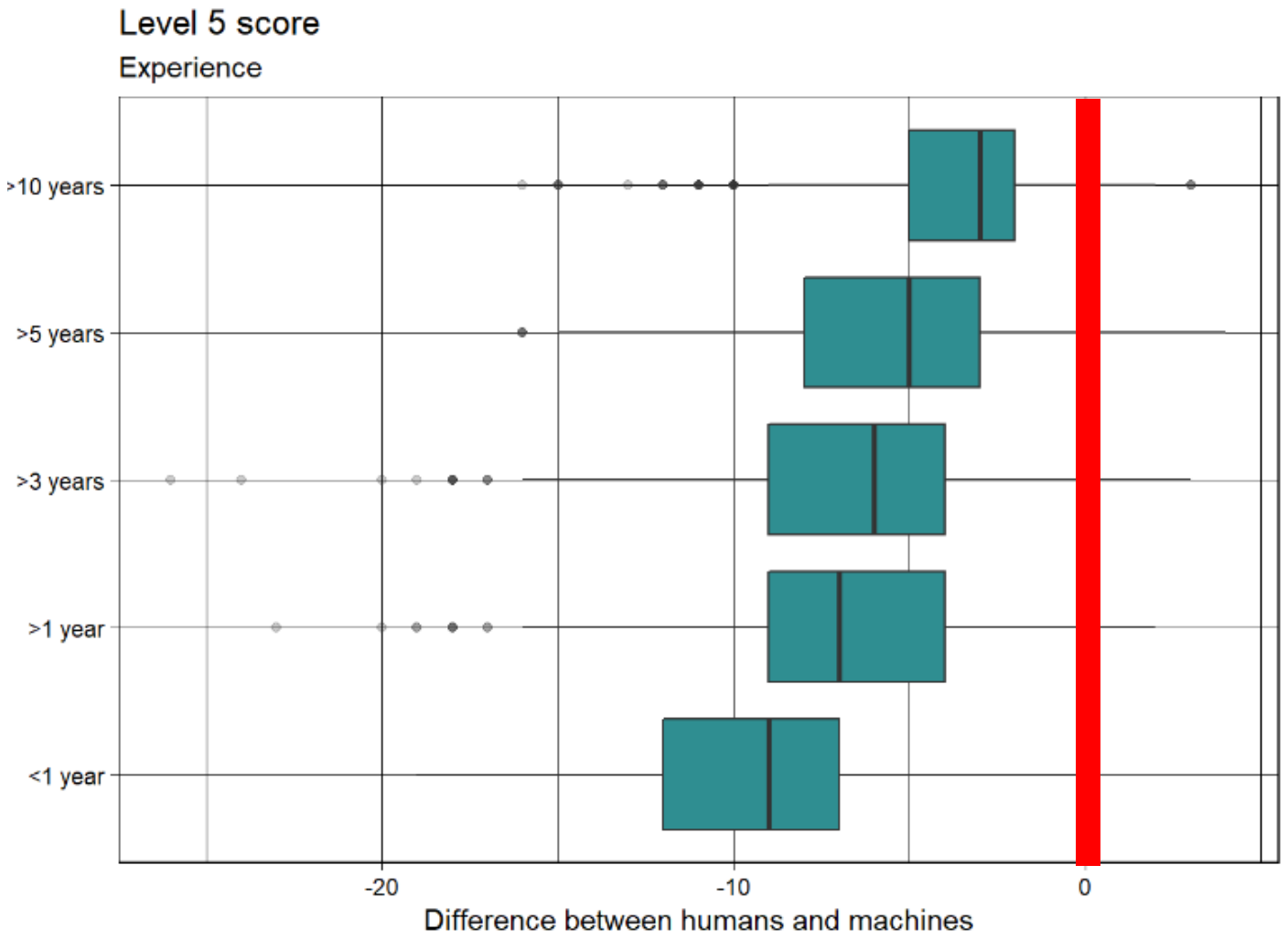

T-Mobile A
20:38
dermachallenge.meduniwien.ac.
Melanoma
Nevus
Seborrheic Keratosis/Solar Lentigo
Actinic Keratosis/Morbus Bowen
Vascular Lesion
Basal Cell Carcinoma
Dermatofibroma

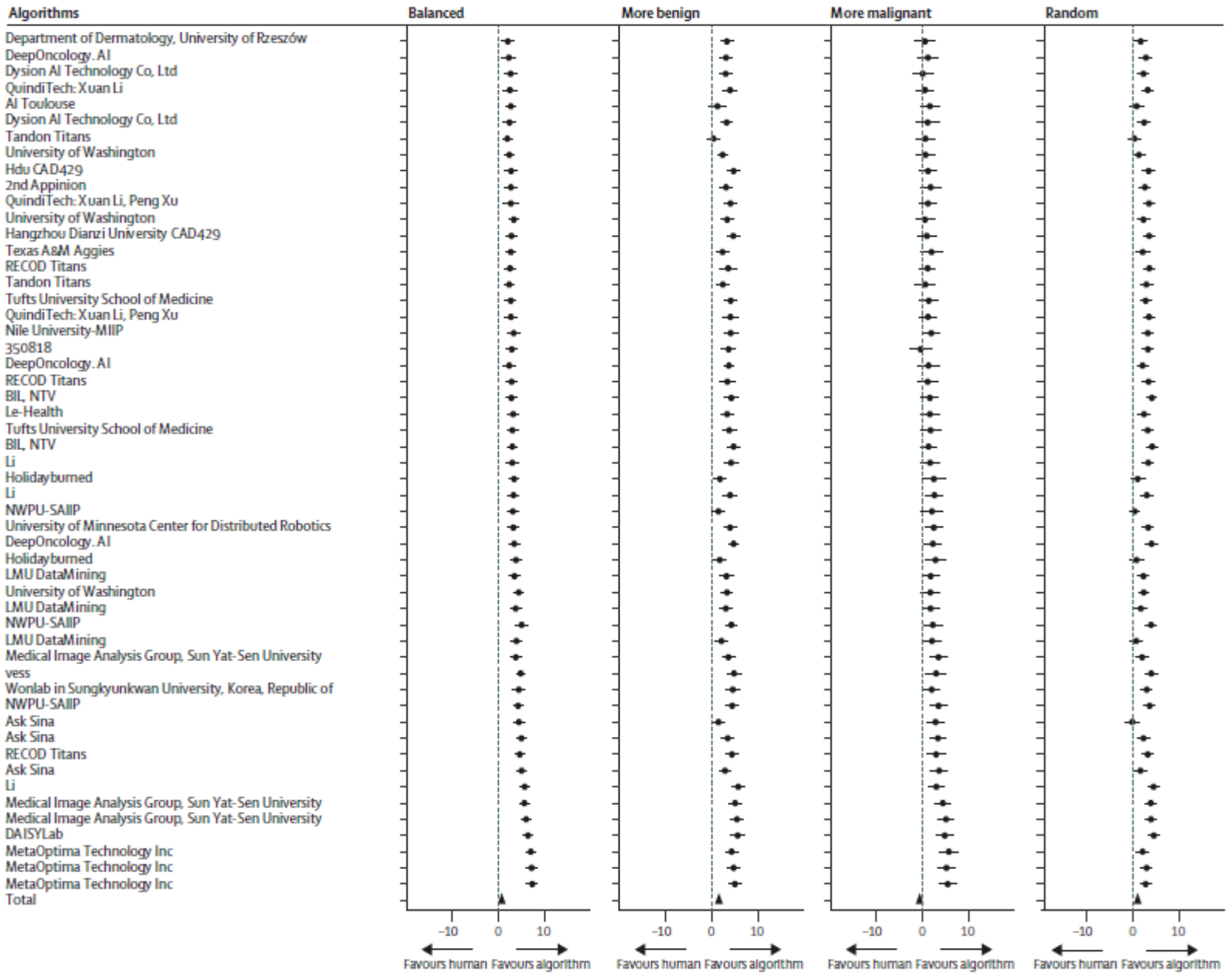

T-Mobile A
20:38
dermachallenge.meduniwien.ac.
Score
You 21:27 The Computer.
Oh no! The current state-of-the-art artificial intelligence system had 6 more correct answer(s).
Your time was 47 seconds.
Play again!
Home
Highscore

THE LANCET
Oncology

Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study

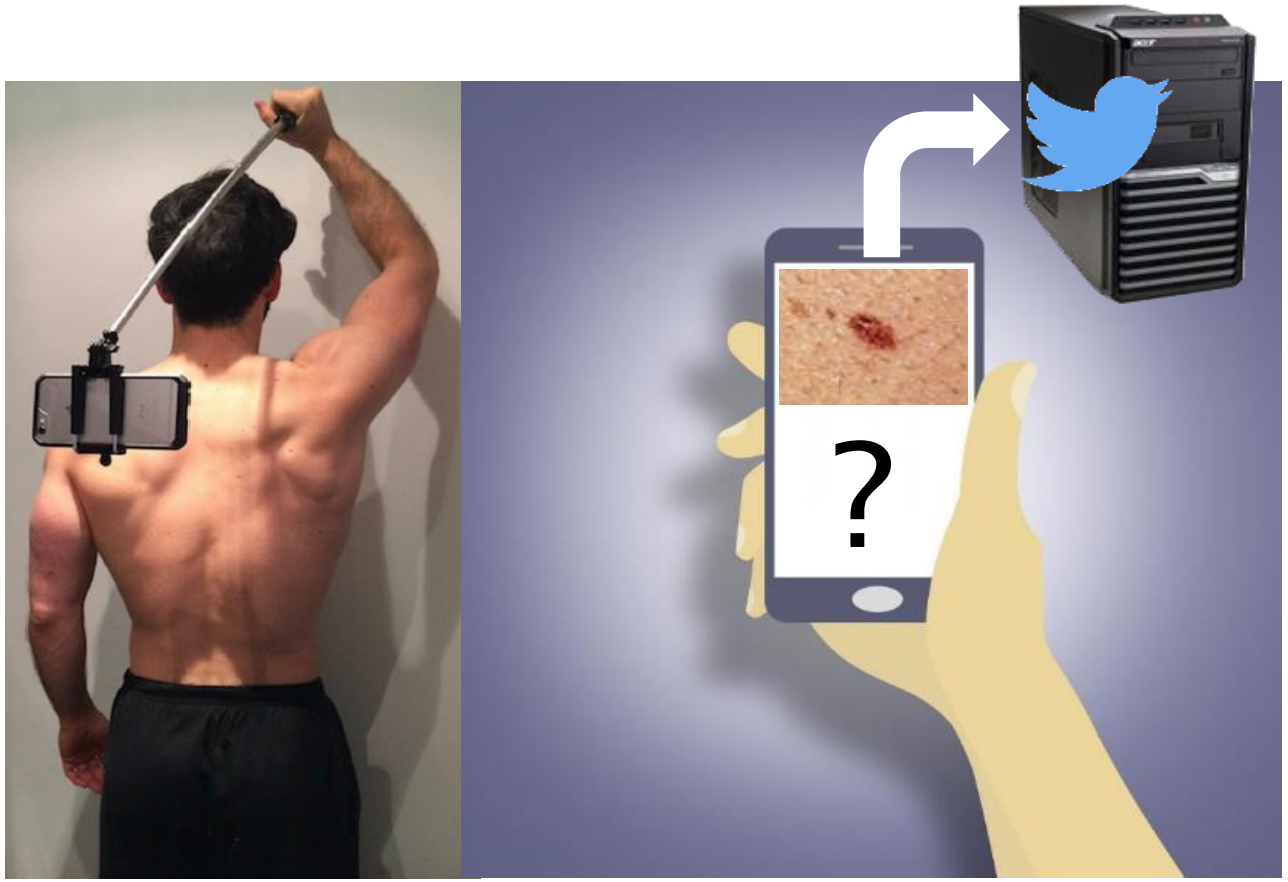
Philipp Tschandl, Noel Codella, Bengü Nisa Akay, Giuseppe Argenziano, Ralph P Braun, Horacio Cabo, David Gutman, Allan Halpern, Brian Helba, Rainer Hofmann-Wellenhof, Aimilios Lallas, Jan Lapins, Caterina Longo, Josep Malvehy, Michael A Marchetti, Ashfaq Marghoob, Scott Menzies, Amanda Oakley, John Paoli, Susana Puig, Christoph Rinner, Cliff Rosendahl, Alon Scope, Christoph Sinz, H Peter Soyer, Luc Thomas, Iris Zalaudek, Harald Kittler

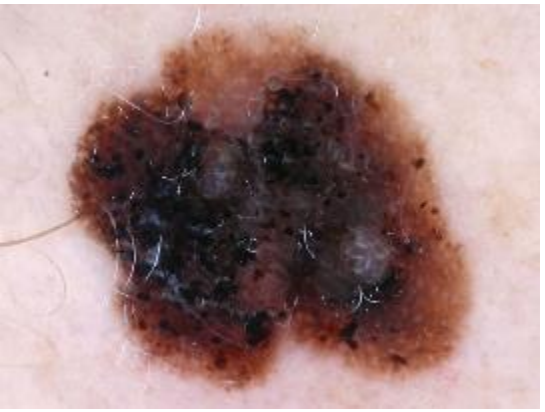



Algorithms
Balanced
More benign
More malignant
Random
Department of Dermatology, University of Rzeszów
DeepOncology.AI
Dysion AI Technology Co, Ltd
QuindiTech: Xuan Li
AI Toulouse
Dysion AI Technology Co, Ltd
Tandon Titans
University of Washington
Hdu CAD429
2nd Appinion
QuindiTech: Xuan Li, Peng Xu
University of Washington
Hangzhou Dianzi University CAD429
Texas A&M Aggies
RECOD Titans
Tandon Titans
Tufts University School of Medicine
QuindiTech: Xuan Li, Peng Xu
Nile University-MIIP
350818
DeepOncology.AI
RECOD Titans
BIL, NTV
Le-Health
Tufts University School of Medicine
BIL, NTV
Li
Holidayburned
Li
NWPU-SAIIP
University of Minnesota Center for Distributed Robotics
DeepOncology.AI
Holidayburned
LMU DataMining
University of Washington
LMU DataMining
NWPU-SAIIP
LMU DataMining
Medical Image Analysis Group, Sun Yat-Sen University
vess
Wonlab in Sungkyunkwan University, Korea, Republic of
NWPU-SAIIP
Ask Sina
Ask Sina
RECOD Titans
Ask Sina
Li
Medical Image Analysis Group, Sun Yat-Sen University
Medical Image Analysis Group, Sun Yat-Sen University
DAISYLab
MetaOptima Technology Inc
MetaOptima Technology Inc
MetaOptima Technology Inc
Total
−10
0
10
Favours human
Favours algorithm

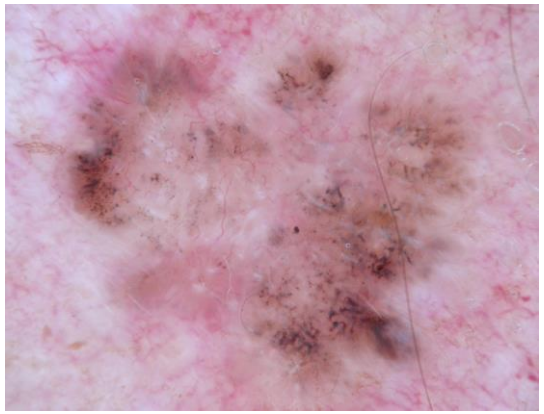
The importance of domain-specific knowledge: Human-assisted AI and AI-assisted humans

Personalized treatment decisions?
This is an AI-hard problem

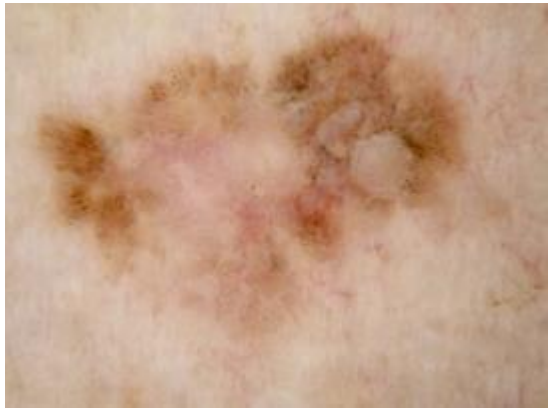




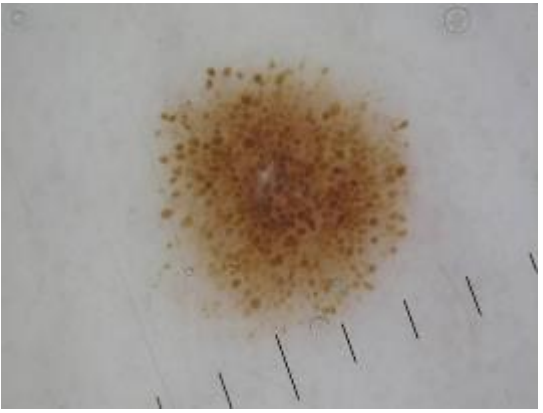
Melanoma



Basal cell carcinoma



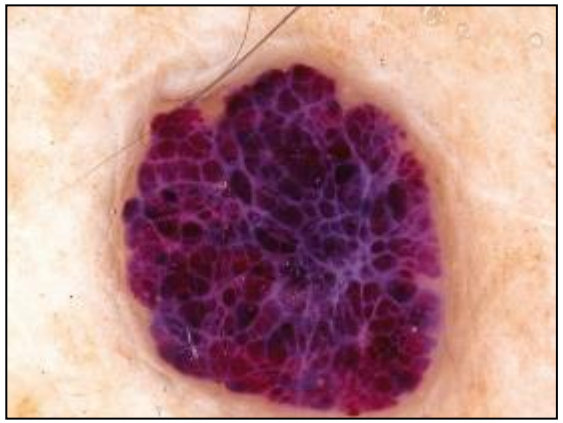
Squamous cell carcinoma



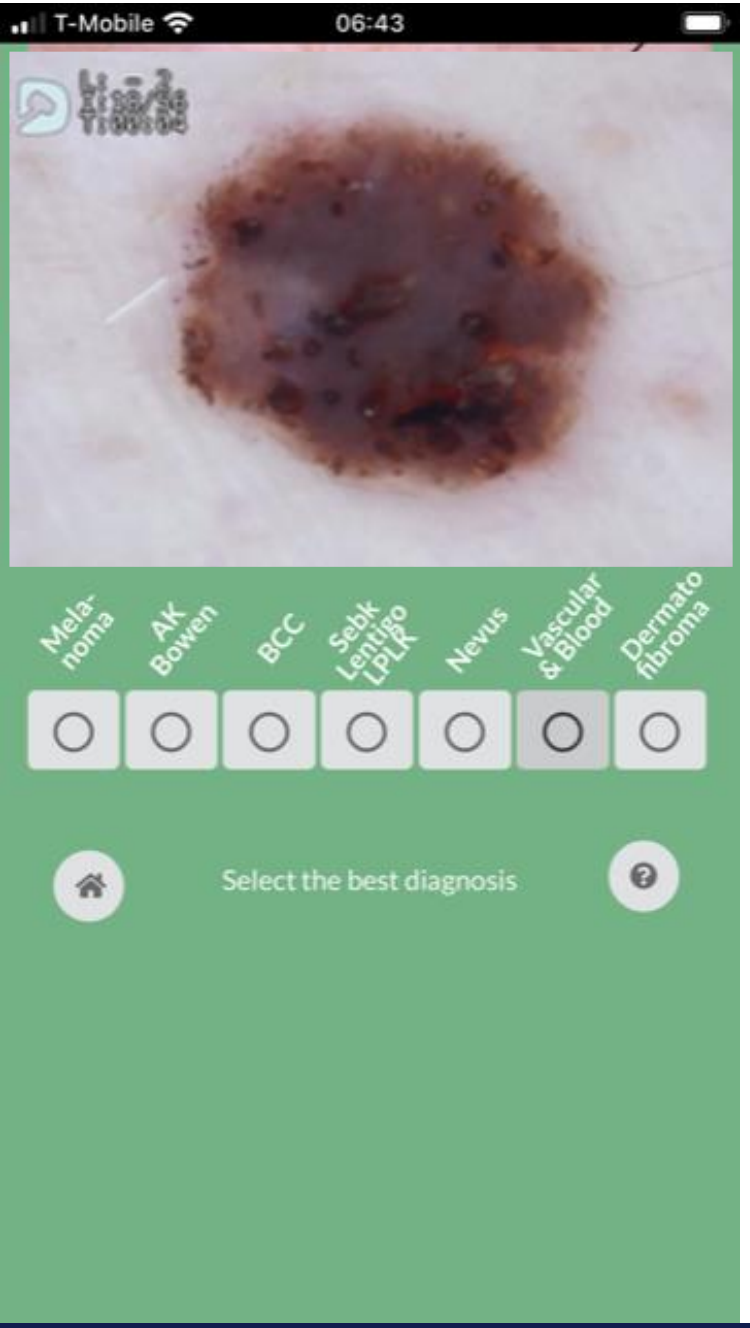
Nevus

Benign keratoses

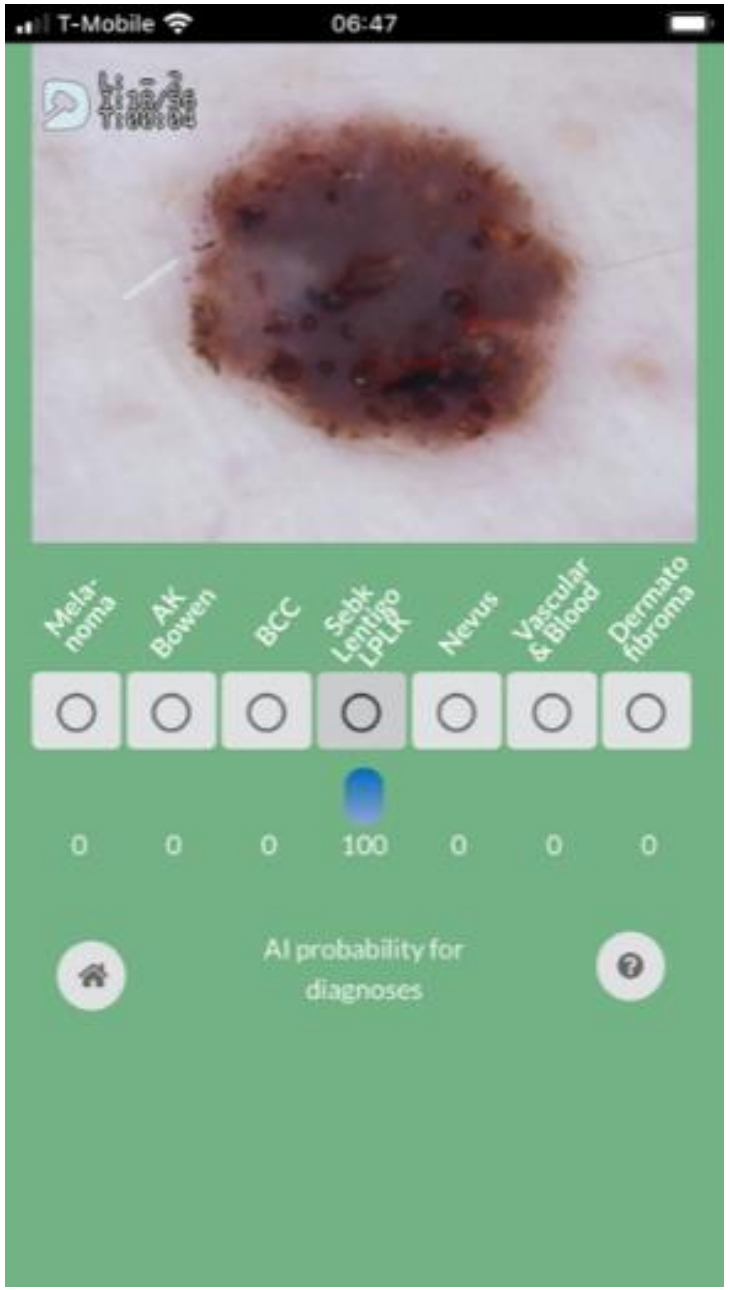
Dermatofibroma



Angioma

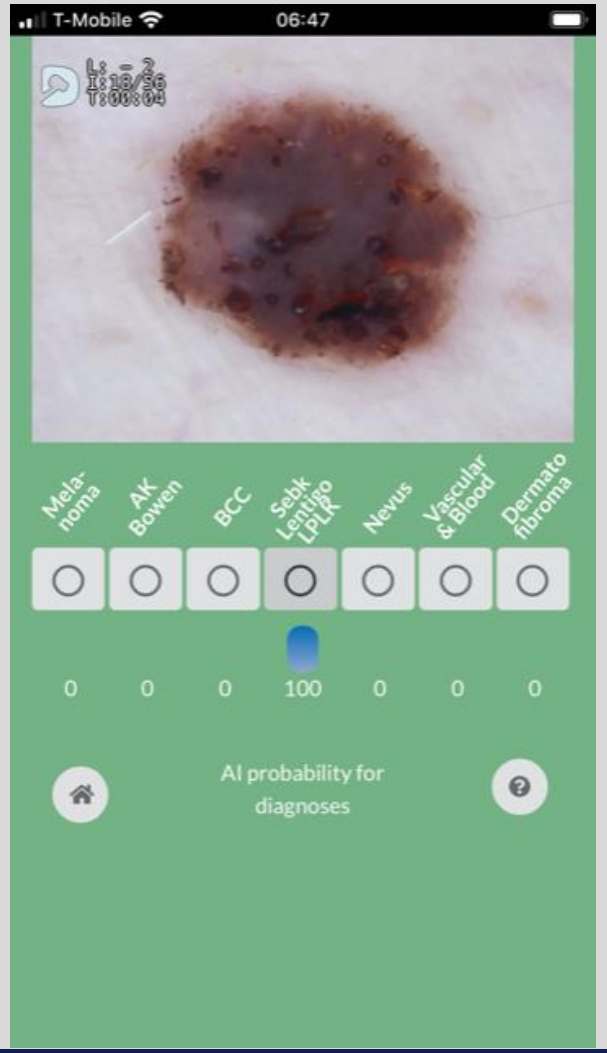


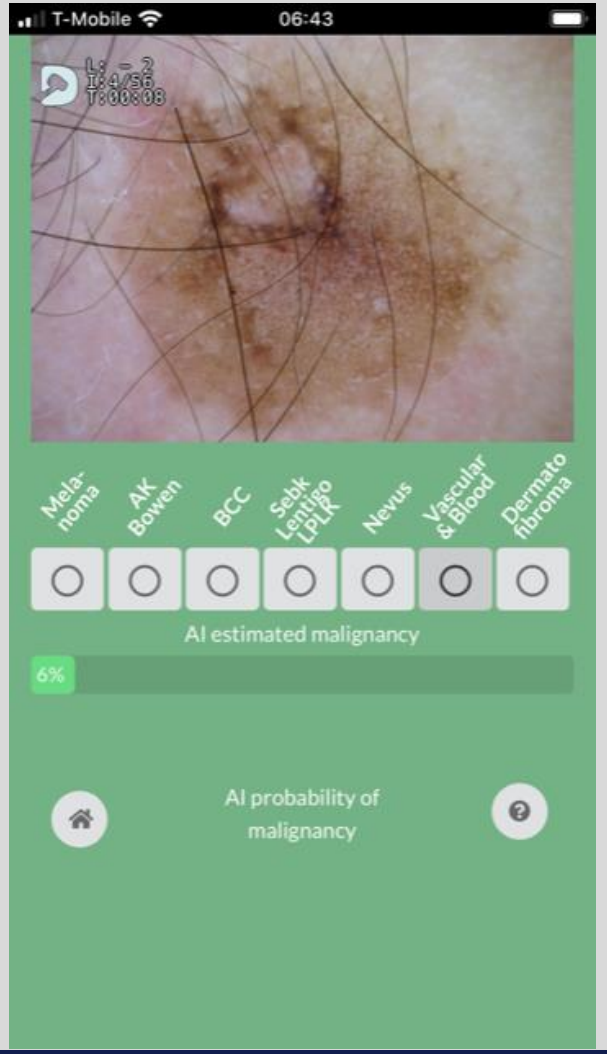
Without AI-support

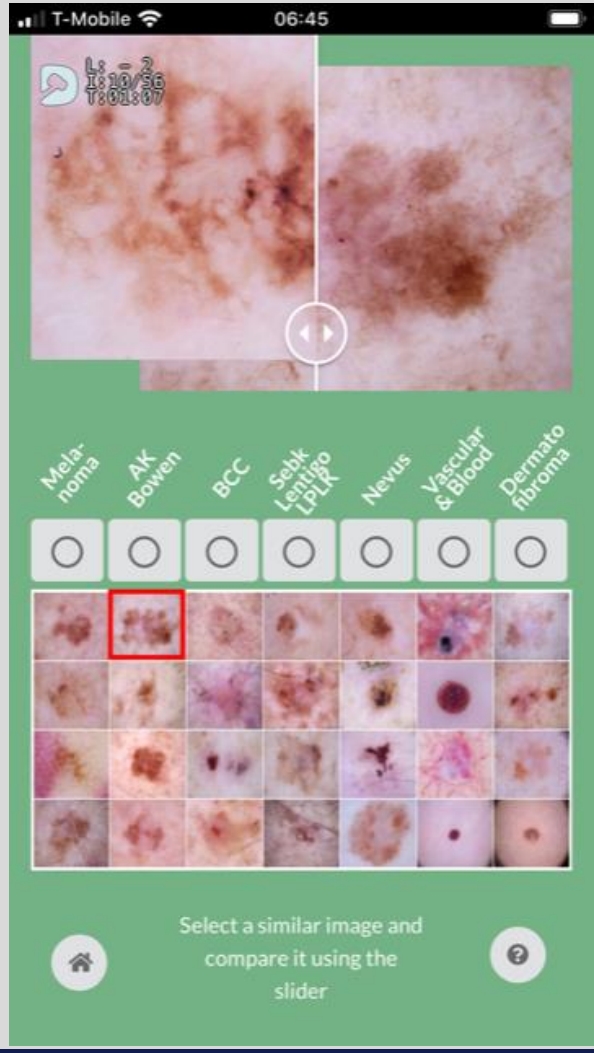


With AI-support

The importance of domain-specific knowledge: Human-AI collaboration

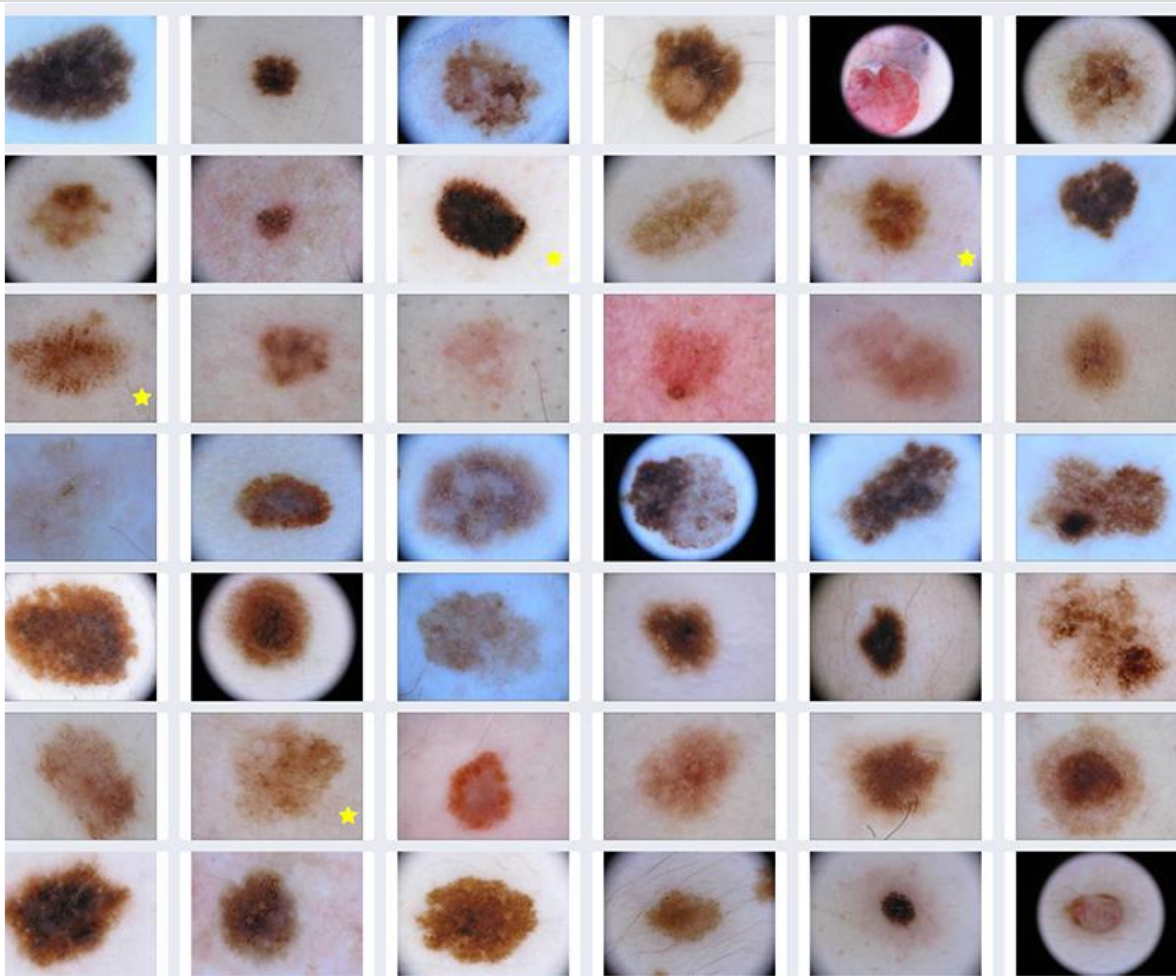






Content based image retrieval:

Tschandl et al, Br J Dermatol 2018, **epub 2018 Sep 12**



ypsono

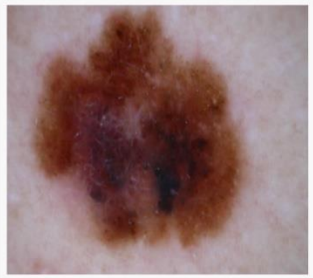
HAM10000 Dataset

Beat the Machine

ViDIR Group

Upload a dermatoscopic image, and use our neural network to **search for similar images** in the HAM10000 dataset!

Choose a dermatoscopic image...



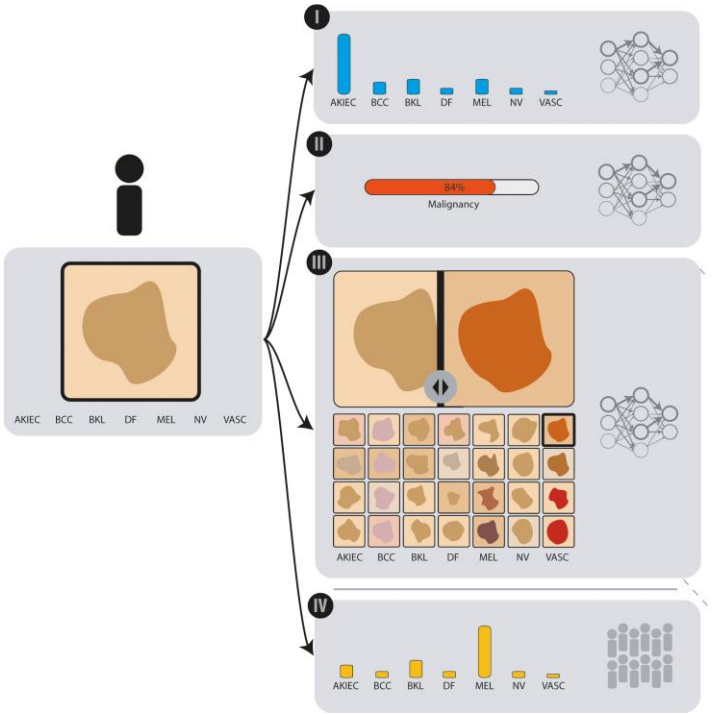
Upload a dermatoscopic image, and use our neural network to **search for similar images** in the HAM10000 dataset!

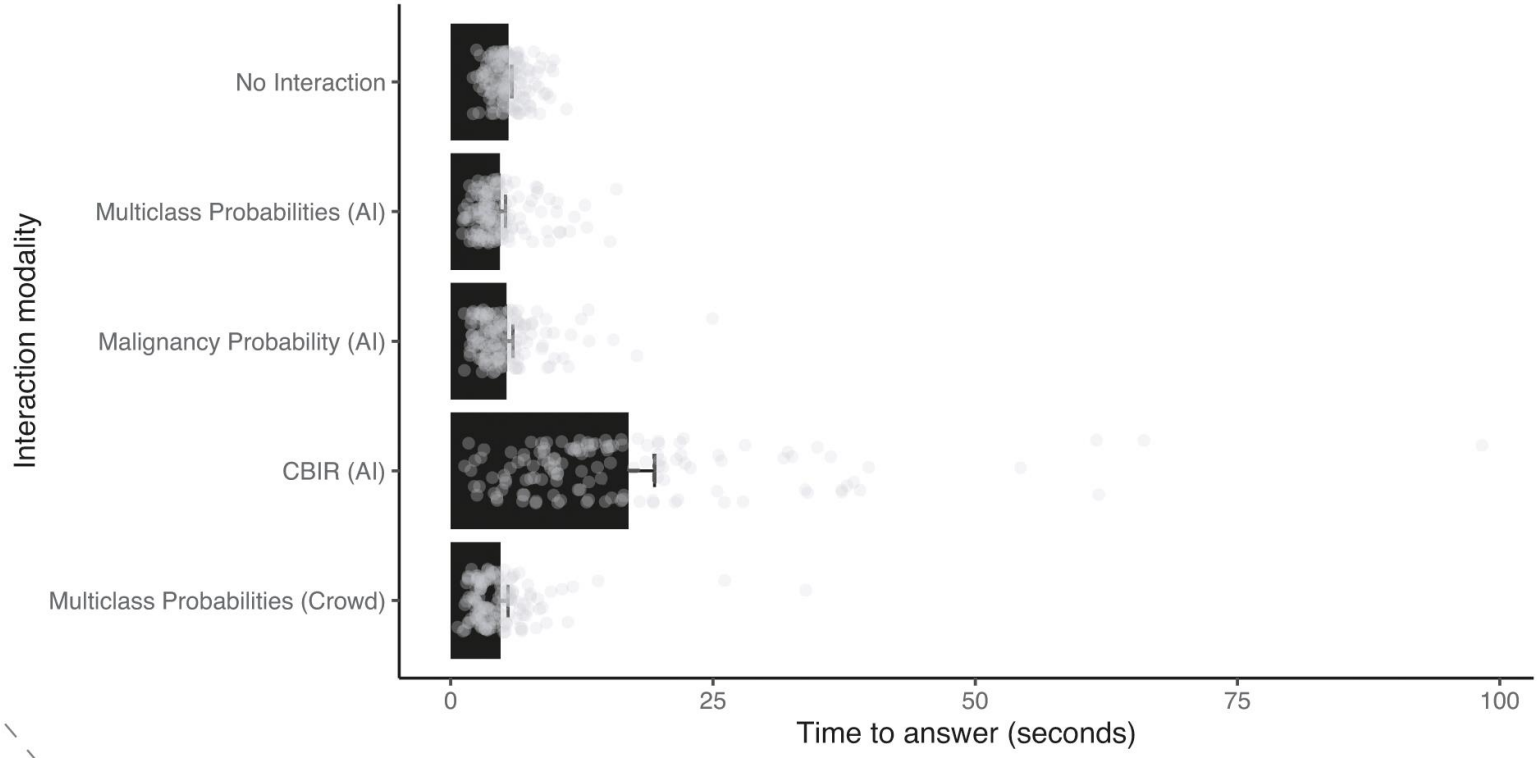
Choose a dermatoscopic image...

nature medicine

Human–computer collaboration for skin cancer recognition

Philipp Tschandl[1,17], Christoph Rinner[2,17], Zoe Apalla[3], Giuseppe Argenziano[4], Noel Codella[5], Allan Halpern[6], Monika Janda[7], Aimilios Lallas[3], Caterina Longo[8,9], Josep Malvehy[10,11], John Paoli[12,13], Susana Puig[10,11], Cliff Rosendahl[14], H. Peter Soyer[15], Iris Zalaudek[16] and Harald Kittler[1]

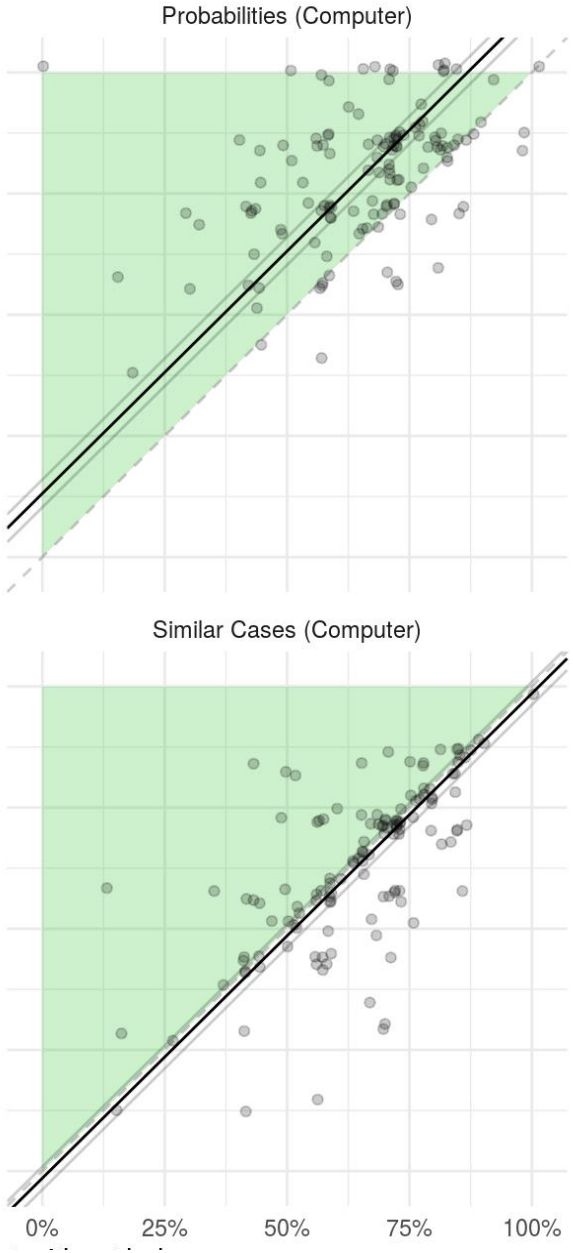




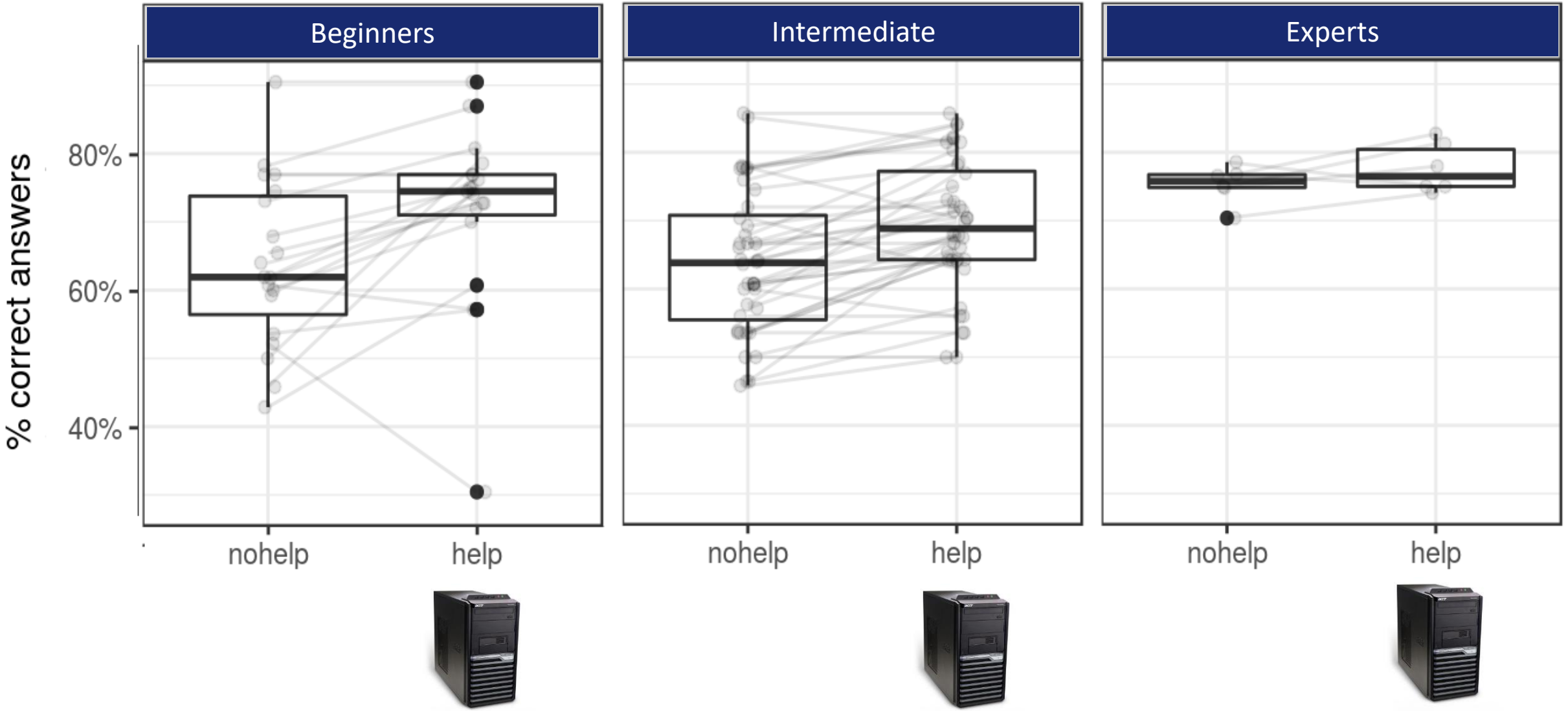
nature medicine

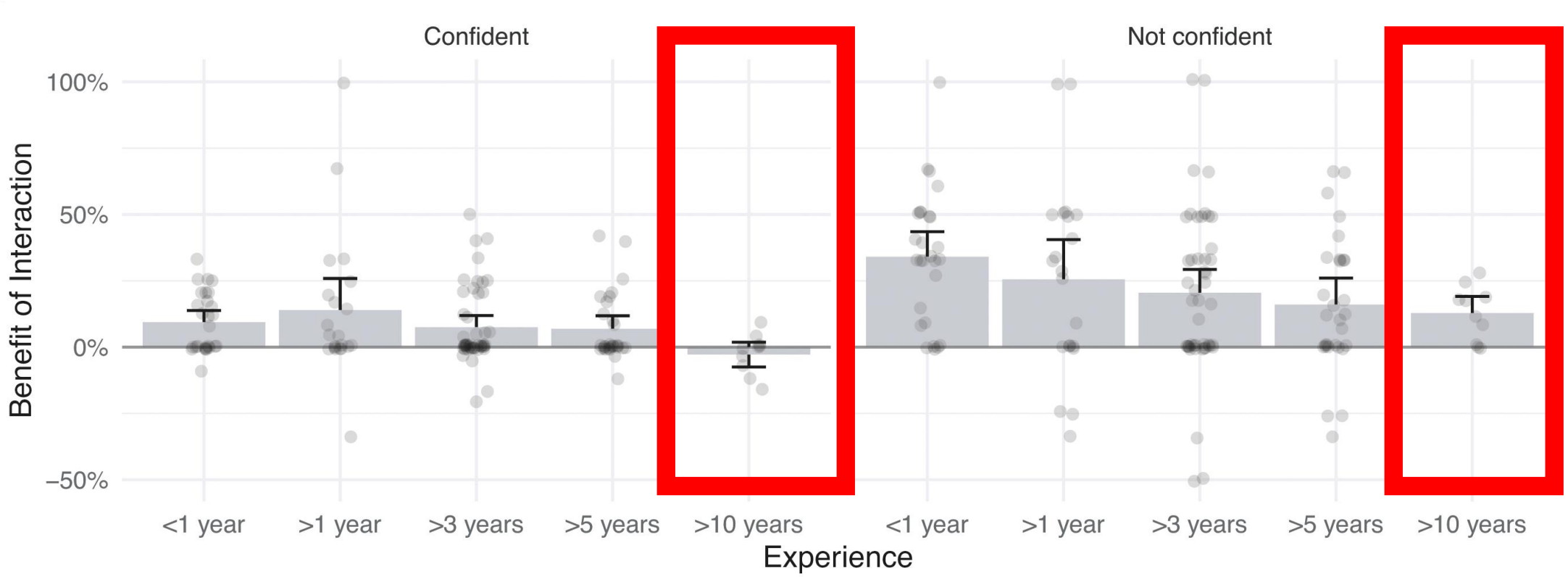
Human–computer collaboration for skin cancer recognition

Philipp Tschandl[1,17], Christoph Rinner[2,17], Zoe Apalla[3], Giuseppe Argenziano[4], Noel Codella[5], Allan Halpern[6], Monika Janda[7], Aimilios Lallas[3], Caterina Longo[8,9], Josep Malvehy[10,11], John Paoli[12,13], Susana Puig[10,11], Cliff Rosendahl[14], H. Peter Soyer[15], Iris Zalaudek[16] and Harald Kittler[1]✉



nature medicine

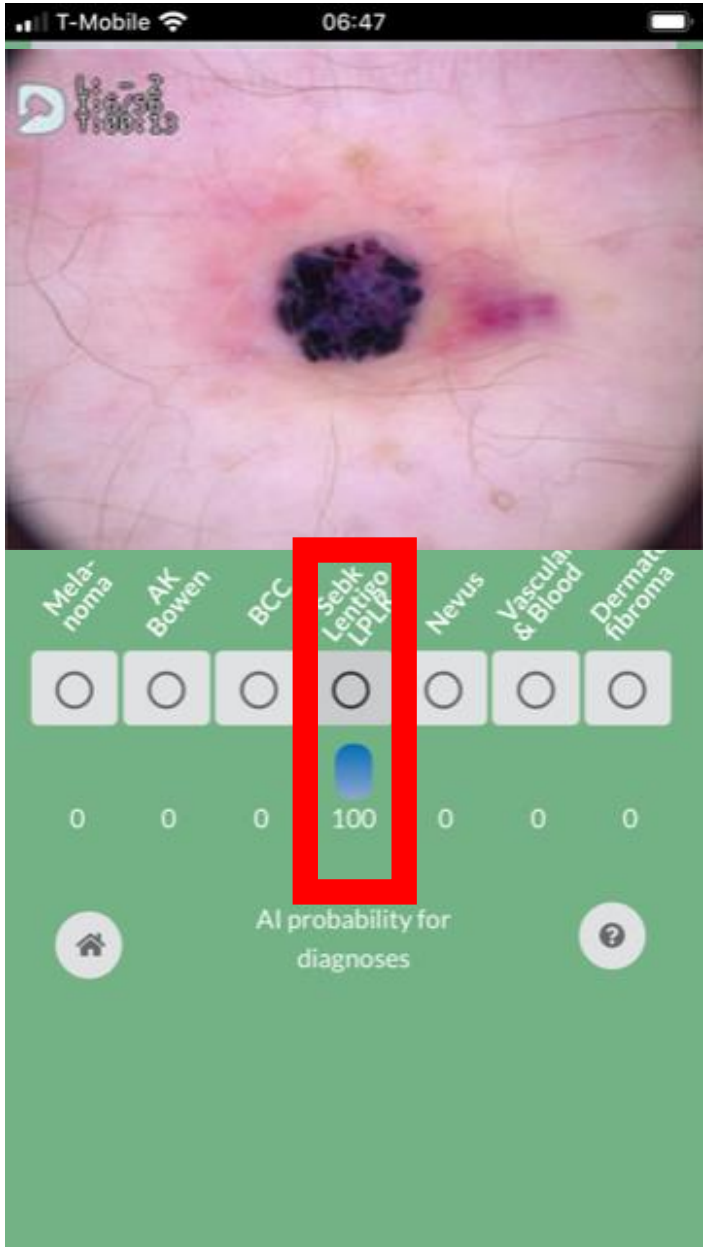



Confident
Not confident
100%
50%
0%
−50%
Benefit of Interaction
<1 year
>1 year
>3 years
>5 years
>10 years
<1 year
>1 year
>3 years
>5 years
>10 years
Experience

nature medicine

Human–computer collaboration for skin cancer recognition

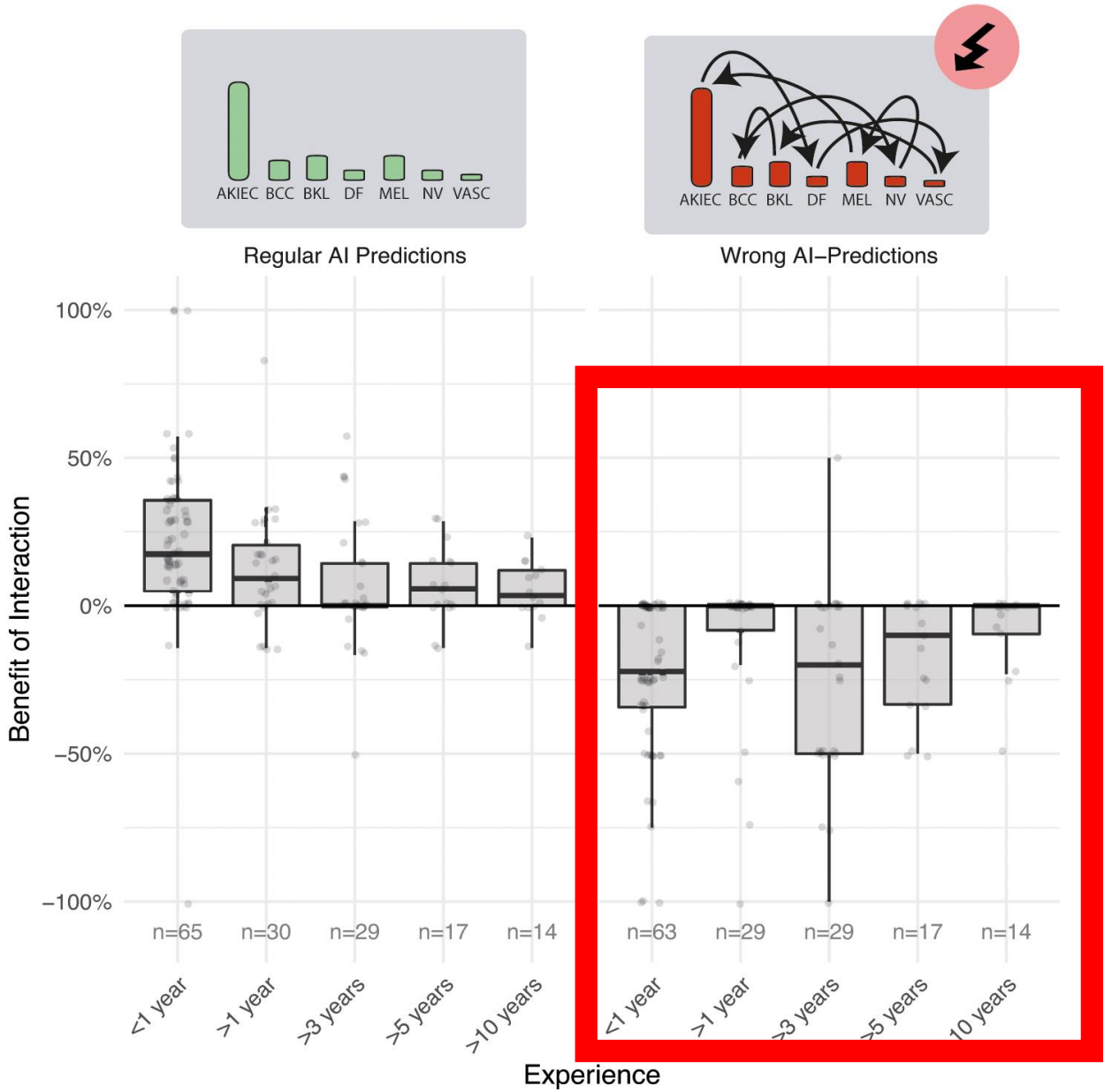
Philipp Tschandl[1,17], Christoph Rinner[2,17], Zoe Apalla[3], Giuseppe Argenziano[4], Noel Codella[5], Allan Halpern[6], Monika Janda[7], Aimilios Lallas[3], Caterina Longo[8,9], Josep Malvehy[10,11], John Paoli[12,13], Susana Puig[10,11], Cliff Rosendahl[14], H. Peter Soyer[15], Iris Zalaudek[16] and Harald Kittler[1]✉



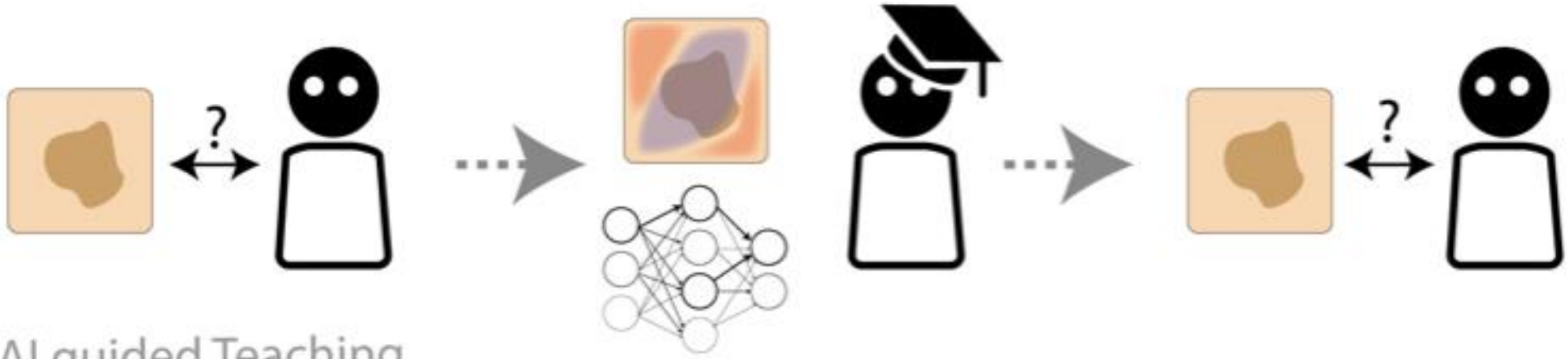
nature medicine

Human–computer collaboration for skin cancer recognition

Philipp Tschandl[1,17], Christoph Rinner[2,17], Zoe Apalla[3], Giuseppe Argenziano[4], Noel Codella[5], Allan Halpern[6], Monika Janda[7], Aimilios Lallas[3], Caterina Longo[8,9], Josep Malvehy[10,11], John Paoli[12,13], Susana Puig[10,11], Cliff Rosendahl[14], H. Peter Soyer[15], Iris Zalaudek[16] and Harald Kittler[1]



Knowledge Transfer

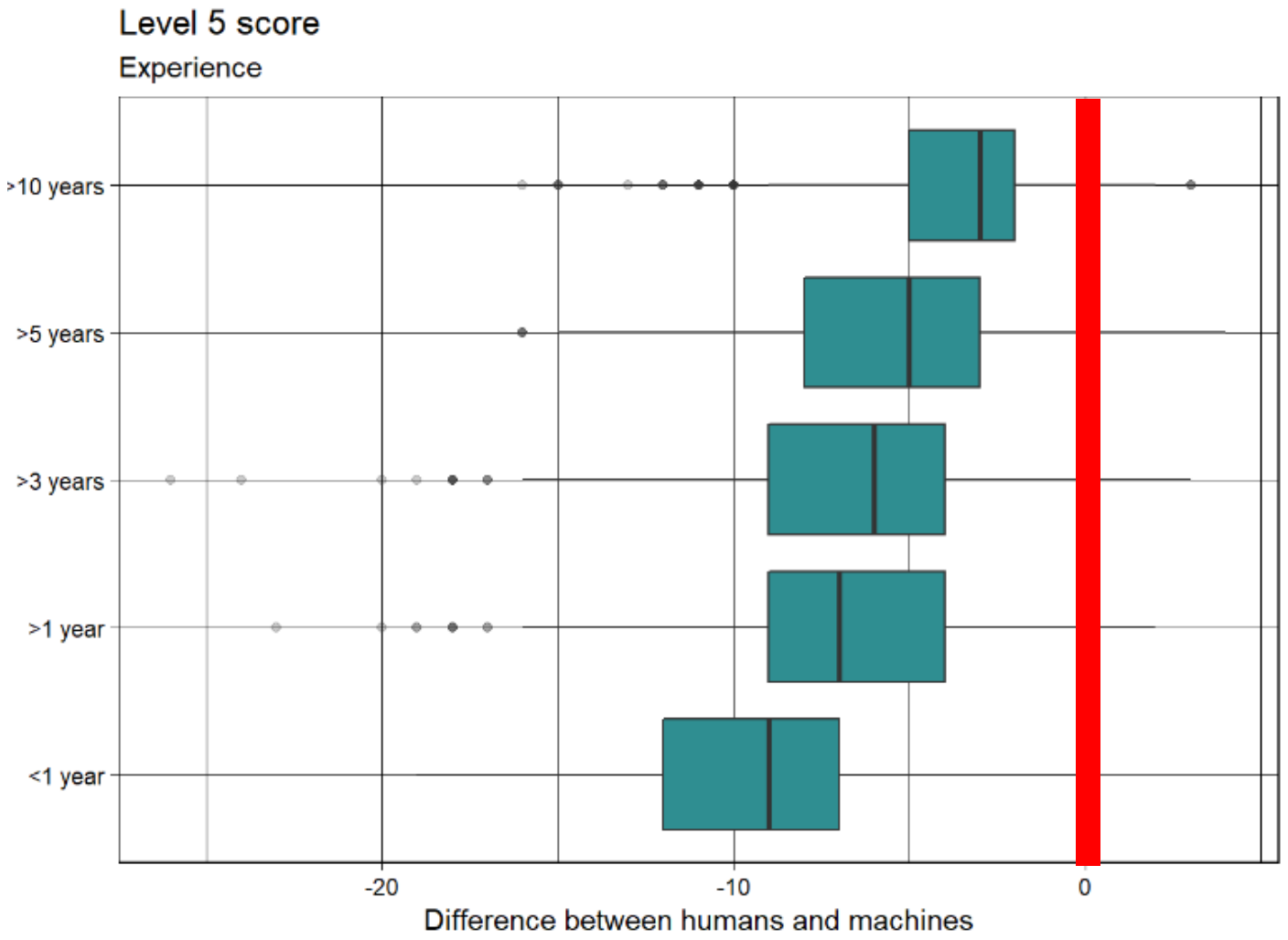


AI guided Teaching

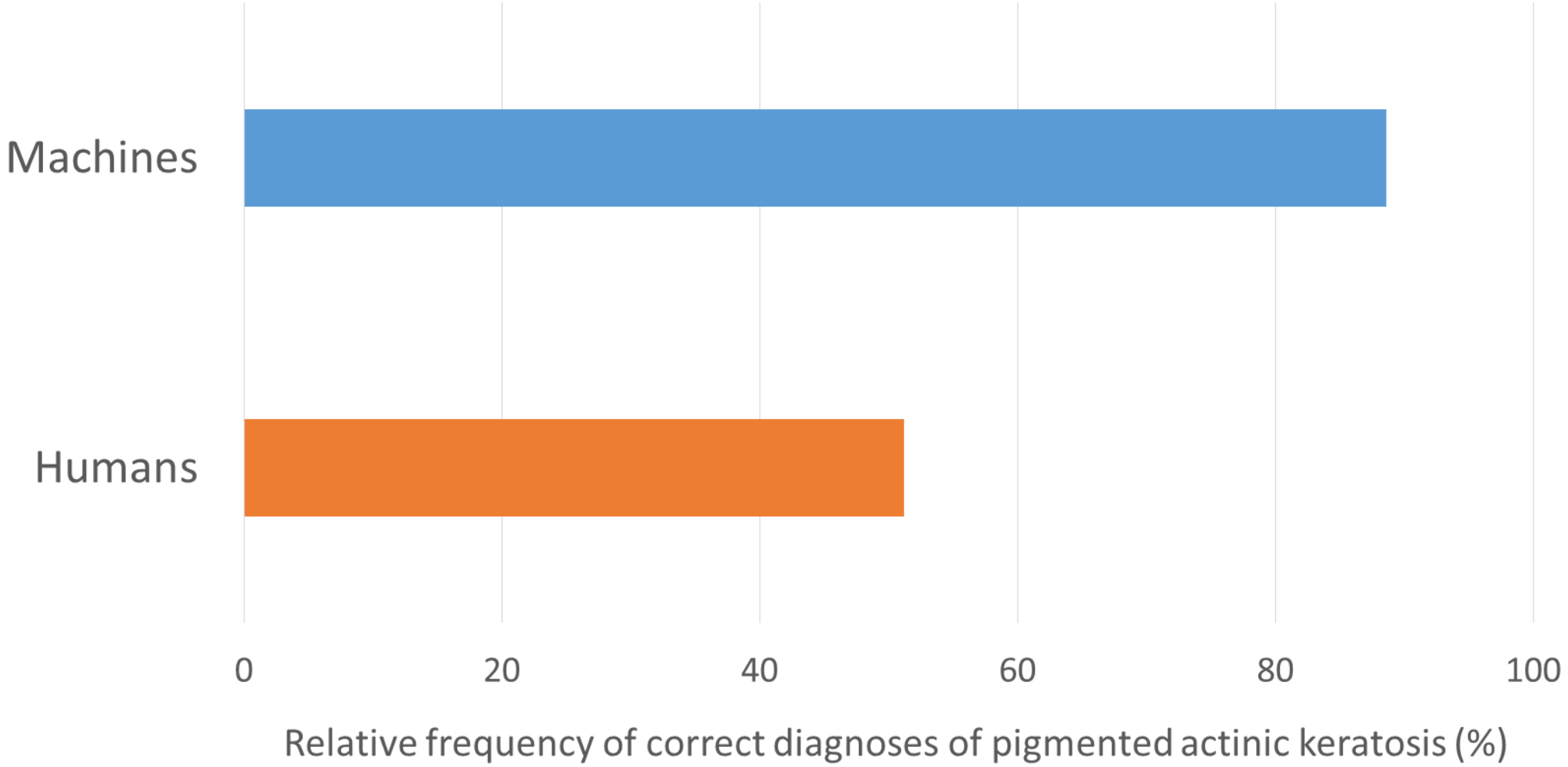
THE LANCET
Oncology

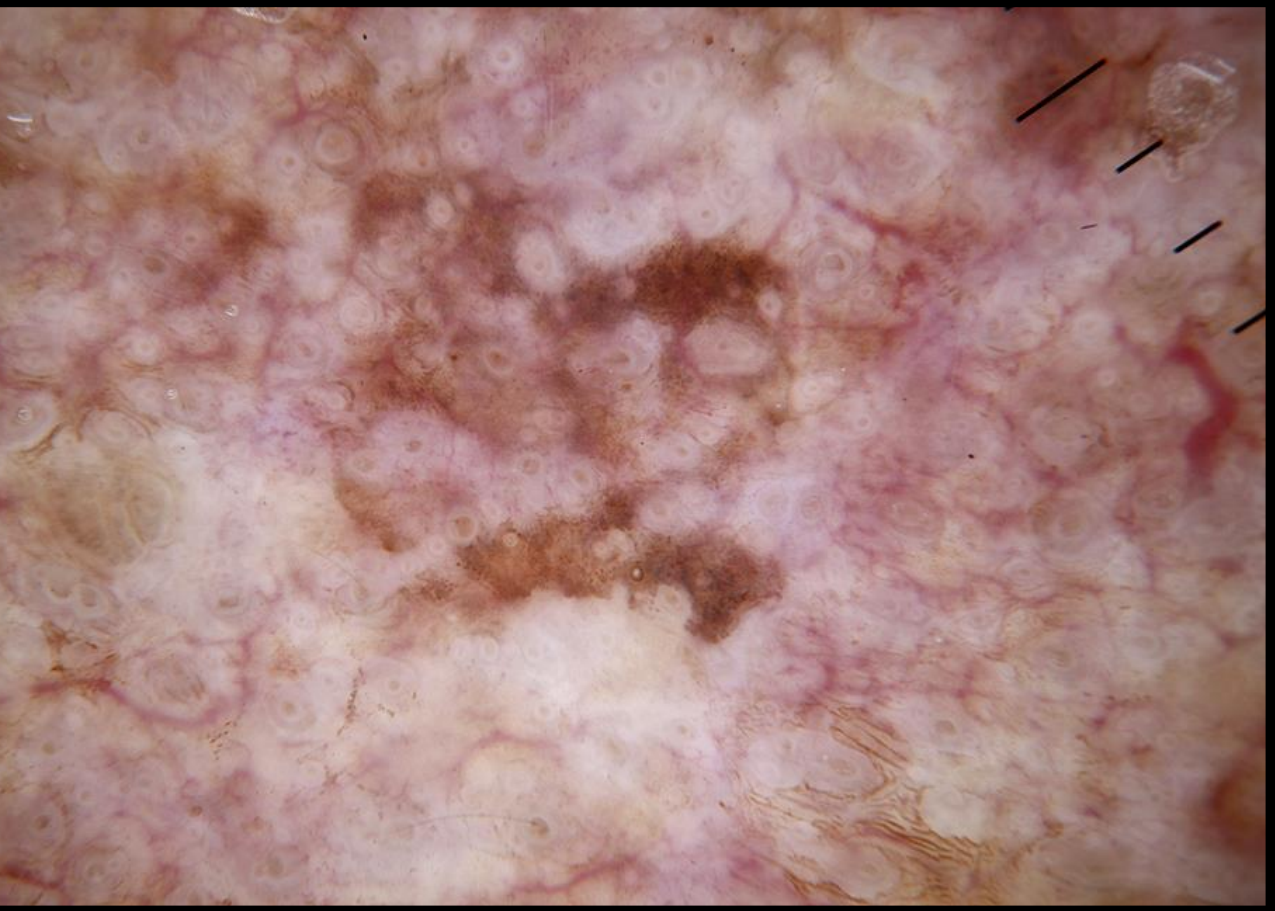
Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study

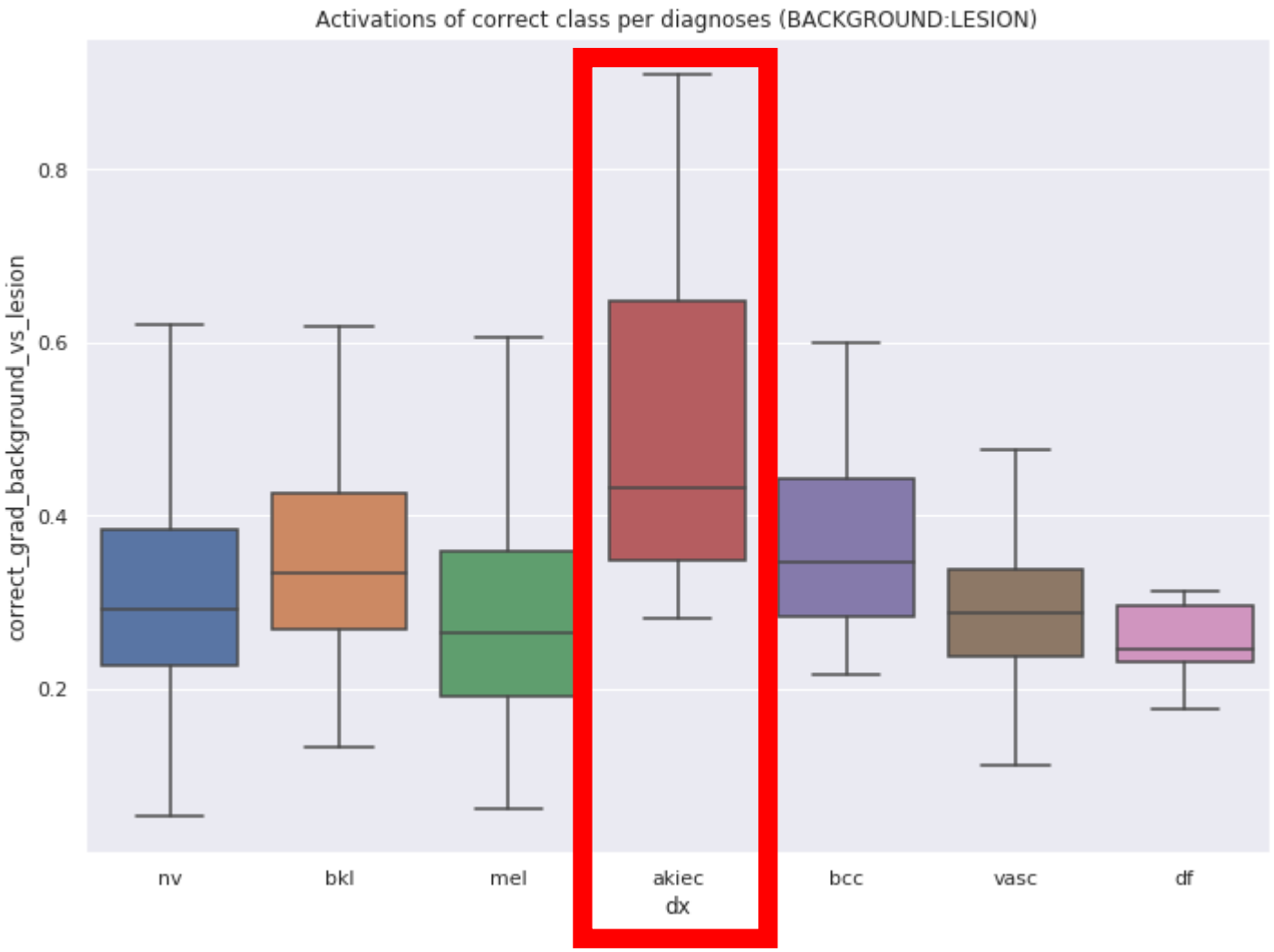
Philipp Tschandl, Noel Codella, Bengü Nisa Akay, Giuseppe Argenziano, Ralph P Braun, Horacio Cabo, David Gutman, Allan Halpern, Brian Helba, Rainer Hofmann-Wellenhof, Aimilios Lallas, Jan Lapins, Caterina Longo, Josep Malvehy, Michael A Marchetti, Ashfaq Marghoob, Scott Menzies, Amanda Oakley, John Paoli, Susana Puig, Christoph Rinner, Cliff Rosendahl, Alon Scope, Christoph Sinz, H Peter Soyer, Luc Thomas, Iris Zalaudek, Harald Kittler

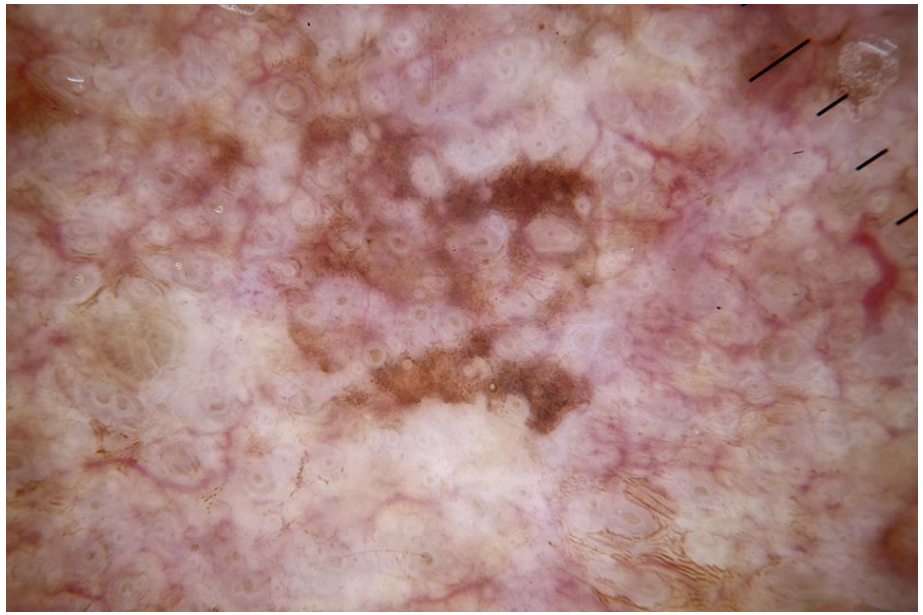


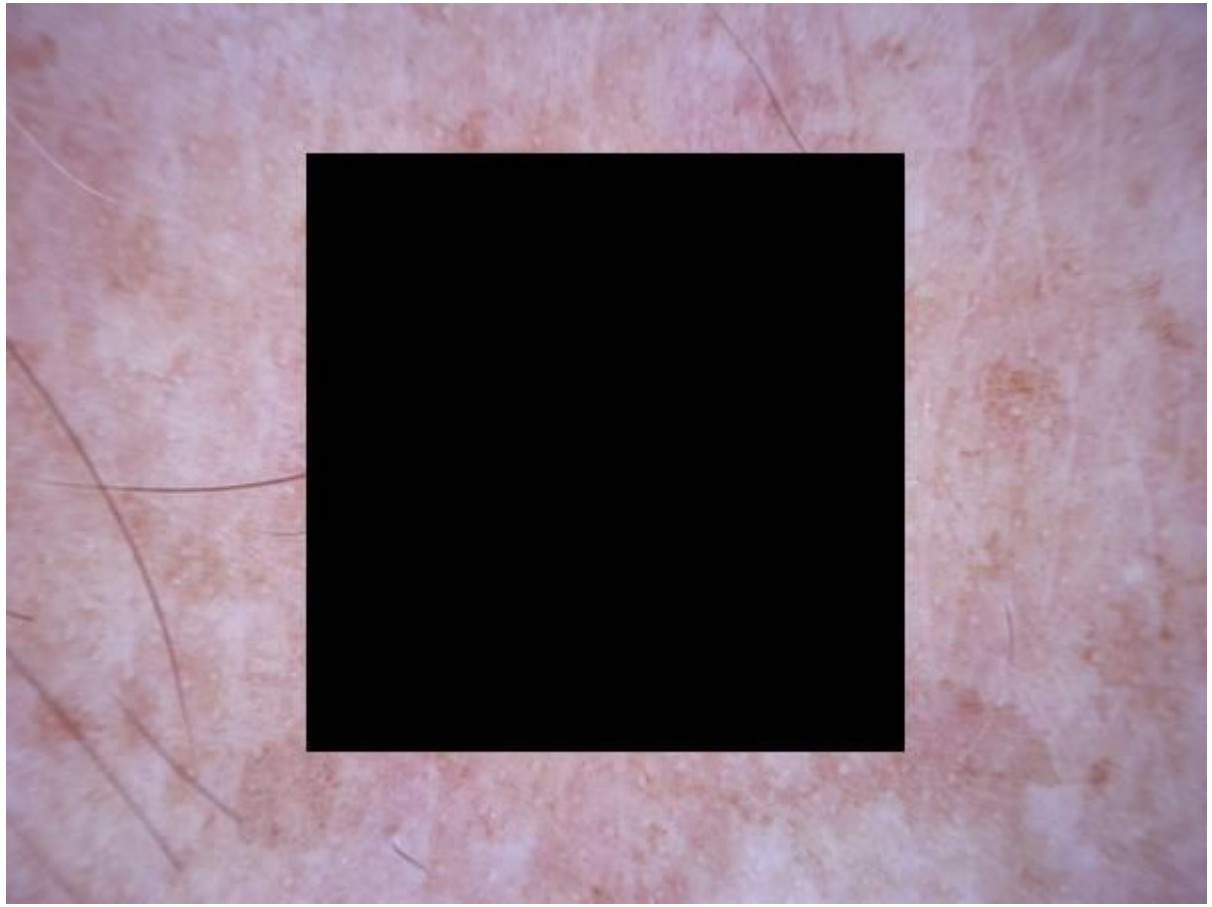
Why did humans loose?

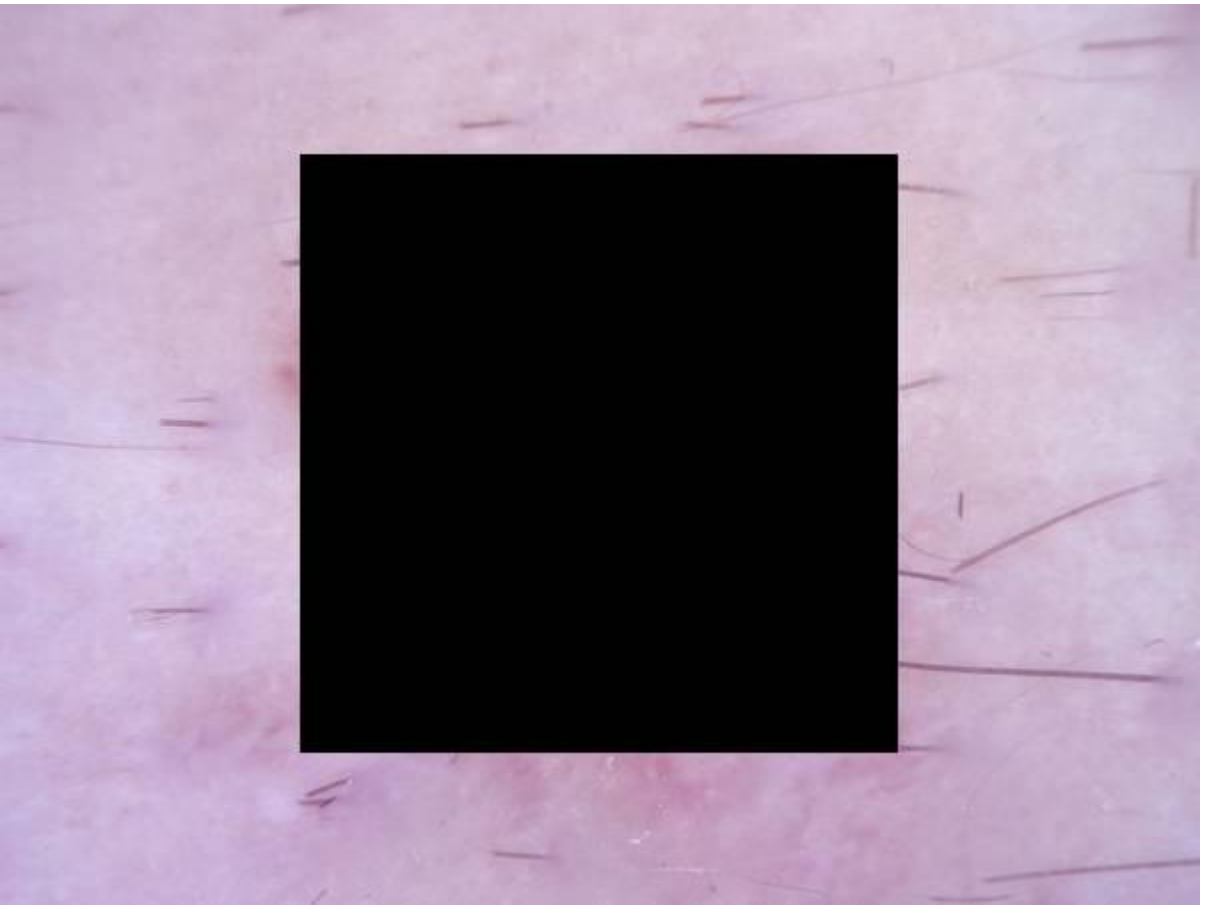





Activations of correct class per diagnoses (BACKGROUND:LESION)
correct_grad_background_vs_lesion
0.8
0.6
0.4
0.2
nv
bkl
mel
akiec
bcc
vasc
df
dx

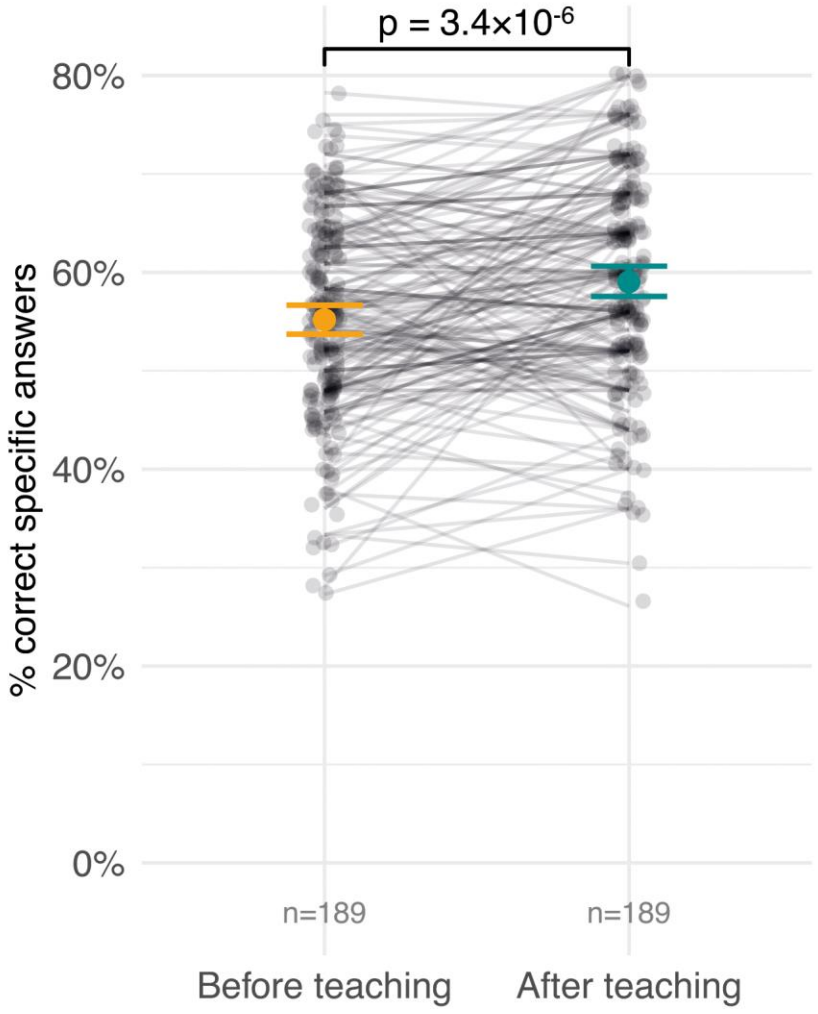


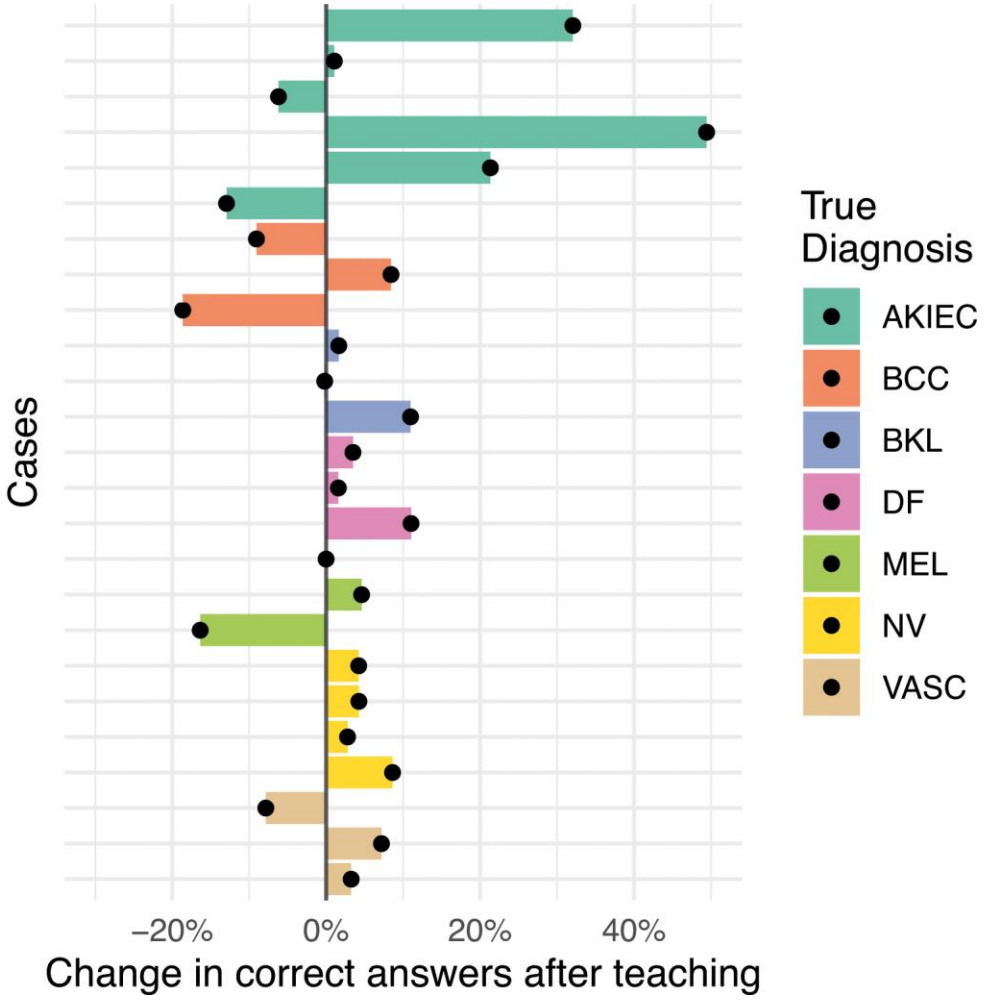




Do humans improve if we tell them?

E | Explainable AI - Teaching





Domain specific knowledge was necessary to translate the results of class activation maps into a human understandable concept

The importance of domain-specific knowledge:

Problem formulation

Data collection

Data preprocessing

Modeling

Interpretation of results

Understand and predict limitations and exceptions

Make the algorithm fit for intended use

Technical skills alone are not enough.

Important parts of the problem are not captured in the data.

Applied computer science

The importance of domain-specific knowledge:

Visit a health care facility to get a feeling for the problem from the physician point of view
(3 weeks is enough)

ViDIR - Vienna Dermatologic Imaging Research Group
harald.kittler@meduniwien.ac.at
philipp.tschandl@meduniwien.ac.at