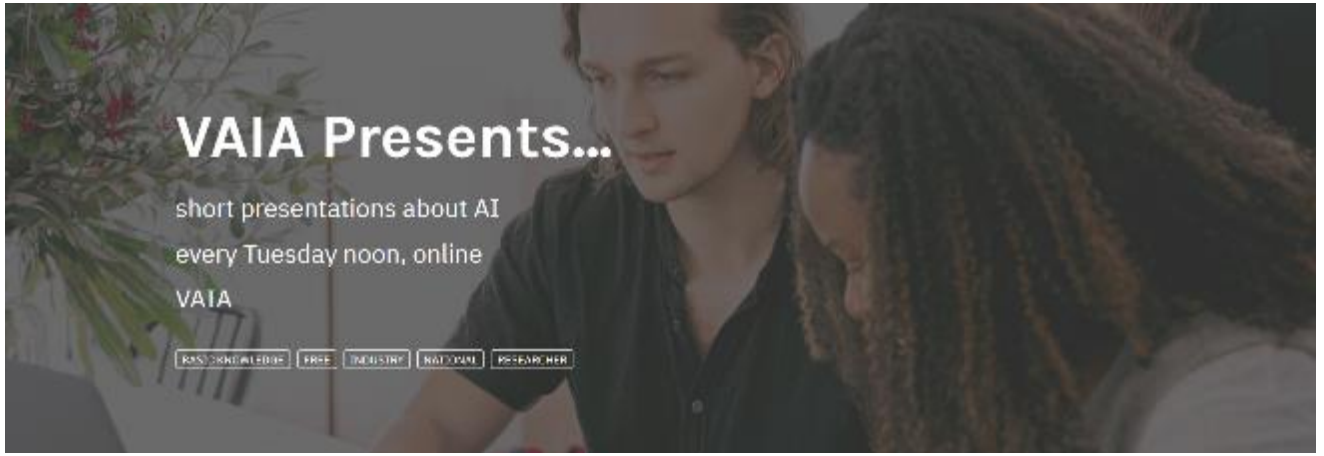
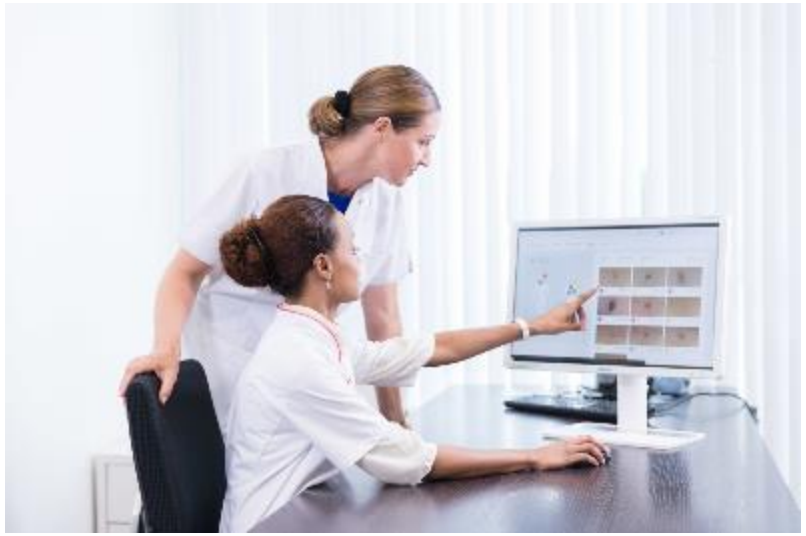
A collaboration of all the universities and universities for applied sciences in Flanders



Artificial intelligence in healthcare: opportunities, learnings and challenges

Tom Kimpe
Jonas De Vylder
Bart Diricx

Barco Healthcare, Belgium
Dec 21st 2021



Agenda

- Introduction
  - Barco Healthcare
  - Demetra
- Adoption of AI in various industries
- AI in Healthcare
- Challenges
  - Criticality
  - Complex ecosystem
  - Regulations
  - Data (bias)
  - Validation methods
  - Changing AI algorithms postmarket
- Towards broad AI use in healthcare
- Conclusions
- Q&A

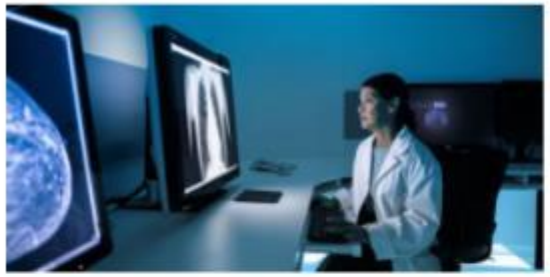
Introduction

ENTERTAINMENT

HEALTHCARE

ENTERPRISE

Barco Healthcare - introduction



Radiology

Mammography

Dermatology

Surgery

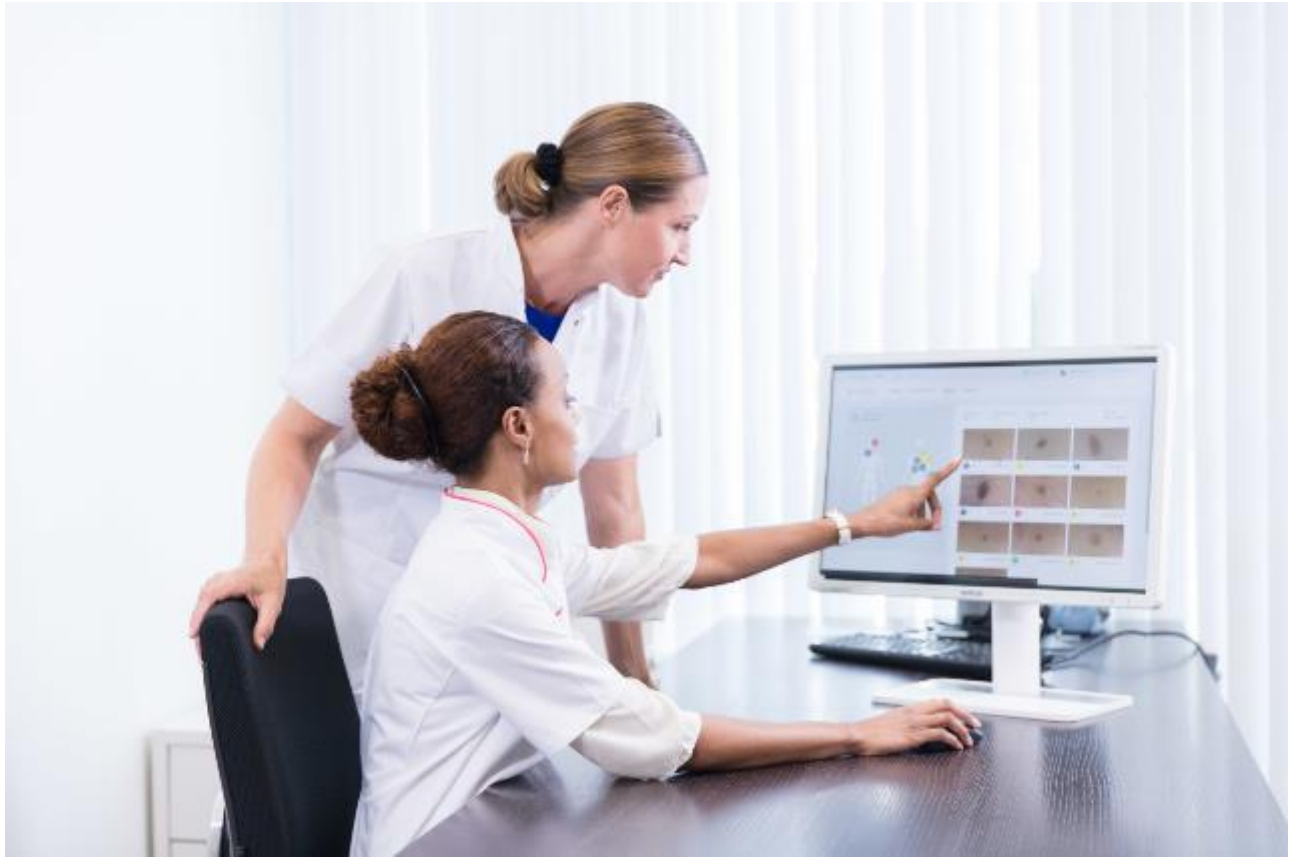
Modality

Oncology

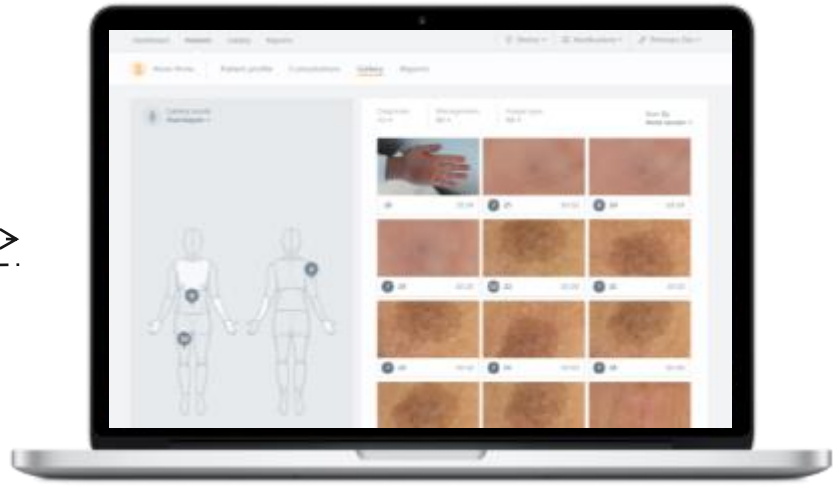
Pathology

Dentistry

Barco Demetra



Barco Demetra offers an easy-to-use and all-in-one skin imaging solution



Demetra Scope + **Demetra Cloud** + **Demetra Web**

https://www.barco.com/en/page/healthcare/demetra

Adoption of AI

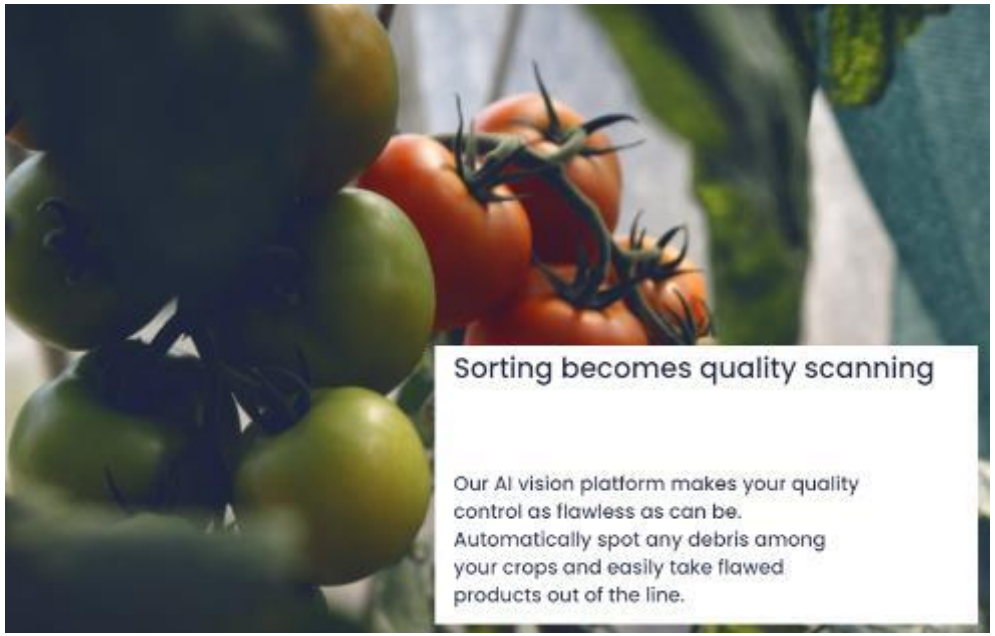
Adoption of AI

Mercedes Beats Tesla to Hands-Free Driving on the Autobahn

By William Wilkes +Get Alerts
9 december 2021 12:30 CET *Updated on 9 december 2021 13:09 CET*

Source: https://www.bloomberg.com/news/articles/2021-12-09/mercedes-beats-tesla-to-hands-free-driving-on-german-autobahn

ROBOVISION



Source: https://robovision.ai/agri-food/

Adoption of AI

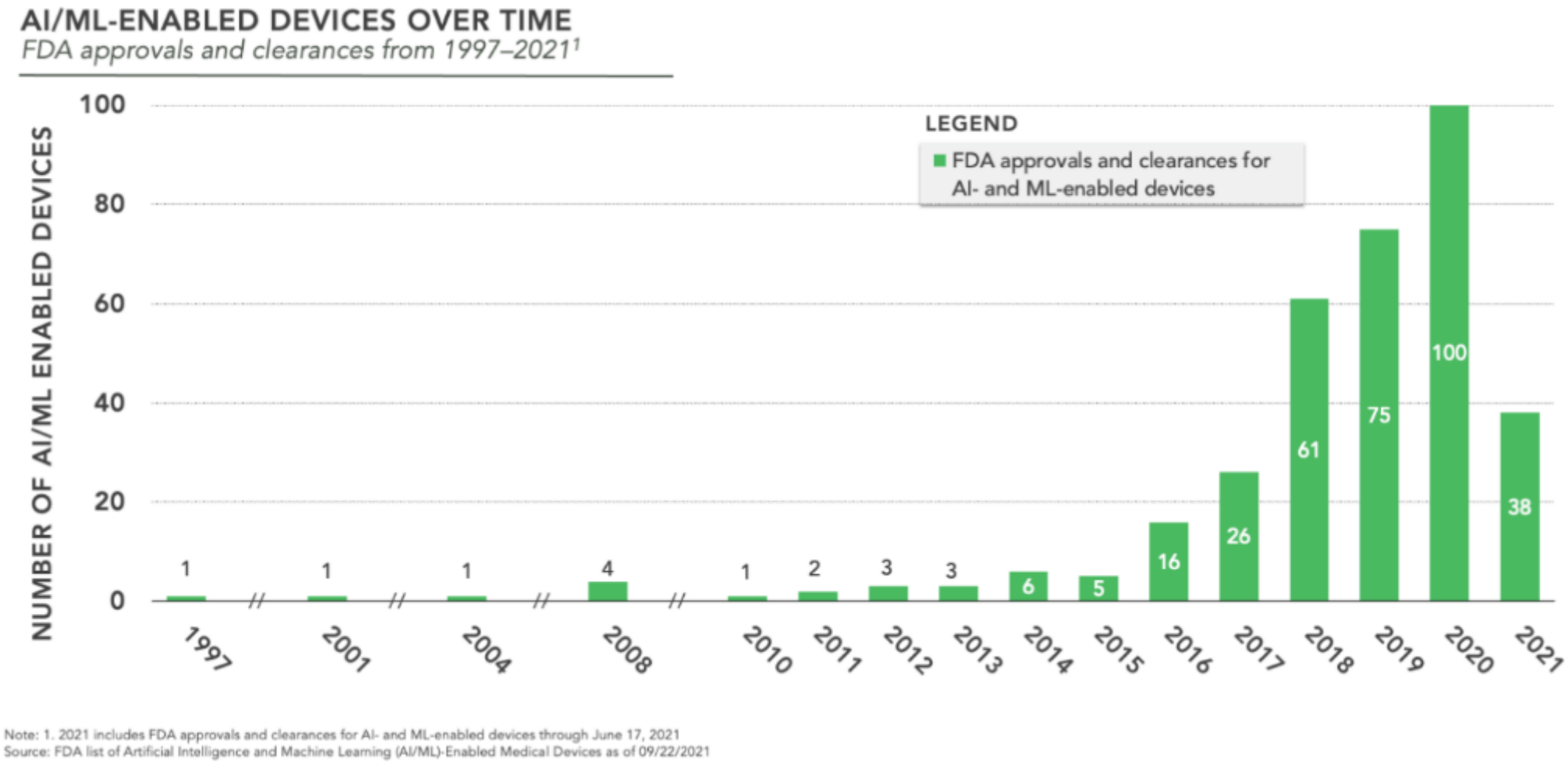
Adoption Rate of Artificial Intelligence by Industry

% firms in an industry that are adopting AI

Rank	Industry	Adoption %
1	Technology and communications	32%
2	Automotive and assembly	29%
3	Financial services	28%
4	Energy and resources	27%
5	Media and entertainment	22%
6	Transportation and logistics	21%
7	Consumer packaged goods	20%
8	Retail	19%
9	Health care	17%
10	Education	17%
11	Construction	16%
12	Professional services	13%
13	Travel and tourism	11%

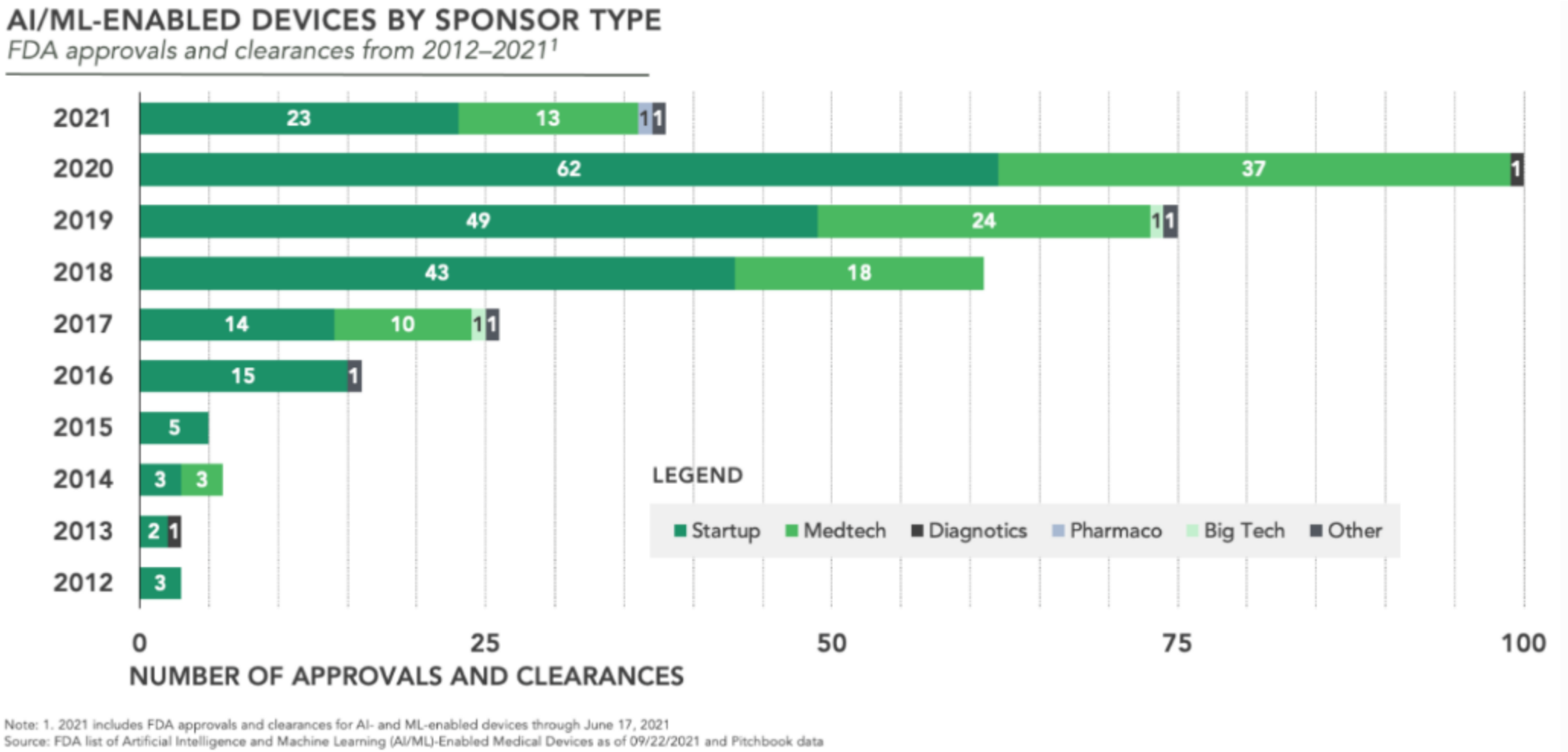
Source: https://priceonomics.com/which-industries-are-investing-in-artificial/

Adoption of AI in healthcare (1)



Source: https://rockhealth.com/insights/pulse-check-an-analysis-of-the-fdas-list-of-ai-and-ml-enabled-devices/

Adoption of AI in healthcare (2)



Source: https://rockhealth.com/insights/pulse-check-an-analysis-of-the-fdas-list-of-ai-and-ml-enabled-devices/

Adoption of AI in healthcare (3)

- Huge promises
  - Higher efficiency & throughput
  - lower costs for payors, society and providers
  - Better quality of care
  - Faster access to care for patients

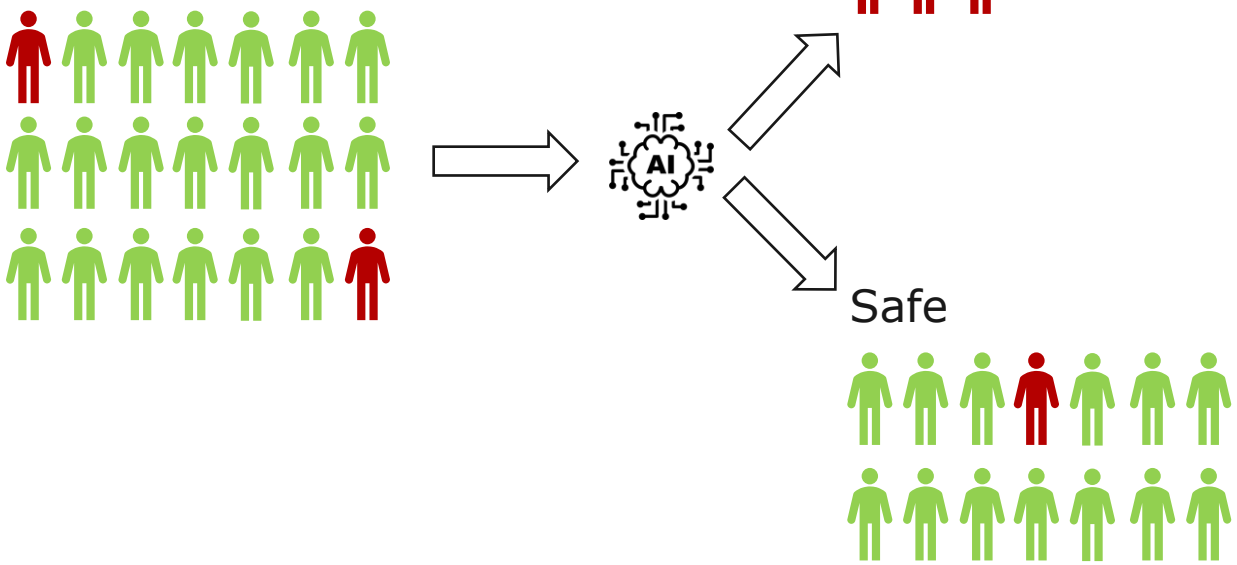
- And yet… true widespread use of AI in clinical setting remains fairly limited
- Many AI algorithms don’t make the step from lab to product, or do not ramp up commercially

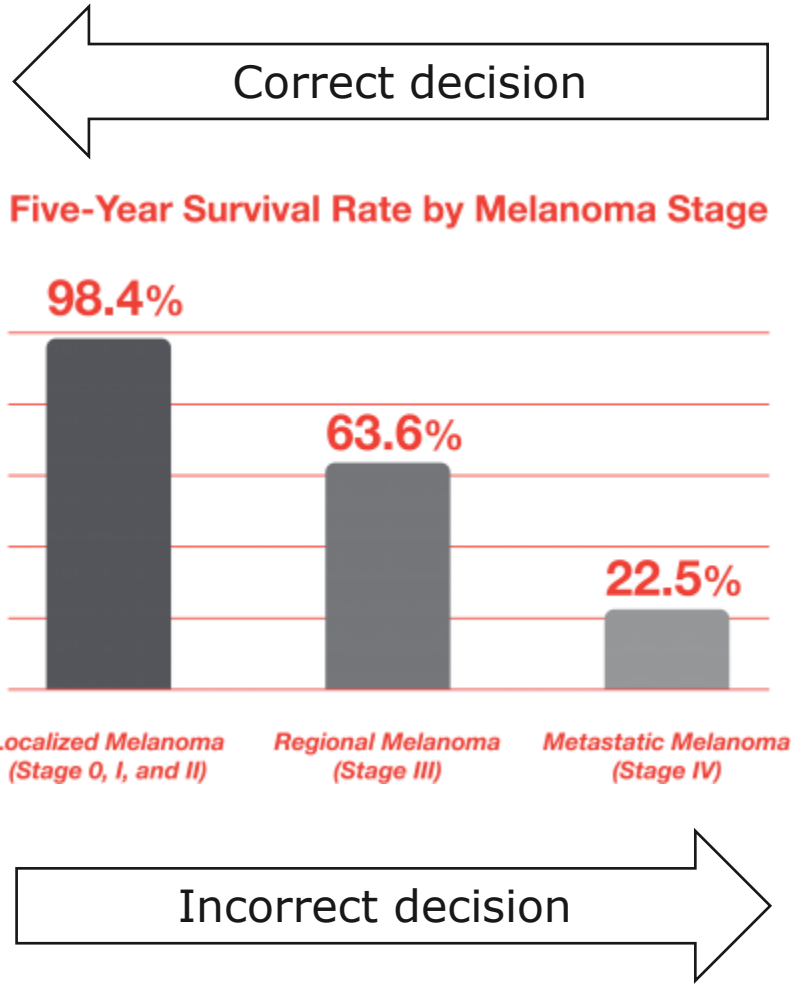
AI in healthcare

What makes AI in healthcare different & challenging?

- Life-critical application
- Complex ecosystem, different levels of specialization…
- Strict regulations, both on process and product level. High entry barrier.
- Data privacy
- Data variability (different standards, ways of working, terminologies)

Healthcare AI: life critical





Define the context and end users

Multi-disciplinary teams → Dermatologist → GP → Secondary care → Self screening

Regulatory certifications

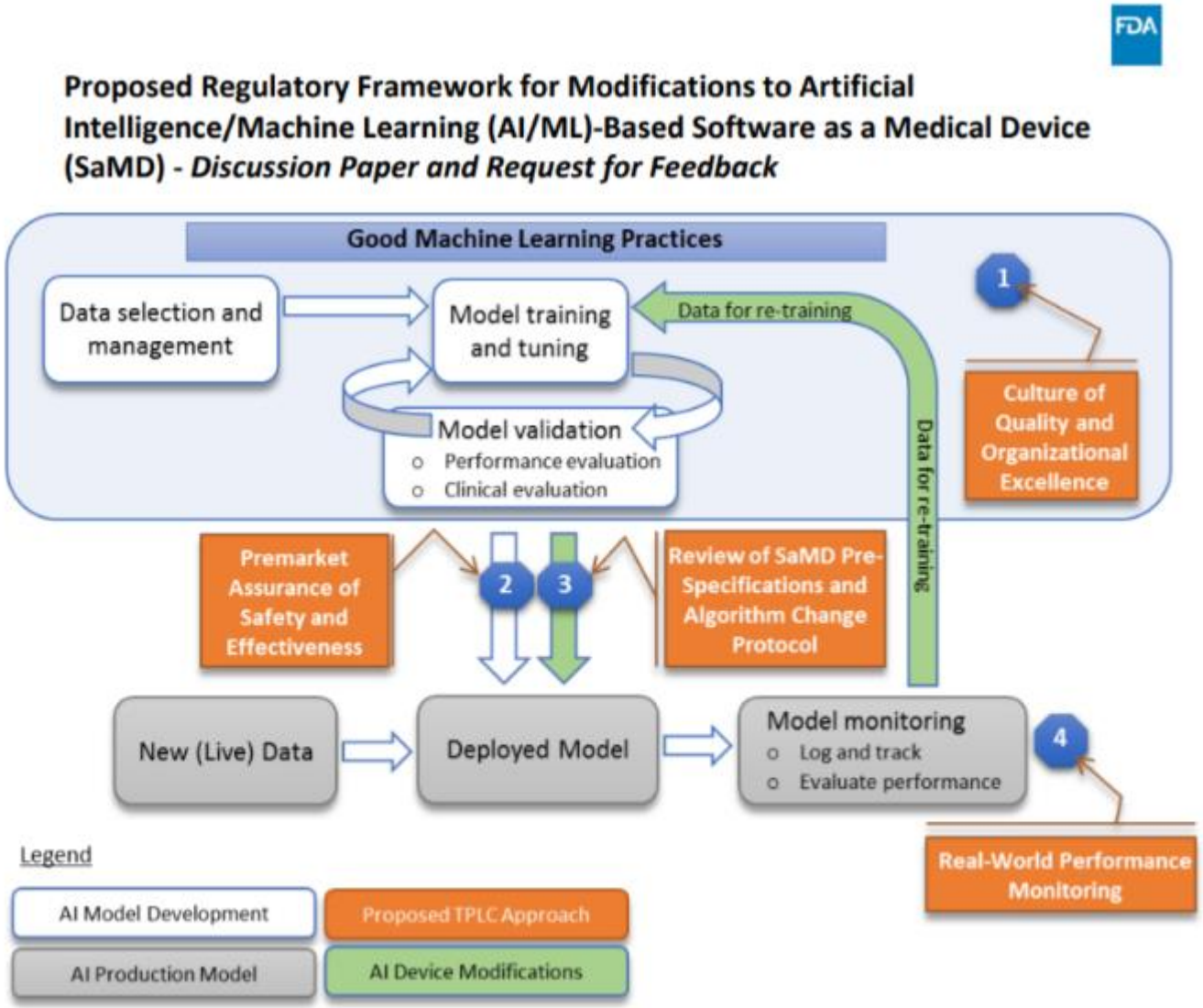
- In order to manufacture and sell medical devices, a range of medical certifications and approvals is required.

- This includes process related certifications
  - Eg. ISO13485
  - Eg. IEC 62304
  - ...

- And product specific certifications
  - Eg. MDR
  - Eg. FDA 510k/PMA

- It typically takes (at least) 12 months to go through the process compliance path, and on average 6-12 months for product certififcation

Regulatory compliance



Source: https://www.fda.gov/media/122535/download

Source: https://www.fda.gov/media/145022/download

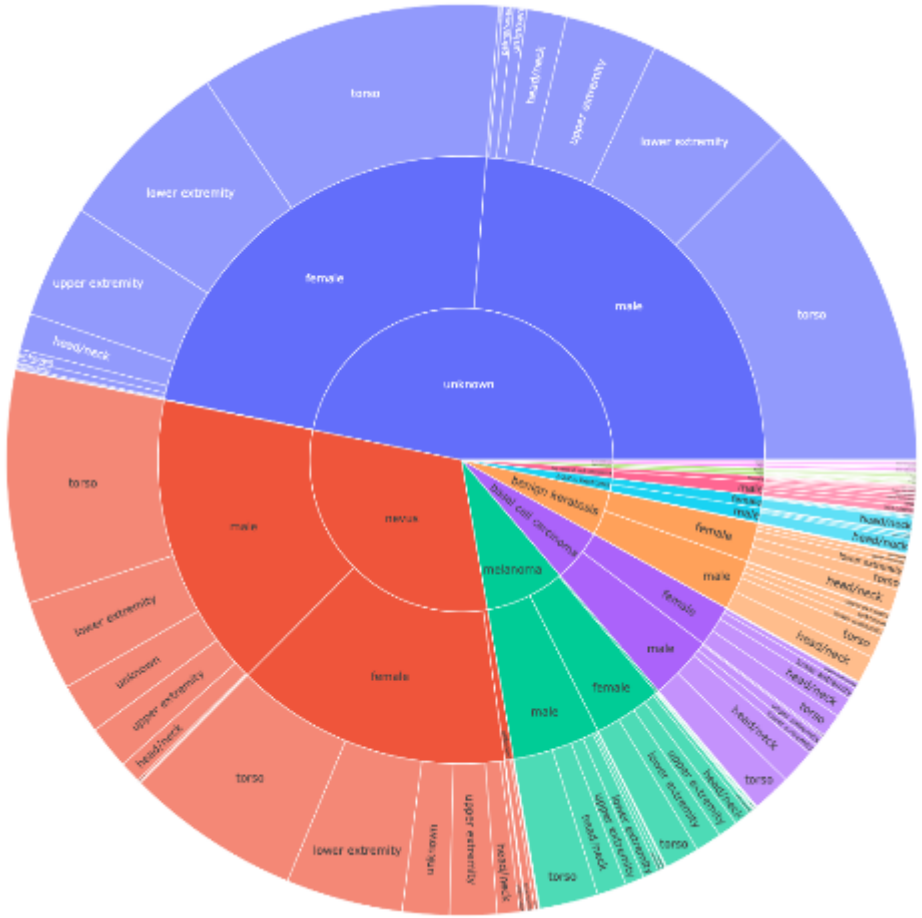
Data and privacy aspects

- Privacy is extremely important and needs to be strictly respected
- Large differences between EU & USA.
- GDPR & HIPAA is very specialized matter. Use experts to guide you.

- Be open and transparent about privacy (EULA, consent, …)
- Being sloppy with privacy rules or breaching privacy rules is a huge risk for your business. (severe fines, reputation damage, …)

HIPAA
Health Insurance Portability and Accountability Act

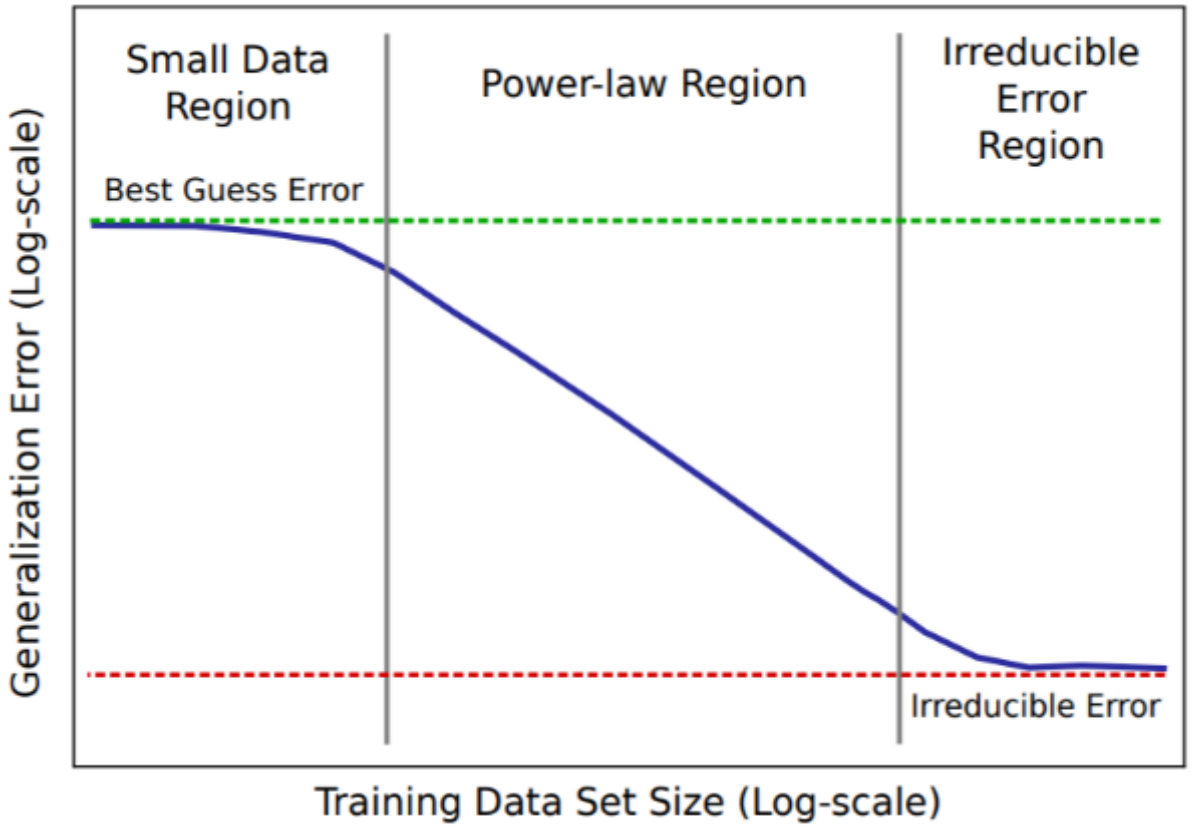
Data challenges for Real clinical AI development



Kaggle challenge data distribution

- Standardization on labels and outcome
- Need to establish ground truth
- Strict regulations for data gathering and product deployment
- “Small datasets” for training
- Long tailed problem, cancer is by definition ill structured
- Clinical features are not the same as the features for cat’s and dogs ->not always properly modelled in SOTA network architectures
- Non-clinical elements in data can introduce potential biases
- Reaching consensus on definitions and desired behavior
  - Actinic keratosis: not cancer, nor benign
  - Different grades of dysplasia for nevi have different treatment strategies
- Unbalanced data: melanoma but also “simple” benign classes severely underrepresented in public data

Estimating training set size requirements



Data collection strategy: preclinical / pre-release

- Find a healthy mix between
  - Public datasets (but check the license)
  - Clinical trials to collect data
  - Licensing of proprietary (competitive or other) datasets

- Important aspects to consider
  - Cost price
  - Speed of getting access to data
  - Similarity of the data
  - Quality of the data
  - Usefulness (can it be used e.g. for regulatory submissions; can you use it for marketing/user guides/brochures)

Source: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T

Design your solution with continuous data collection in mind

- Structured data versus unstructured data

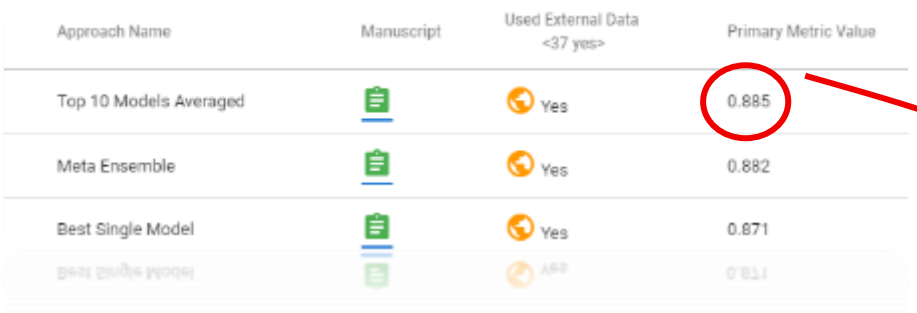
- Logs, logs, logs
  - E.g. Corrections user makes to annotations
  - Correlating user decisions with output of (prototype) algorithms
  - Features used or not used
  - Timing behavior

- Database structure and set of tools that allows for ease access and use of the collected data

- But always respect privacy…

AI in the lab and in literature vs in real clinical practice

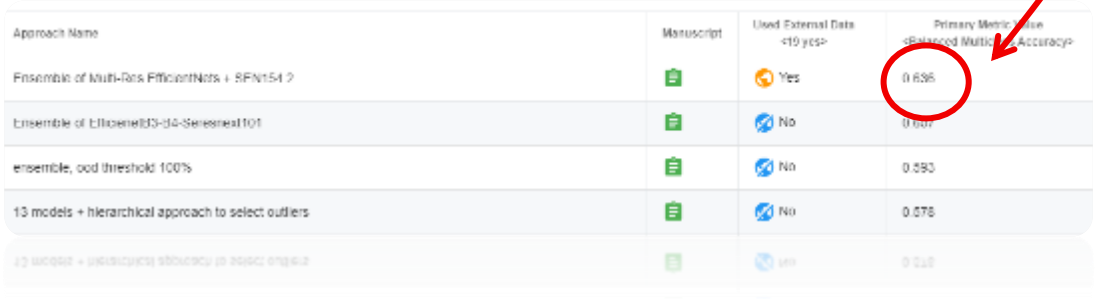
- ISIC Challenge 2018 leaderboard



Approach Name	Manuscript	Used External Data <37 yes>	Primary Metric Value
Top 10 Models Averaged		Yes	0.885
Meta Ensemble		Yes	0.882
Best Single Model		Yes	0.871

-0.25

- ISIC Challenge 2019 leaderboard



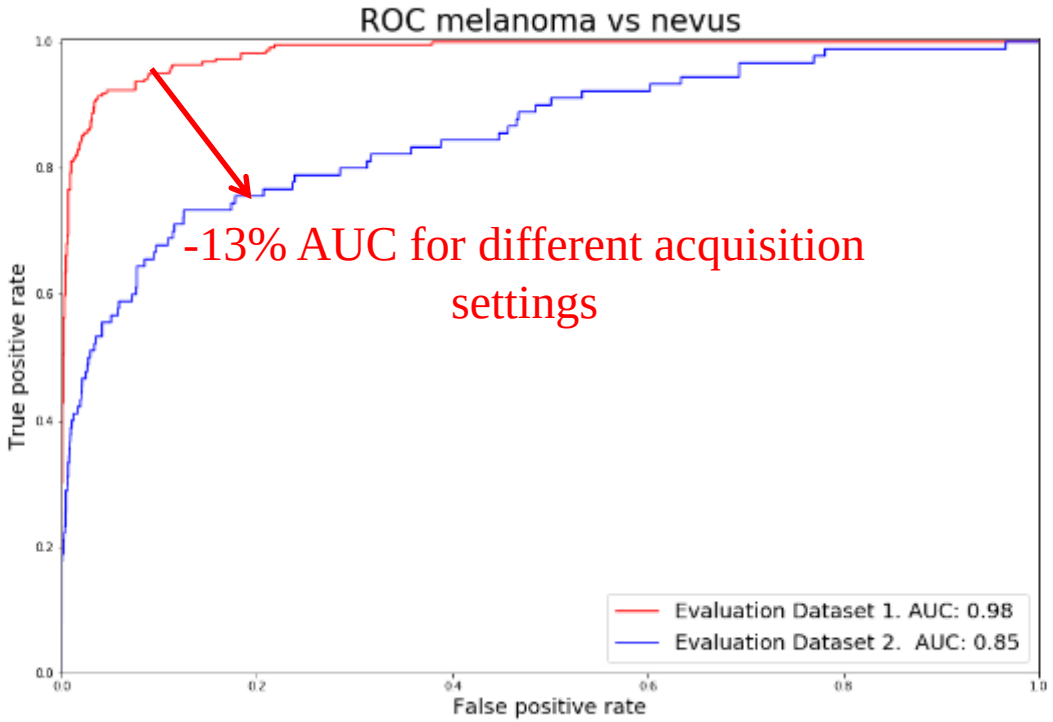
Approach Name	Manuscript	Used External Data <19 yes>	Primary Metric Value <Balanced Multi-class Accuracy>
Ensemble of Multi-Res EfficientNets + SEResNet152		Yes	0.636
Ensemble of EfficientB3-B4-SeResnet101		No	0.607
ensemble, ood threshold 100%		No	0.593
13 models + hierarchical approach to select outliers		No	0.578

- *"Algorithms tended to perform worse on images from other dermoscopic data sources, not present in public datasets."**
- *"Not all datasets in line with what the dermatologists are doing when visiting a patient in the clinical practice."**
- *"Although machine-learning algorithms outperformed human experts in nearly every aspect, higher accuracy in a diagnostic study with digital images does not necessarily mean better clinical performance or patient management"* **
- Many publications and reports are based on these datasets and challenges.

- ➢ Reported performance does not translate to clinical practice
- ➢ Scientific proof not representing real clinical usage is discarded by FDA and other notified bodies

* *Combalia, M, et al. "BCN20000: Dermoscopic lesions in the wild." arXiv preprint arXiv:1908.02288 (2019).*
** *Tschandl P, et al. "Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study". Lancet Oncol. (2019)*

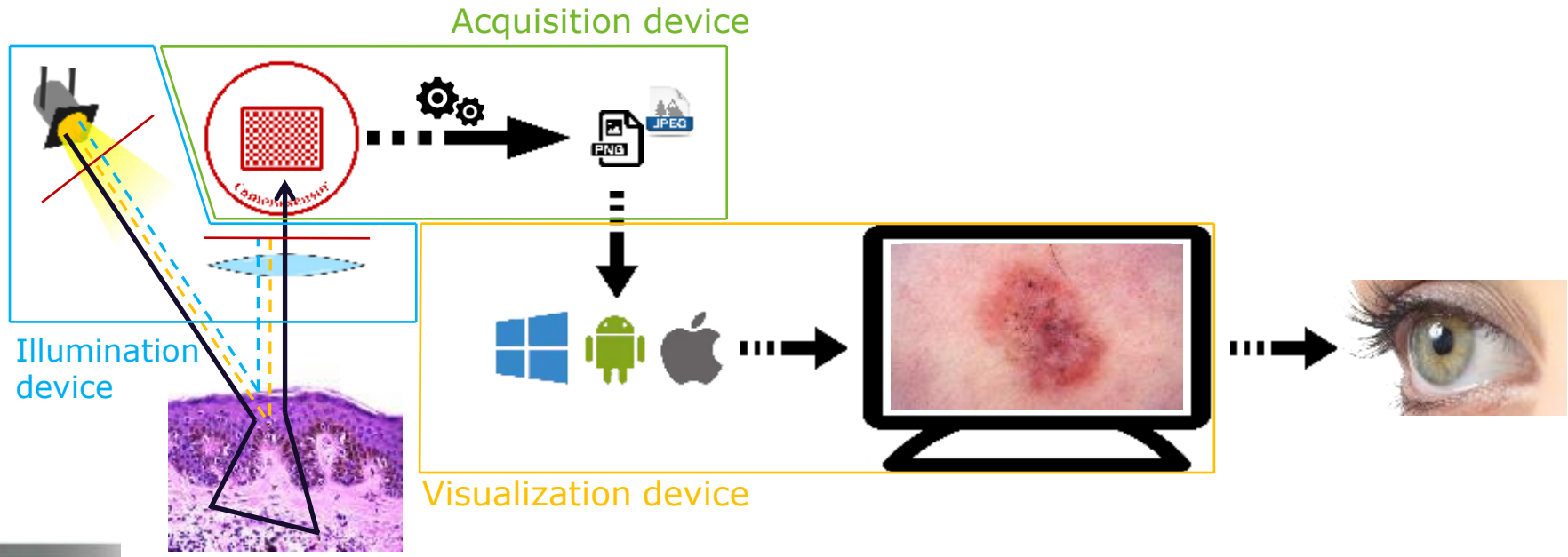
Impact of acquisition variability on AI



De Vylder et al. "The impact of consistent image capturing for artificial intelligence–based systems", Journal of the American Academy of Dermatology, 2020

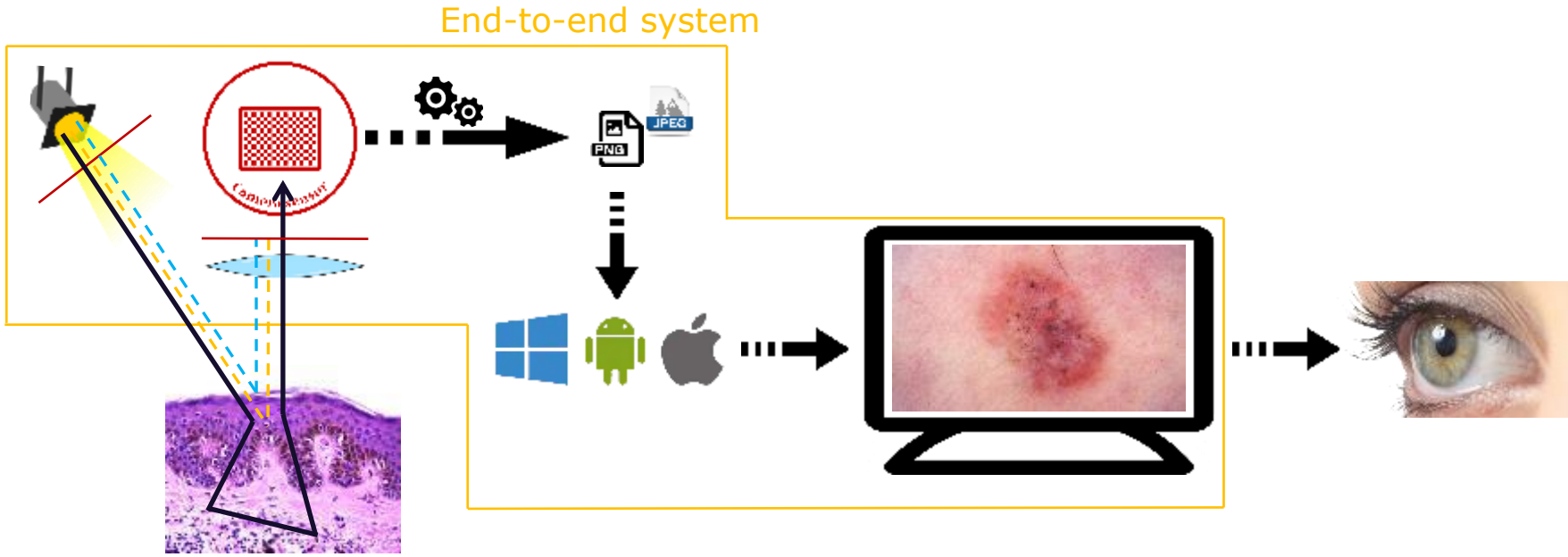
- AI is known to be sensitive to difference between the training and serving images.
- Even images that appear similar too a dermatologist can result in significant different AI output if the image acquisition and quality are not consistent.
- If images are collected with a wide range of acquisition devices (eg. different types/brands of dermatoscopes) and settings, then this will have large negative impact on AI performance

Standard digital dermatoscope

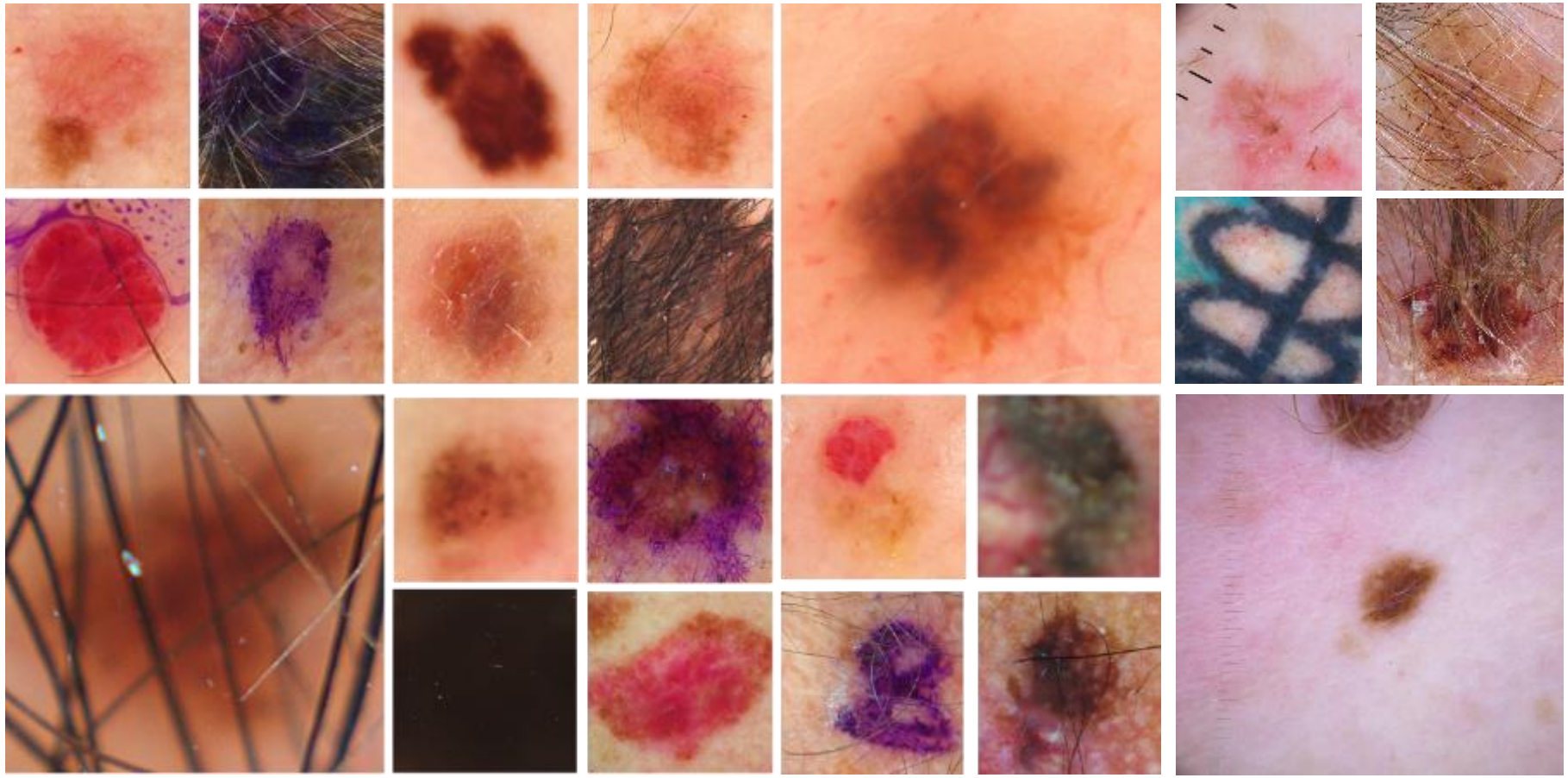


- Analogue dermatoscope is combined with digital acquisition device
- Resulting image to be exported and reviewed on an external display
- Acquisition device image processing is unaware of illumination characteristics

Demetra digital dermatoscope



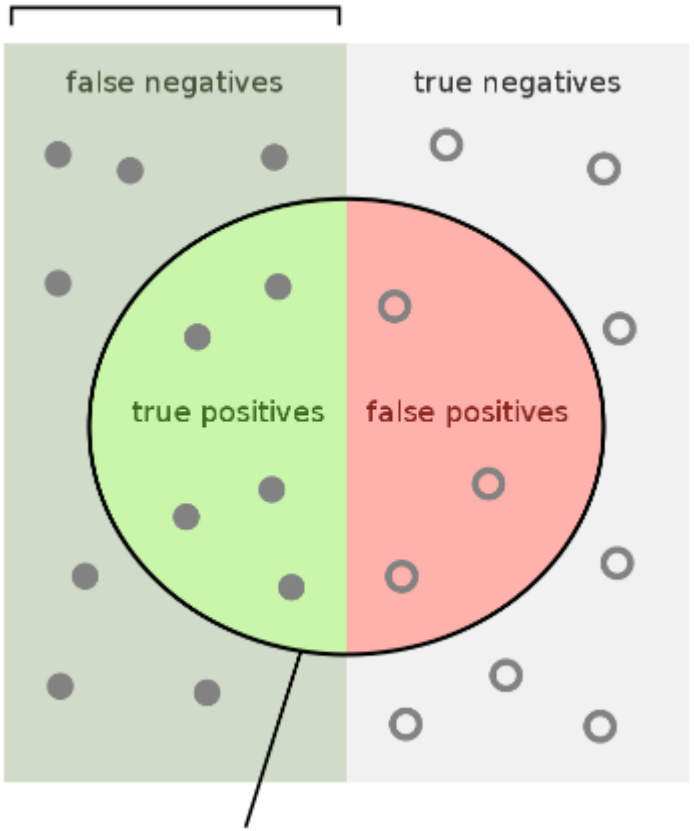
- → Illumination, acquisition, processing and visualization integrated in a single device
- → Resulting image is visualized on the acquisition device (or external display)
- → Color correction can be done end-to-end [2]



- Combining dedicated artifact detectors and generic outlier detectors is essential for data cleaning
- Can be used to improve AI training, but also to prevent inference on “bad” images

Performance evaluation of AI algorithms in HC

True malignant lesion

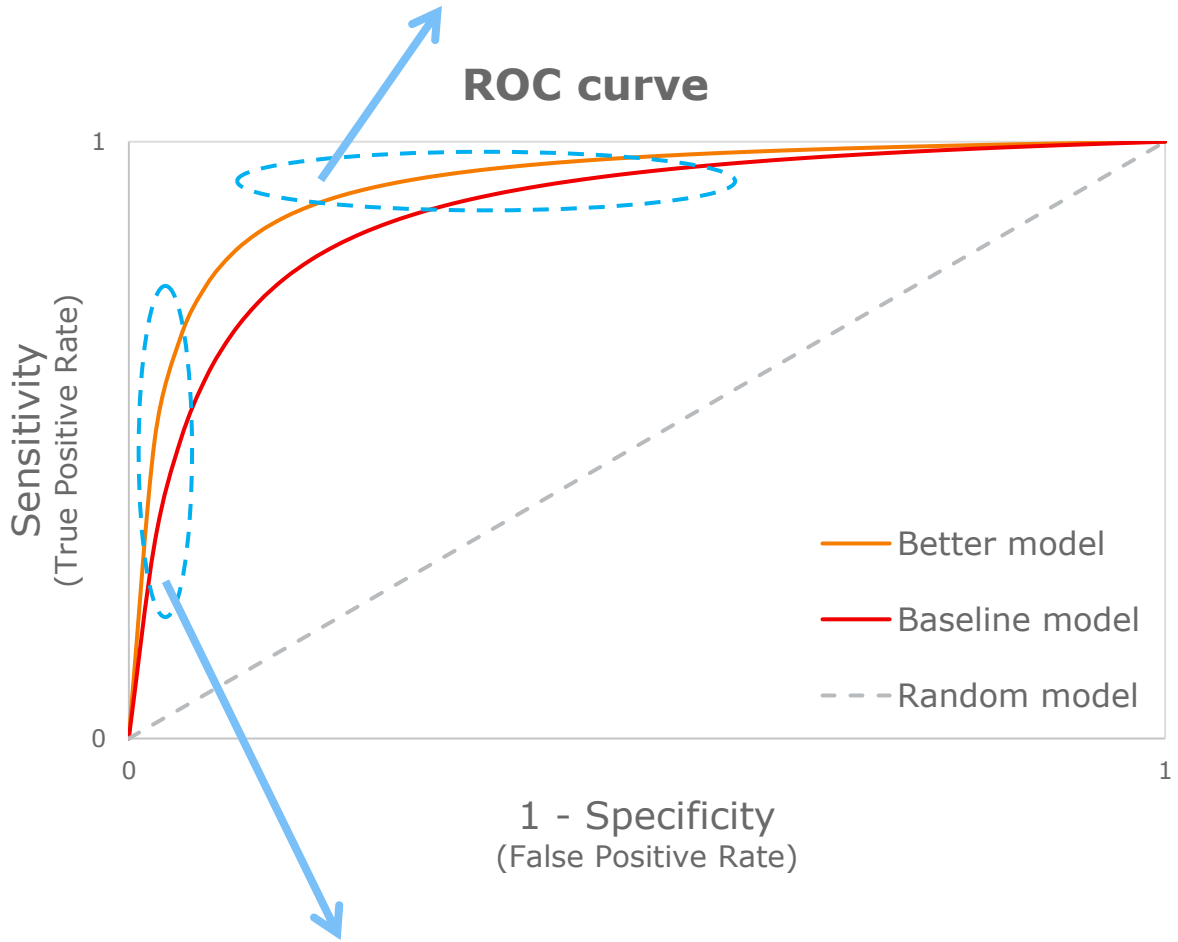


Lesions scored
as malignant by AI

Sensitivity = $\frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$

Specificity = $\frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$

More sensitive: avoid missing malignant lesions

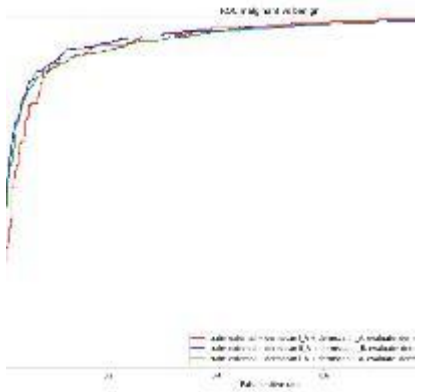


More specific: avoid unnecessary lesion excisions

→ Better AI model has higher Area Under the Curve (AUC)

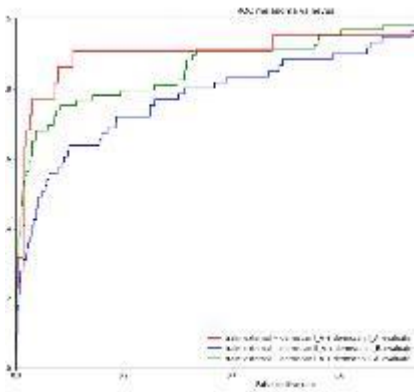
Validation methods

What you optimize for is not the same as the desired outcome



ROC curves

- Good for binary problems
- Extensions exist for multi-classes, e.g. one-against-vs-other

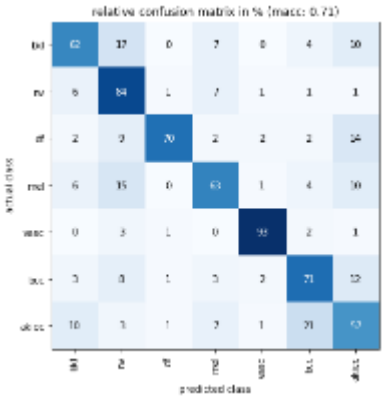


Data sharding

- Similar metrics, but evaluated on different classes, populations, context,…
- Evaluate fairness

Human evaluation

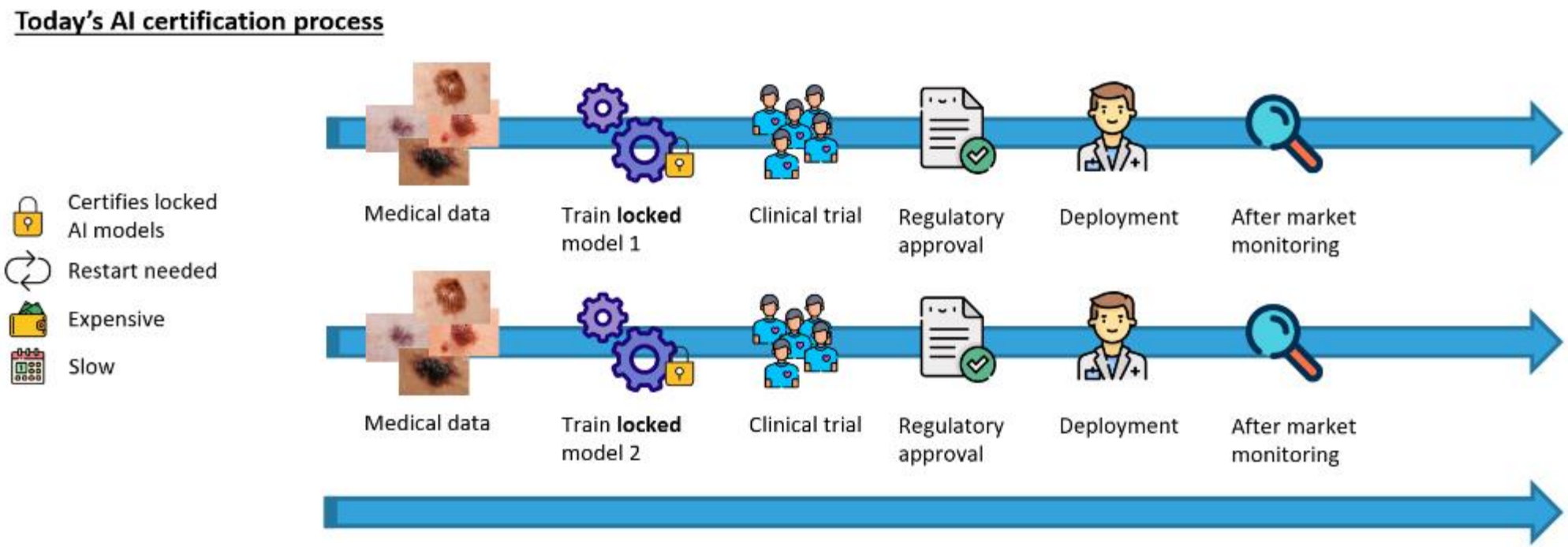
- Ground truth does not always exist
- Does the AI actually help the end user



Confusion metrics

- Insights per class
- Insights between which classes is confusion

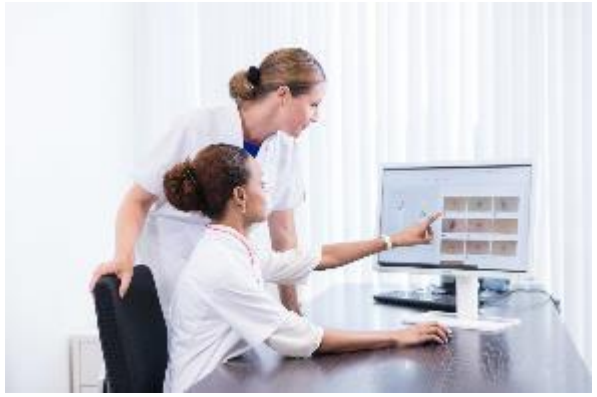
Changing AI algorithms postmarket



Towards broad AI use in healthcare

Being more than a pure replacement of manual task.

- Digitizing skin imaging supported by AI could open up novel clinical pathways and improve quality and efficiency of diagnosis
  - Improved patient pathways (priority routes GP -> dermatologist)
  - AI supported screening & prevention
  - Tele dermatology
  - Second opinion / expert consultation
  - Improved treatment follow-up
  - Population wide trend analysis

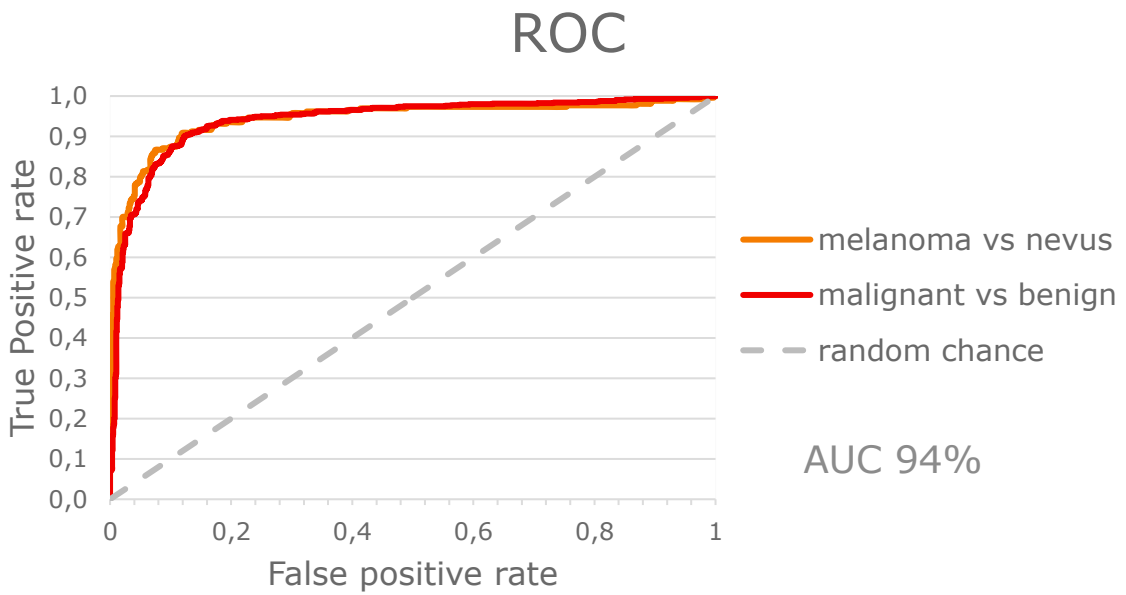


- However, this requires a clear vision and rigorous discipline when defining and designing the new digital solution
  - Cloud connectivity required, all data in the cloud (with respect for privacy & security)
  - Structured data, EMR integration

AI based risk scoring: potential to change the patient pathway

Estimate of the risk on malignancy (including MM, BCC, AK, BD & SCC)

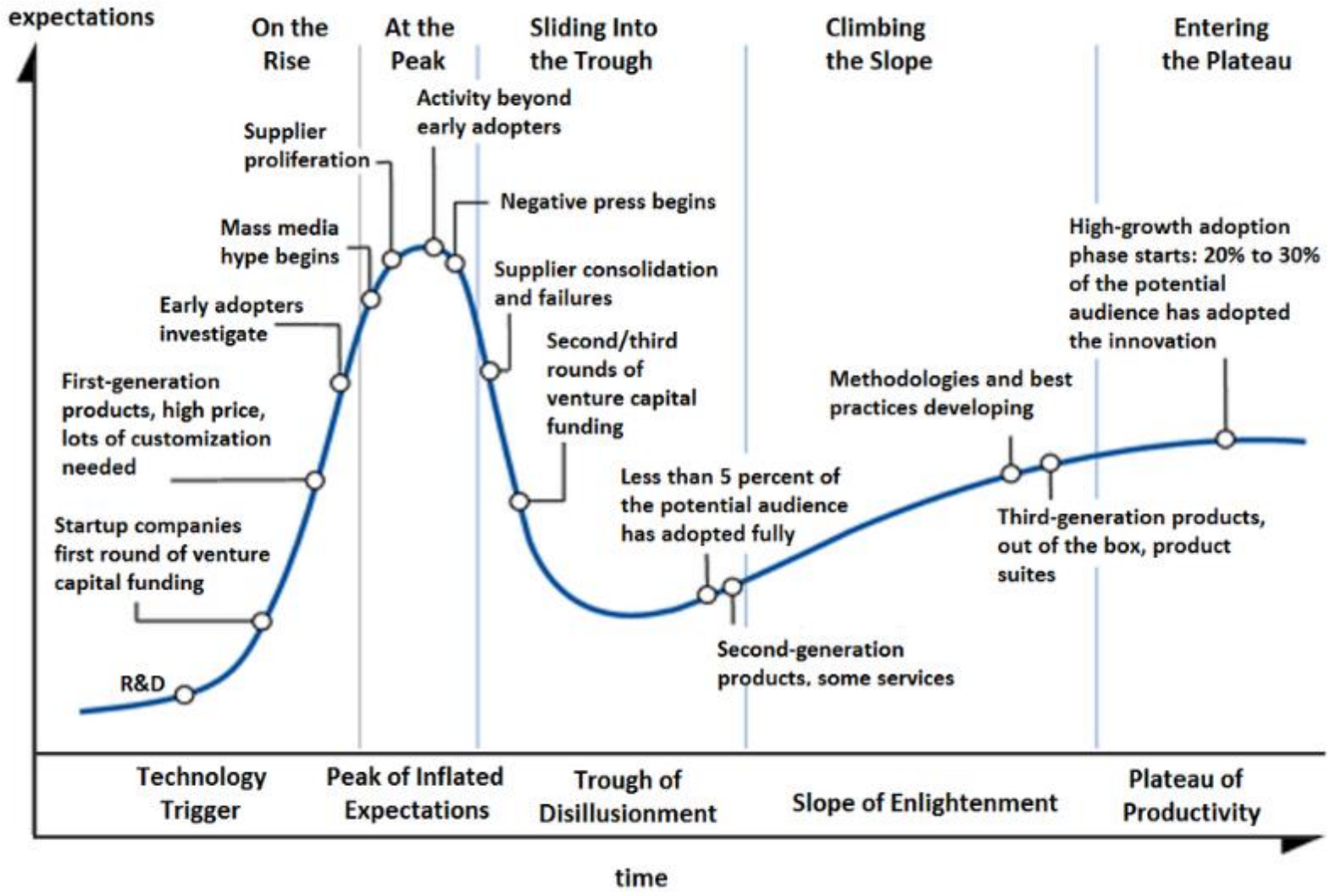
In research phase, not available for clinical use yet



Drivers for broad adoption of AI in healthcare

- Beyond mere automation of human tasks
- Beyond scattered niche applications (volume + integrated solution)
- Standard interfaces, integration with EMR & other data sources
- True attractive use cases (ROI), disruptive change of workflow
- Stable, flexible regulatory framework

Where is AI really?



Source: https://it.wikipedia.org/wiki/File:Hype-Cycle-General.png

Conclusions

- An AI solution can be so much more than the mere replacement of a a human tasks
- Understanding the (future) clinical workflow, early on involving all stakeholders and having a clear vision is key.
- Combination of technological, business & payor model innovation can speed up solution adoption.
- Artificial intelligence in healthcare is begining to take the step from niche use to broader adoption. Exciting innovations to come.

Contact: tom.kimpe@barco.com

@Tom_Kimpe

www.barco.com