

# AI applications for Data Management

Andre Dekker

Medical Physicist | Professor of Clinical Data Science

Maastricht UMC+ | Maastricht University | Maastricht Clinic

Modern Radiation Oncology: multidisciplinary in the era of OMICS and AI guided oncology

32<sup>nd</sup> Residential Course

Rome | Oct 17, 2022 | 15:45-16:00

# Disclosures

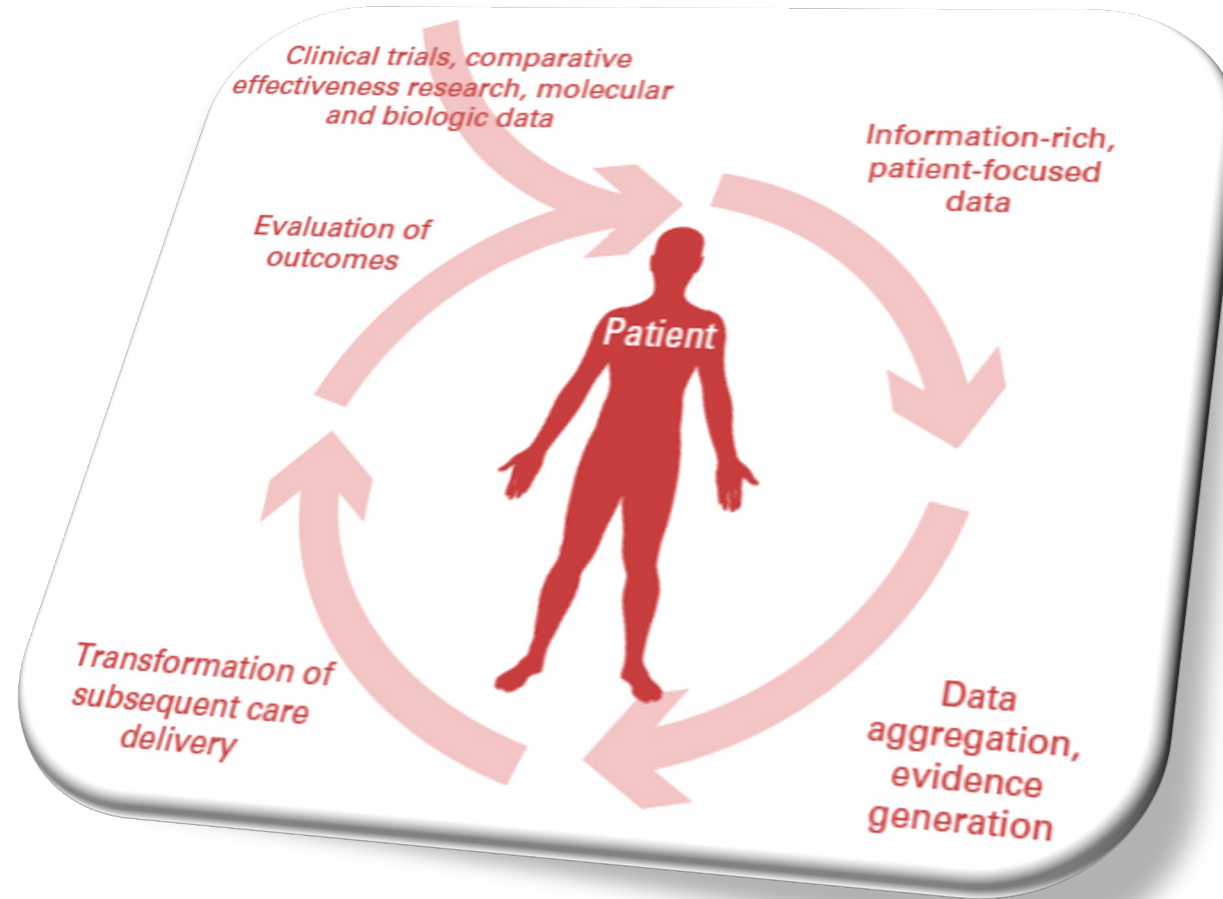
Research collaborations incl. funding, consultancy and speaker honoraria

- Pharma: Roche, Janssen, Bristol-Myers Squibb
- MedTech/Data: Varian, Siemens, Medtronic, Philips, Mirada Medical, IQVIA
- Health insurance: CZ Health Insurance

Spin-offs and commercial ventures

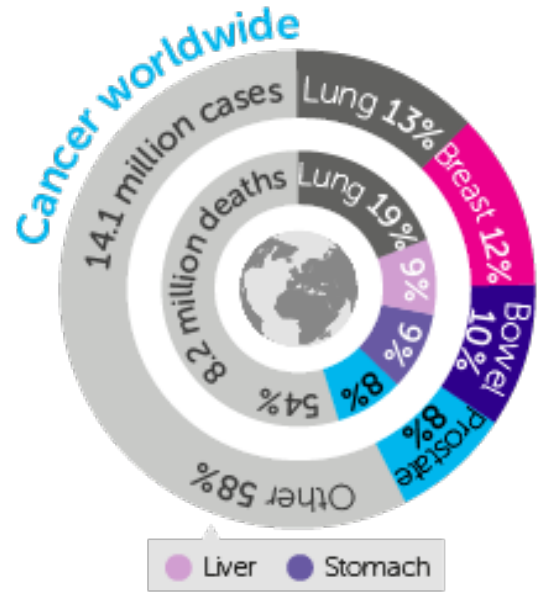
- MAASTRO Innovations B.V.
- Medical Data Works B.V.
- Various patents on medical machine learning & Radiomics

# Why data is important



## Data Management for AI Applications

# Data



**Oncology**  
 2007-2017  
 150M patients  
 0.1-10GB per patient

**15-1500PB**  
**80% unstructured**

**Hospitals**

- China: 25.000
- India: 35.000
- Germany: 2.000
- France: 2.300
- Italy: 1.100
- USA: 5.500
- Australia: 1.400

**TOTAL ~100.000**

# CORAL: Community in Oncology for Rapid Learning

- Toronto
- Boston
- Michigan
- Ottawa
- Philadelphia
- Tampa

- Poznan
- Oslo
- Bergen
- Odense
- Aachen
- Basel
- Zurich
- Nicosia
- Rome
- Aarhus

- Cardiff
- Manchester
- Leeds
- Sheffield
- Oxford
- Cambridge
- Hull

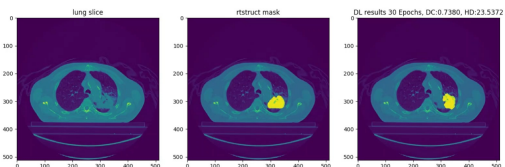
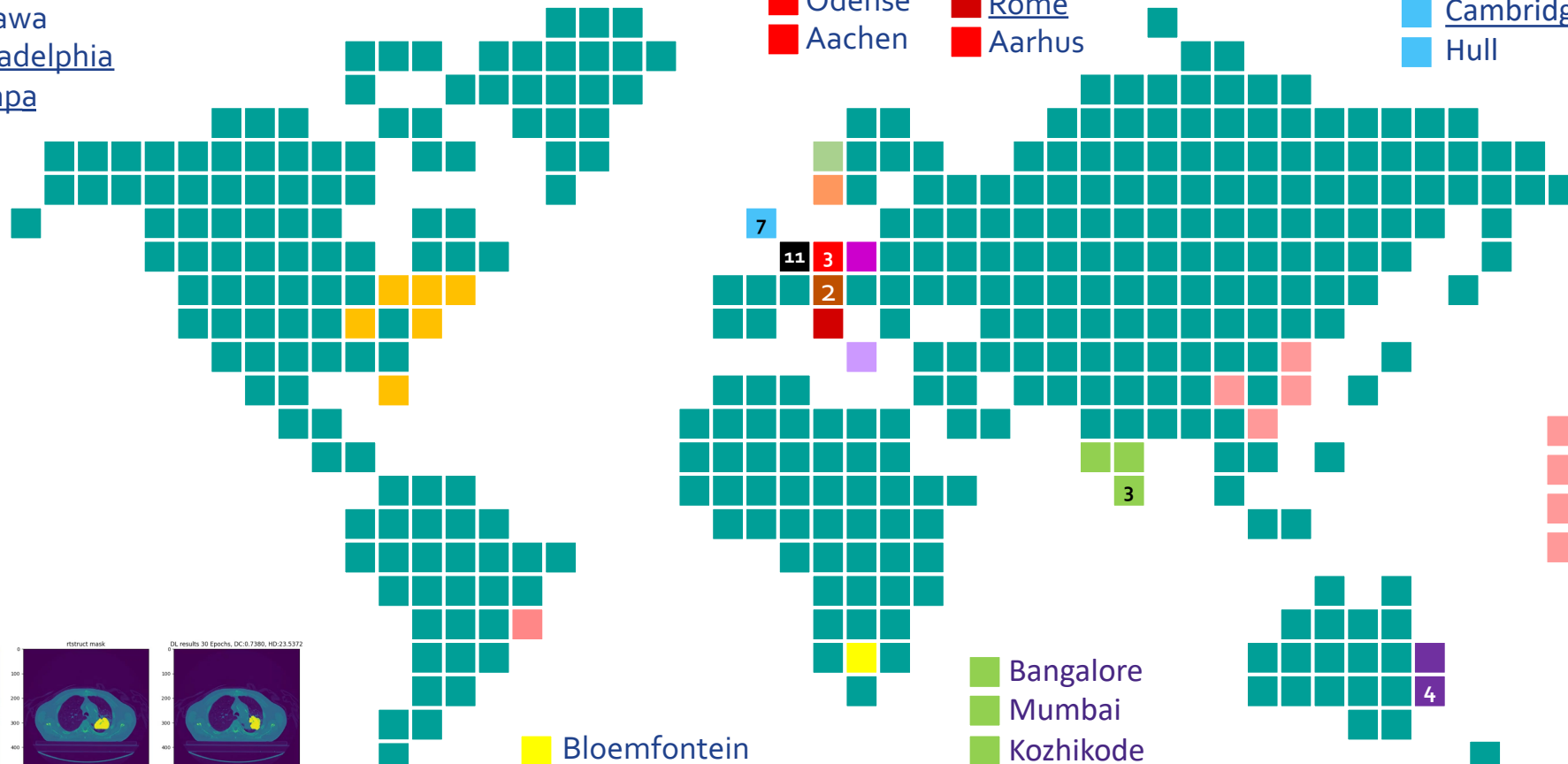
- Hasselt
- Eindhoven
- Maastricht
- Liege
- Nijmegen
- Amsterdam
- Rotterdam
- Zwolle
- Groningen
- Tilburg
- Leuven

- Shanghai
- Suining
- Tianjin
- Shantou

- Liverpool
- Wollongong
- Newcastle
- Westmead
- Coffs Harbour

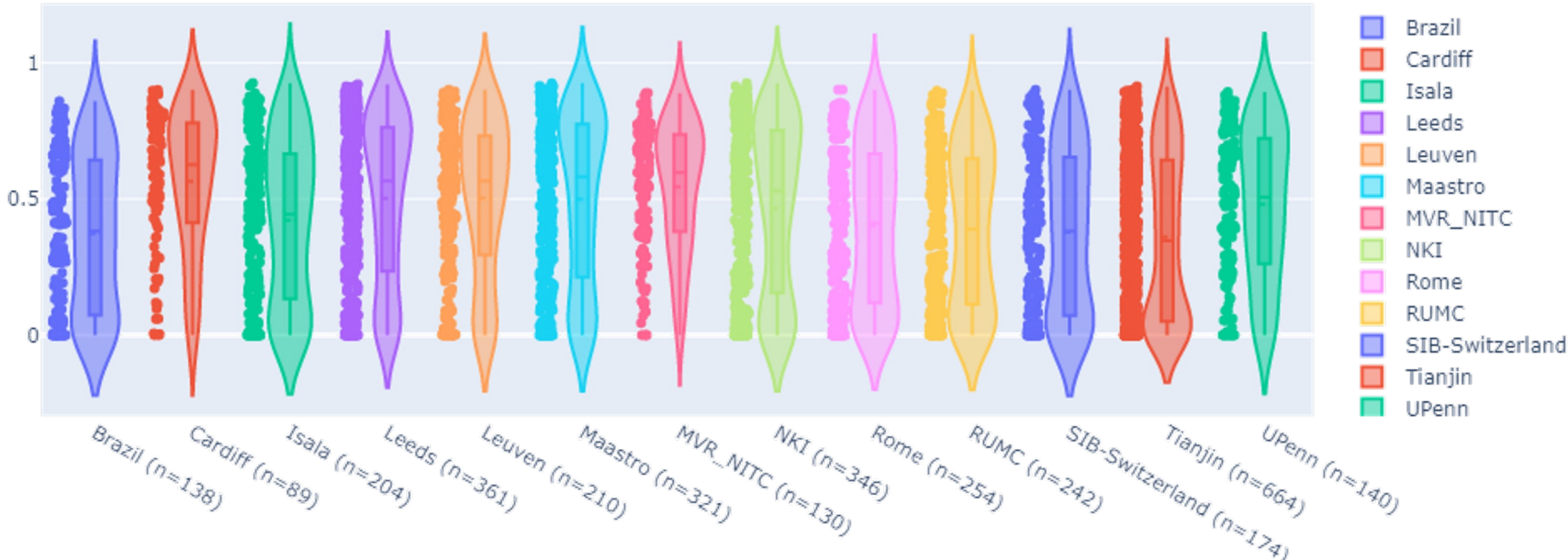
- Bloemfontein
- Sao Paulo

- Bangalore
- Mumbai
- Kozhikode
- Hyderabad
- Vellore



# AI for Lung GTV delineation – Oct 11, 2022

Mean Dice



## AI Applications for Data Management



# Extract data - TNM

ThoraxMass visible in the left upper lobe with a maximum size estimated at 8-46 of 4.7 x 3.0 cm. Possible involvement in mediastinum. Satellite nodes visible at 8-41 with an estimated size of 1.3 cm. Lymph node visible at station 7 with a size of circa 5,2 cm. No lymph nodes visible at contralateral side. Small consolidation middle lobe. No indication of atelectasis.

- Present
- Involved
- Context Modifier
- Context Target

**TNM-8 Lung**

English Classify

T2b N0 M

**Primary Tumor**

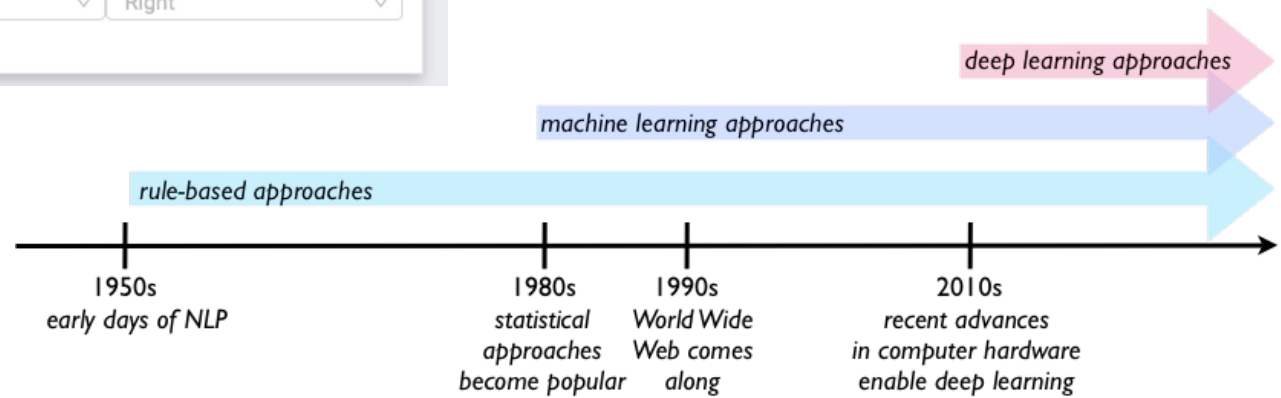
4.7 cm Left Side

Present Involved

Satellite Nodule (T3)  Ipsilateral Tumor (T4)

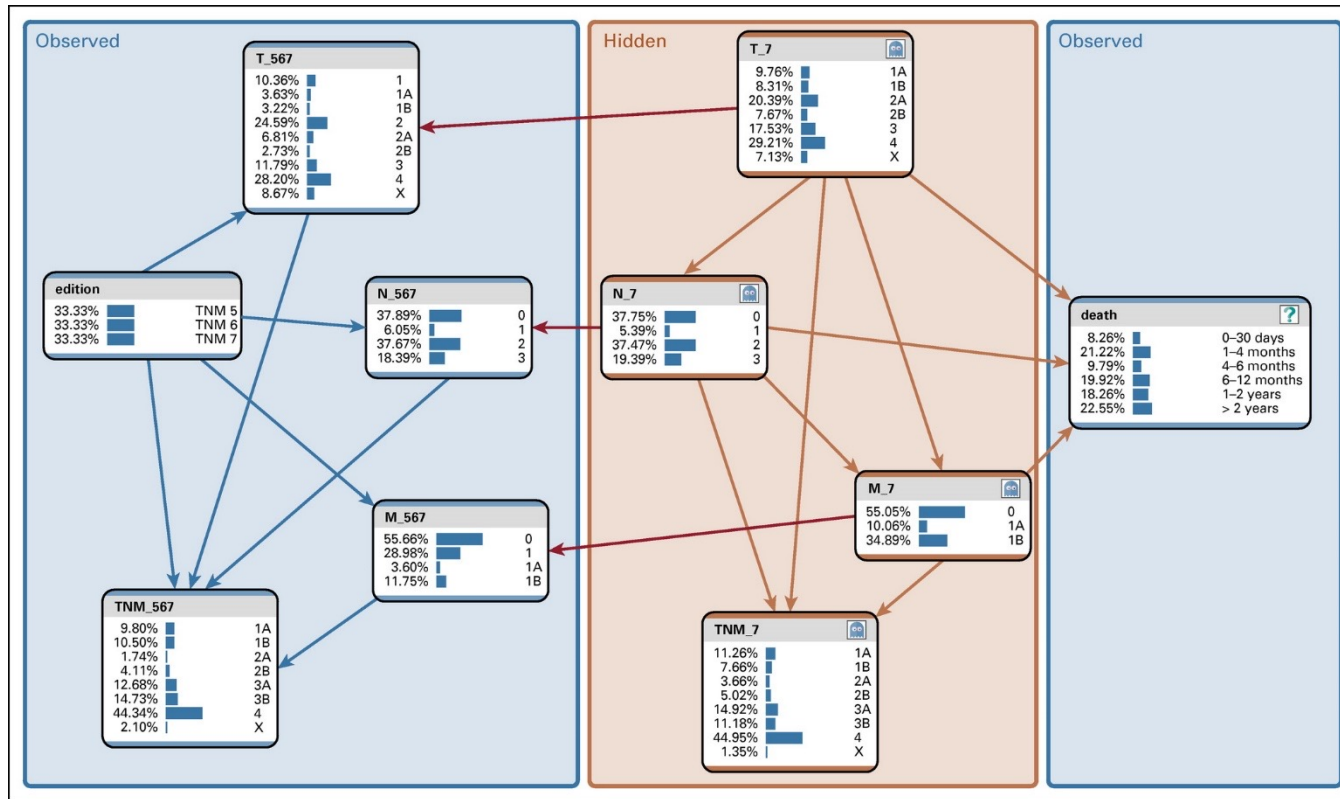
**Lymph Nodes**

Left Right



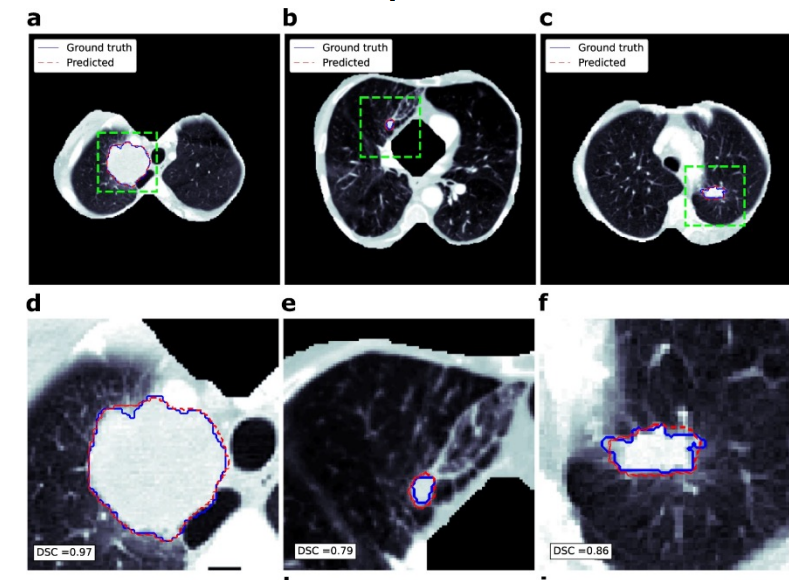
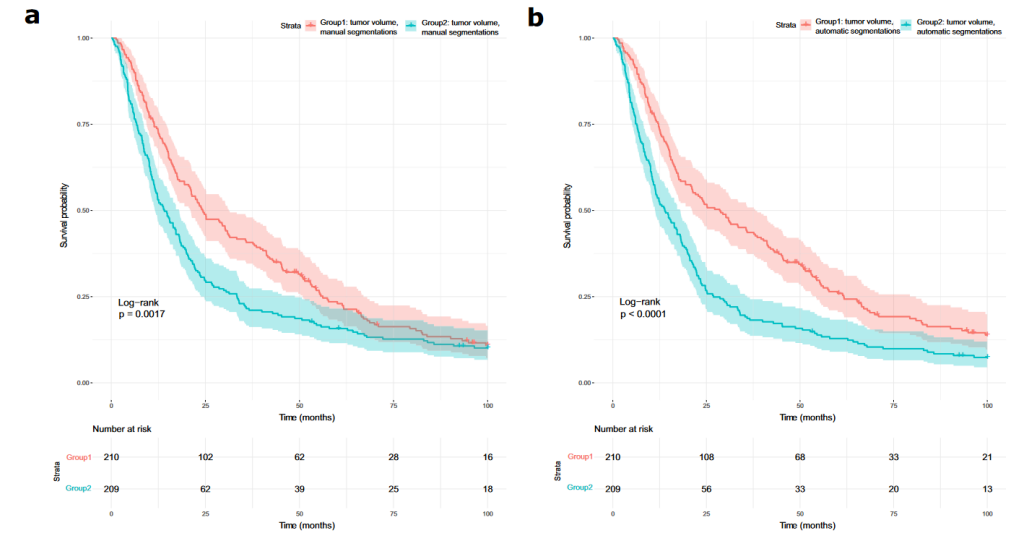
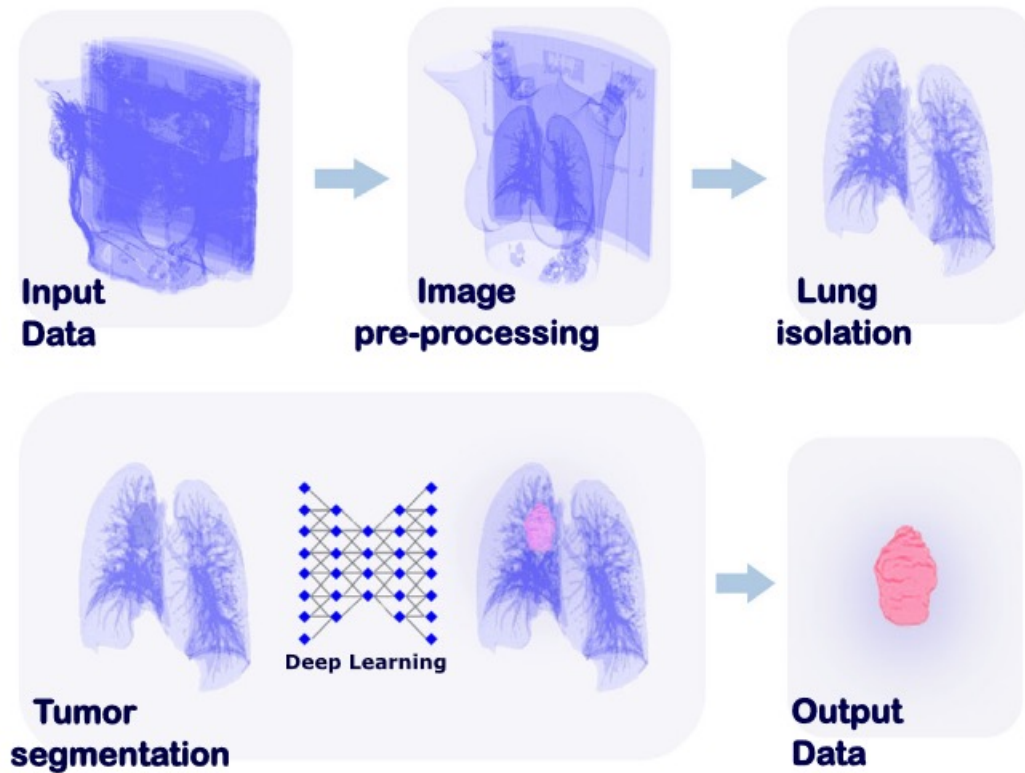
Nobel et al. Insights Imaging 12, 77 (2021).

# Impute data - TNM Reclassification



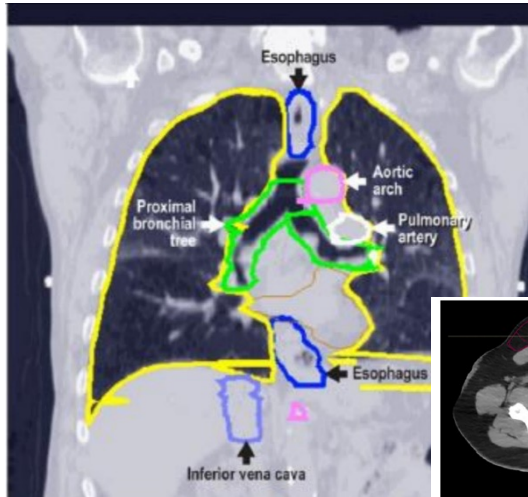
n = 146,084

# Standardize data - GTV & RECIST

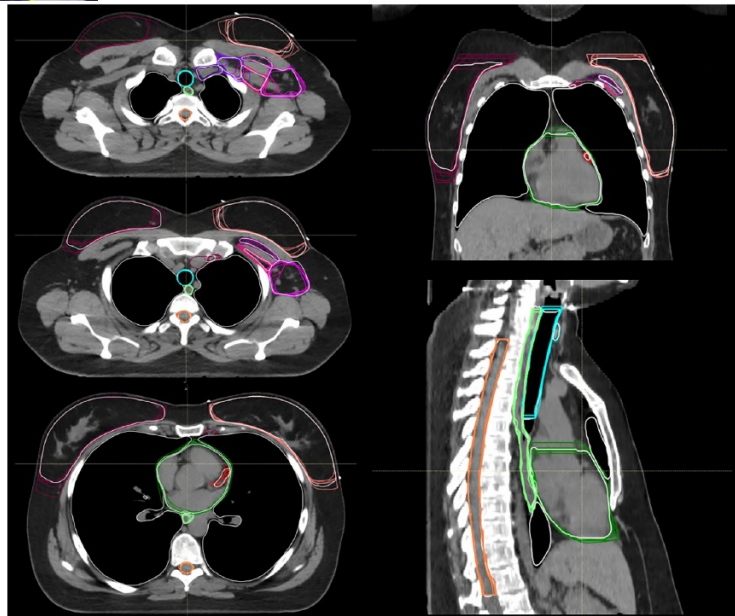


Primakov et al. Nature Communications 2022

# Standardize/add data - organs at risk (OARs) – Recalculate dose

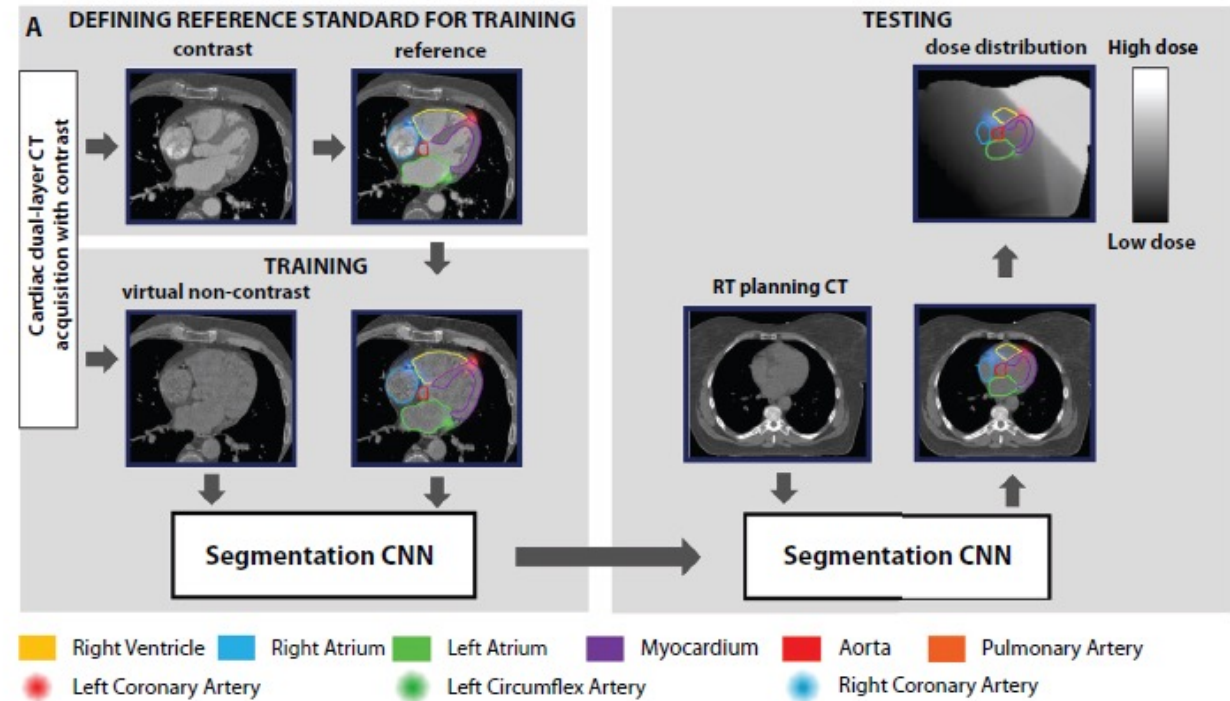


*Radiother Oncol* 2018;126:312.



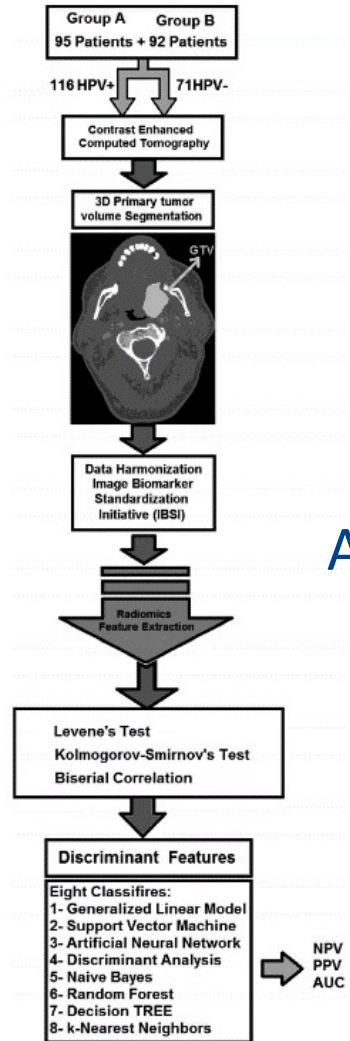
A representative patient case with DL contours in white and manual contours from 5 different oncologists in various colours.

*Radiother Oncology* 2022;173:62

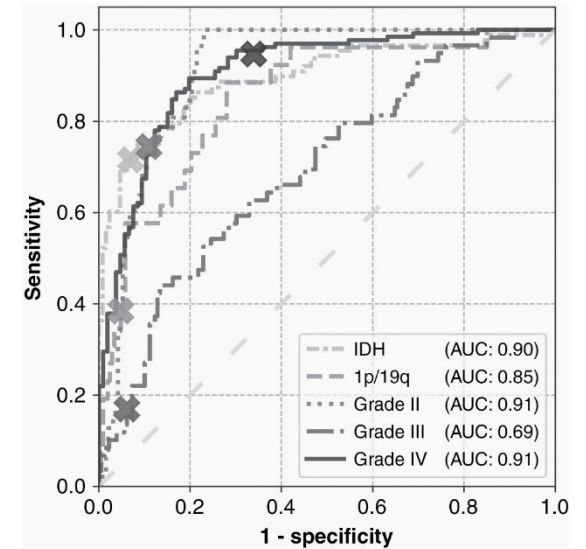
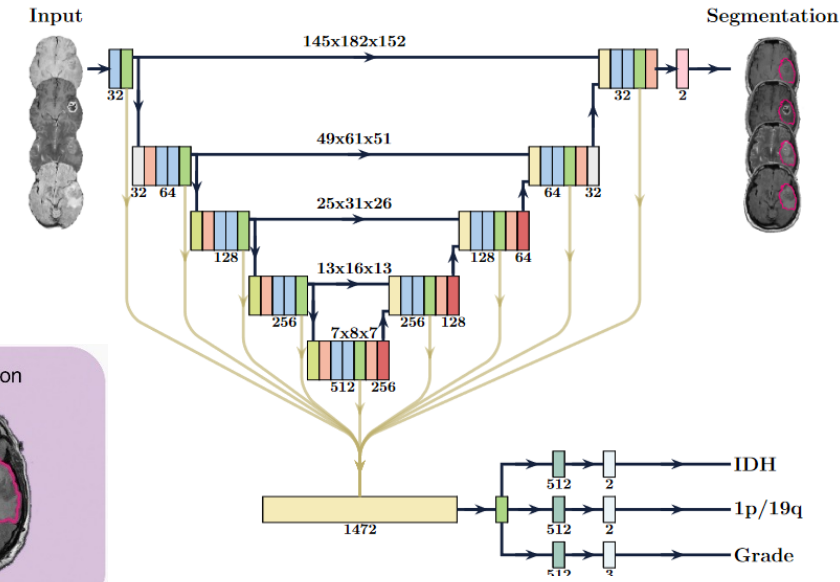
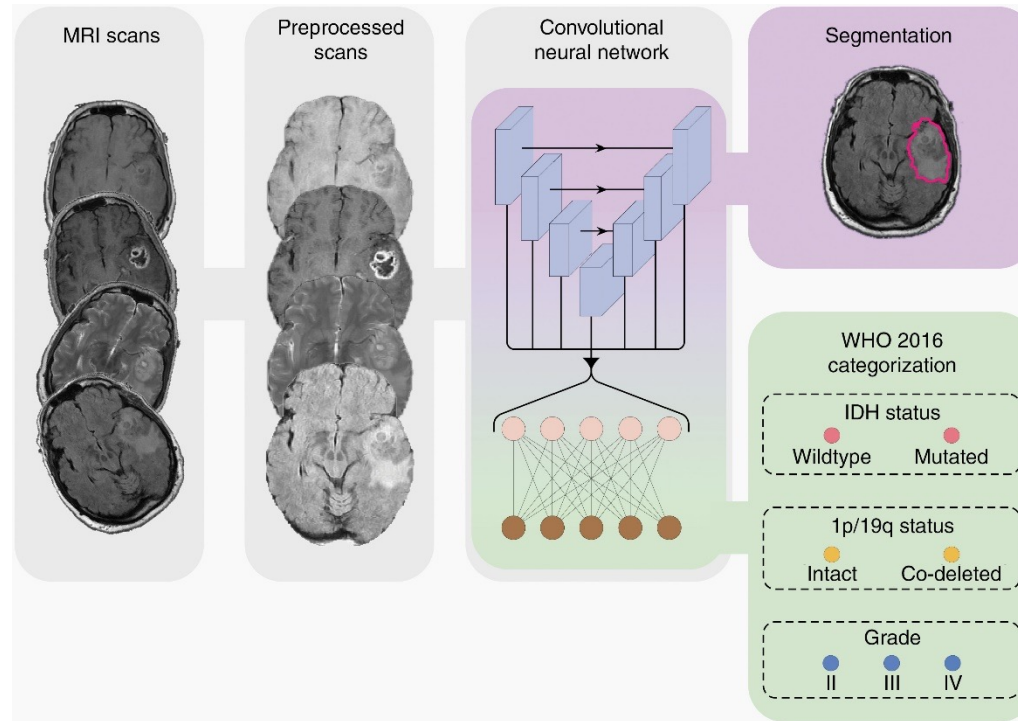


*Int J Radiation Oncol Biol Phys* 2022;112:611

# Add data - Tumor characterization



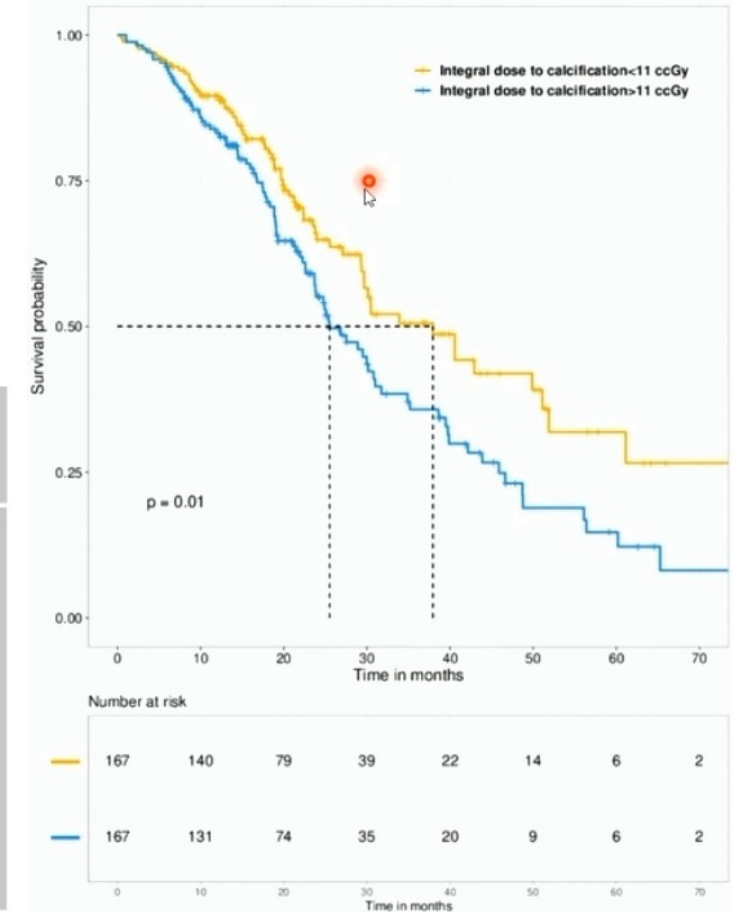
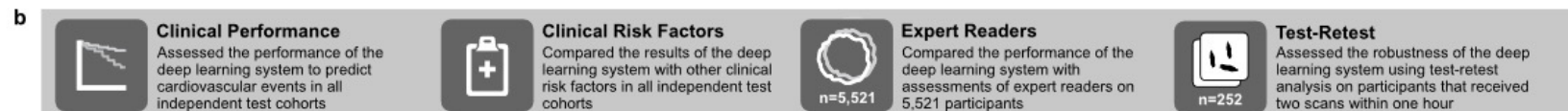
HPV  
AUC~0.85



Med Phys 2020;47:563

Van der Voort et al. Neuro Oncology. 2022

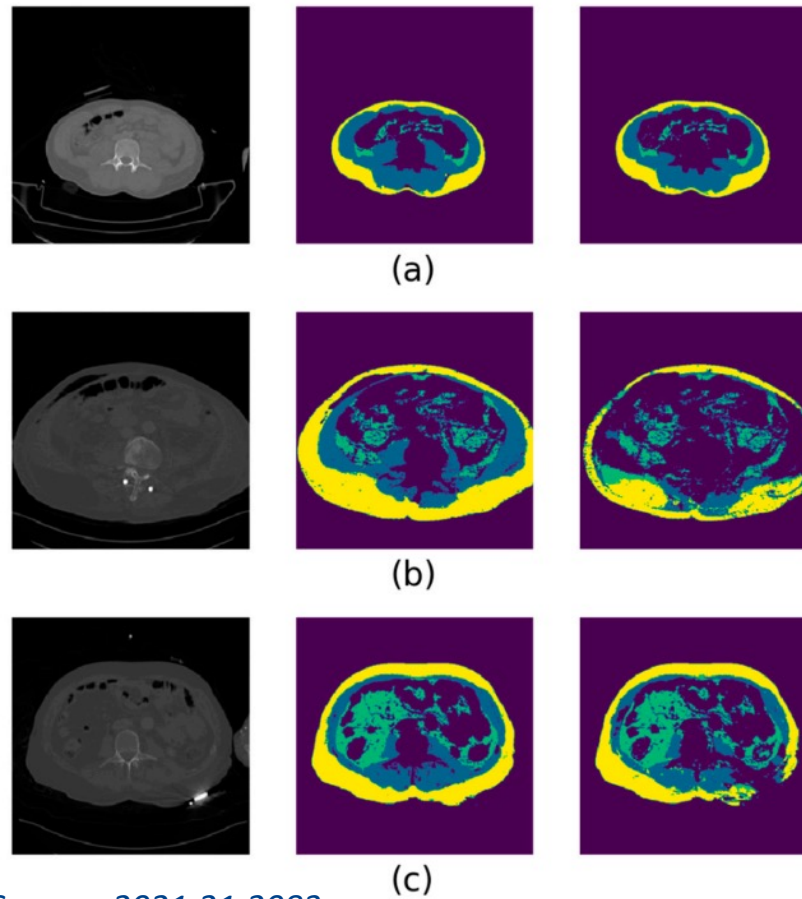
# Add data - Comorbidities Cardiac



*Zeleznik et al. Nat Commun. 2021 Jan 29;12(1):715*

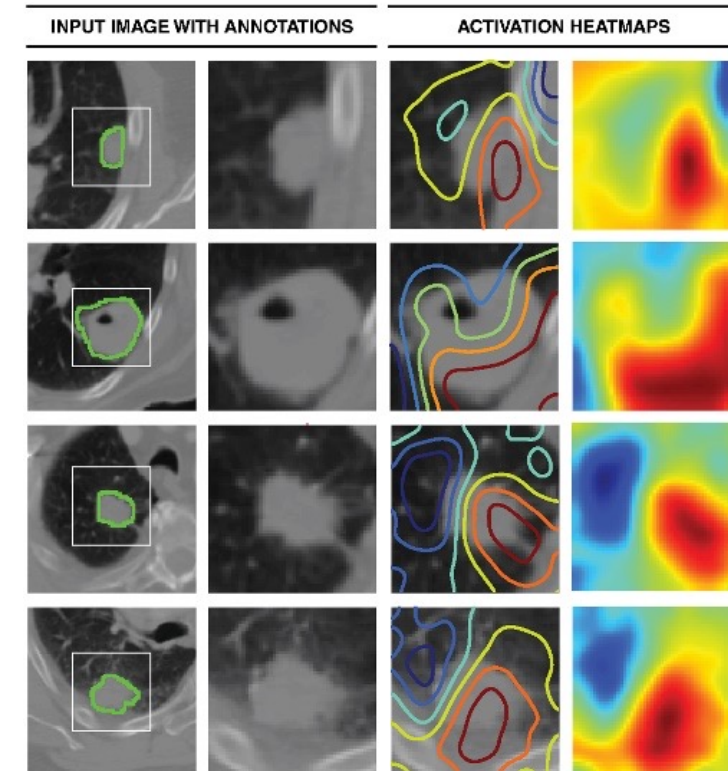
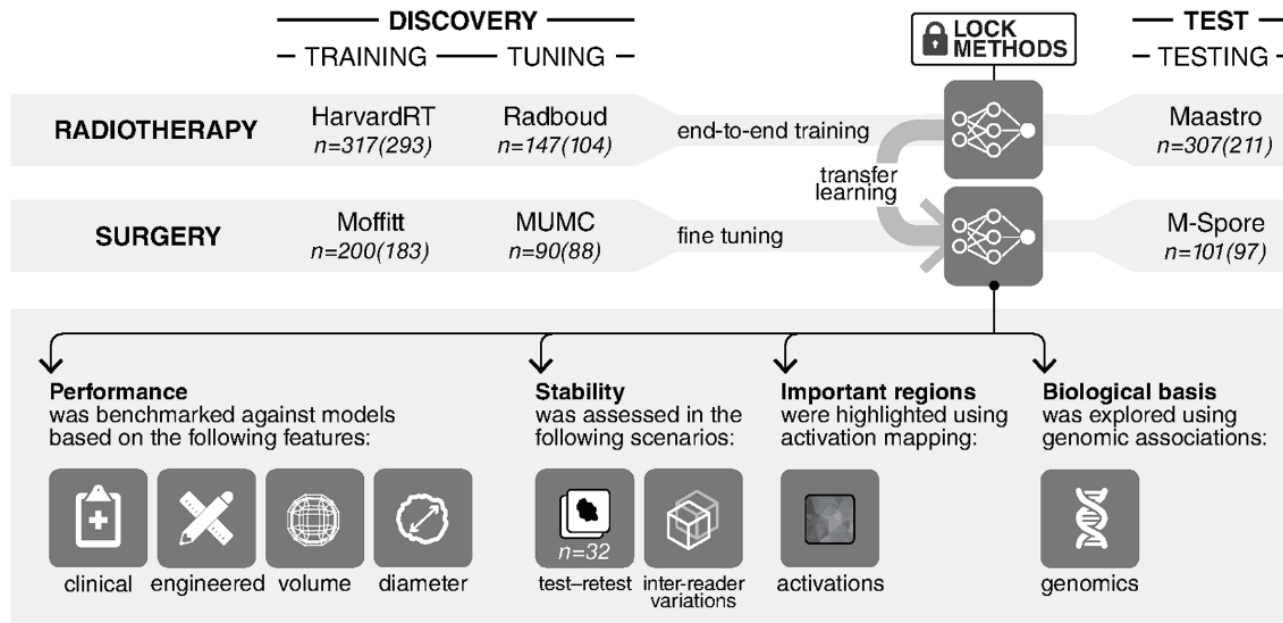
*Abravan et al. Med Phys 2020 7;78:173-178*

## Add data - Comorbidities Muscle strength, sarcopenia



*Sensors 2021;21:2083*

# Add data – Things you did not know were (that) important

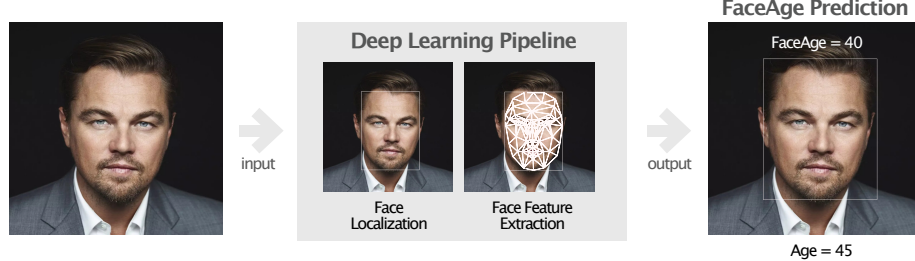


*PLoS Med* 2018 15(11): e1002711.



# Add data – Things you did not know were (that) important - FaceAge

## a | FaceAge Algorithm

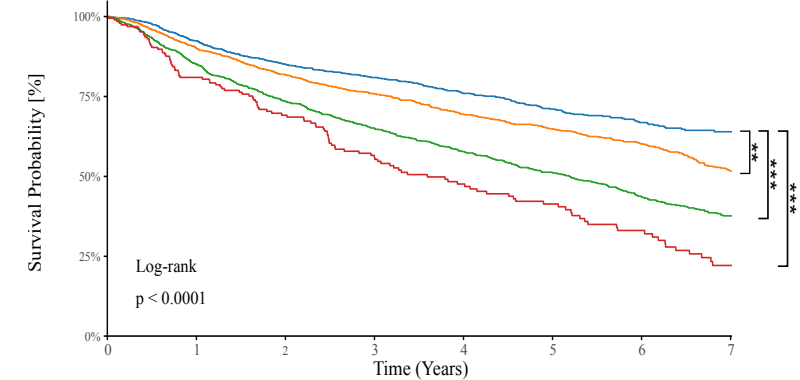
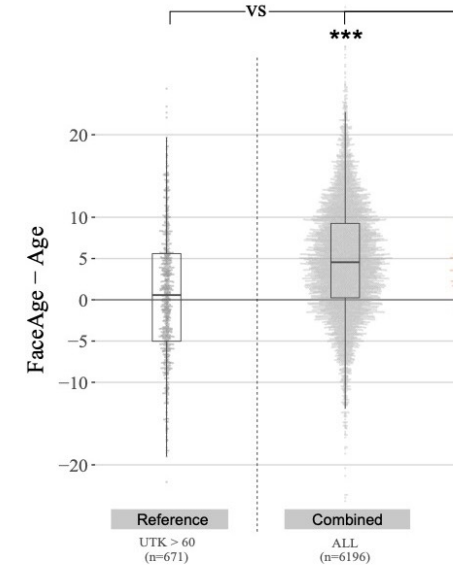


## b | Data

Discovery Datasets (n=58,851)	Clinical Validation Datasets (n=6,196)		
<b>Training Dataset   IMDB-WIKI</b> 56,304 famous and healthy individuals Manually curated and focused on older individuals	<b>MAASTRO-Biobank</b> 4,906 patients Stage I-III multiple cancer types treated curatively Basic clinical information including clinical outcomes	<b>Harvard-Thoracic</b> 573 patients Stage I-III thoracic cancer treated curatively Detailed clinical information and outcomes	<b>Harvard-Palliative</b> 717 patients Stage IV metastatic disease treated palliative Detailed clinical information and outcomes
<b>Technical Validation   UTK</b> 2,547 healthy individuals with matched ages and ethnicity labels	Publicly available data		

## c | Clinical Experiments

Human Subjects Experiments				
FaceAge vs Age	Lifestyle Factors	Cancer Types and Stages	Clinical Outcomes	Prediction Models
Expert VS FaceAge VS Expert+FaceAge 10 humans, n=100 cases (Harvard-Palliative)				



Number at risk	0	1	2	3	4	5	6	7
FaceAge ≤ 65	1385	1067	861	700	536	409	250	128
65 < FaceAge ≤ 75	2035	1565	1300	1062	815	634	419	228
75 < FaceAge ≤ 85	1292	983	791	628	487	377	259	153
FaceAge > 85	194	143	114	88	64	49	32	17

Zalay, Bontempi et al. Submitted

## Key Messages – AI application for Data Management

- Only RCT based evidence is not feasible, we need to have complementary evidence. Especially in Medical Technology.
- We need data management for AI as AI needs a lot of data
  - Federated Data Infrastructures
- AI can help in data management to
  - Extract data from unstructured text data (NLP)
  - Standardize data (OAR/GTV delineation)
  - Add/impute data retrospectively (TNM8, HPV, new OARs, comorbidities)
  - Discover new data that may be important (pleural wall, faceage)

# Acknowledgements

## Netherlands

MAASTRO, Maastricht, Netherlands  
 Radboudumc, Nijmegen, Netherlands  
 Erasmus MC, Rotterdam, Netherlands  
 Leiden UMC, Leiden, Netherlands  
 Elizabeth Twee Steden Ziekenhuis,  
 Tilburg, Netherlands  
 Catharina Hospital, Eindhoven,  
 Netherlands  
 Isala Hospital, Zwolle, Netherlands  
 NKI Amsterdam, Netherlands  
 UMCG, Groningen, Netherlands  
 IKNL, Utrecht, Netherlands

## Europe

Policlinico Gemelli & UCSC, Roma, Italy  
 UH Ghent, Belgium  
 UZ Leuven, Belgium  
 Cardiff University & Velindre CC, Cardiff,  
 UK  
 CHU Liege, Belgium  
 Uniklinikum Aachen, Germany  
 LOC Genk/Hasselt, Belgium  
 The Christie, Manchester, UK  
 State Hospital, Rovigo, Italy  
 St James Institute of Oncology, Leeds, UK  
 U of Southern Denmark, Odense,  
 Denmark  
 Greater Poland Cancer Center, Poznan,  
 Poland  
 Oslo University Hospital, Oslo, Norway

Aarhus Universitetshospital, Aarhus,  
 Denmark

Bank of Cyprus Oncology Center, Nicosia,  
 Cyprus  
 Weston Park Hospital, Sheffield, UK  
 Hull University Teaching Hospitals NHS  
 Trust, Hull, UK

Addenbrookes' Hospital, Cambridge, UK  
 Oxford University Hospitals NHS  
 Foundation Trust, Oxford, UK  
 Haukeland University Hospital, Bergen,  
 Norway

## Africa

University of the Free State,  
 Bloemfontein, South Africa

## Asia

Fudan Cancer Center, Shanghai, China  
 CDAC, Pune, India  
 Tata Memorial, Mumbai, India  
 Suining Central Hospital, Suining, China  
 HGC Oncology, Bangalore, India  
 MVRCC&NITC, Calicut, Kerala, India  
 Apollo Hospitals, Hyderabad, India  
 CMC Vellore, Vellore, India  
 Tianjin Medical University, Tianjin, China  
 Cancer Hospital of Shantou University,  
 Shantou, China

## North America

ROG, Philadelphia, PA, USA  
 MGH, BWH, Harvard, Boston, MA, USA  
 University of Michigan, Ann Arbor, USA  
 Princess Margaret CC, Canada  
 Ottawa Hospital Research Institute,  
 Ottawa, Canada

## South America

Albert Einstein, Sao Paulo, Brazil

## Australia

University of Sydney, Australia  
 Westmead Hospital, Sydney, Australia  
 Liverpool and Macarthur CC, Australia  
 ICCC, Wollongong Australia  
 Calvary Mater, Newcastle, Australia  
 North Coast Cancer Institute, Coffs  
 Harbour, Australia

## Industry

Varian, Palo Alto, CA, USA  
 Philips, Bangalore, India  
 Sohard GmbH, Fuerth, Germany  
 Microsoft, Hyderabad, India  
 Mirada Medical, Oxford, UK  
 CZ Health Insurance, Tilburg, NL  
 Siemens, Malvern, PA, USA  
 Roche, Woerden, NL



## Clinical Data Science research aims

1. Get access to all data of all people in the world
2. Learn personalized health prediction models from data
3. Apply prediction models to improve health

Cancer, Alzheimer's, Cardiovascular disease, Diabetes, Heart Failure, Parkinson's, Irritable Bowel Disease, Orthopedic Surgery, Rheumatoid Arthritis, Pediatric Surgery, Balance disorders, Hip dysplasia

Thank you for your attention