Clinical Data Science

Maastro
Maastricht University
Maastricht UMC+

Federated Learning from FAIR data

Andre Dekker
Professor of Clinical Data Science | Medical Physicist
Maastricht University | Maastro Clinic | Maastricht UMC+

From Technology to Treatment - Advancing Precision Medicine
Zürich | 29-08-2025 | 15:15 – 16:00 (35 +10)

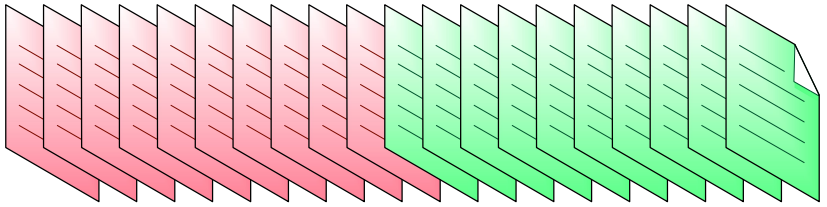
Disclosures

- Research collaborations incl. funding, consultancy and speaker honoraria
  - Pharma: Roche, Janssen, Bristol-Myers Squibb
  - MedTech/Data: Varian - Siemens, Philips, Sohard, Mirada Medical, IQVIA

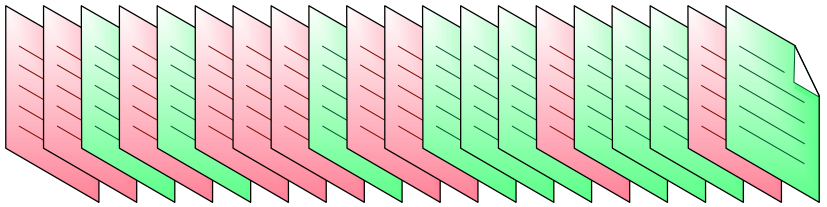
- External advisory roles
  - Care: MD Anderson Cancer Center, Peter Munk Cardiac Center
  - Charities: Hanarth Fund, NovoNordisk Foundation
  - Other: International Atomic Energy Agency, Luxemburg National Research Fund

- Spin-offs and commercial ventures
  - Maastro Innovations B.V.
  - **Medical Data Works B.V.**
  - Various patents on medical machine learning & Radiomics

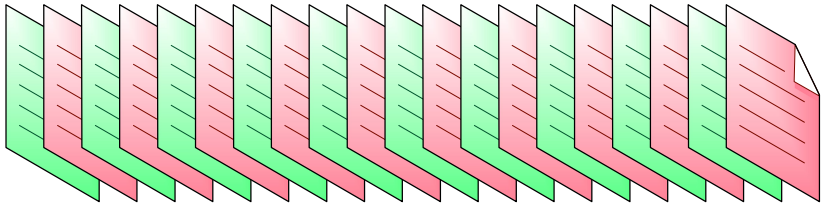
Prediction of Survival



AUC
1.00



AUC
0.72



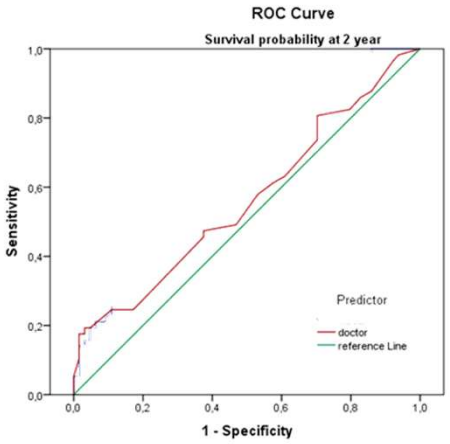
AUC
0.50

Less likely to survive

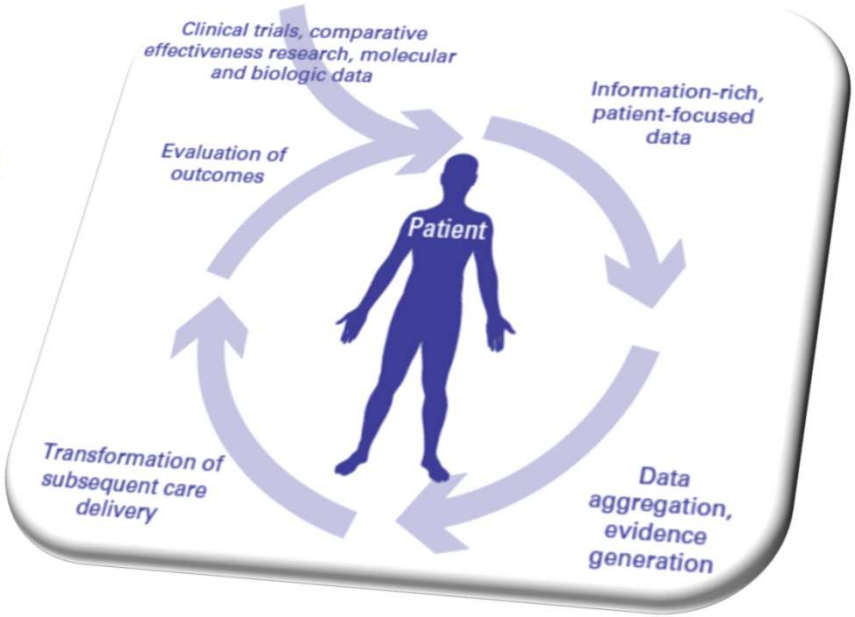
More likely to survive

Actual survival

Learning health care system



Lung Cancer
2 year survival
158 patients
5 MDs
Prospective
AUC: 0.56
2006

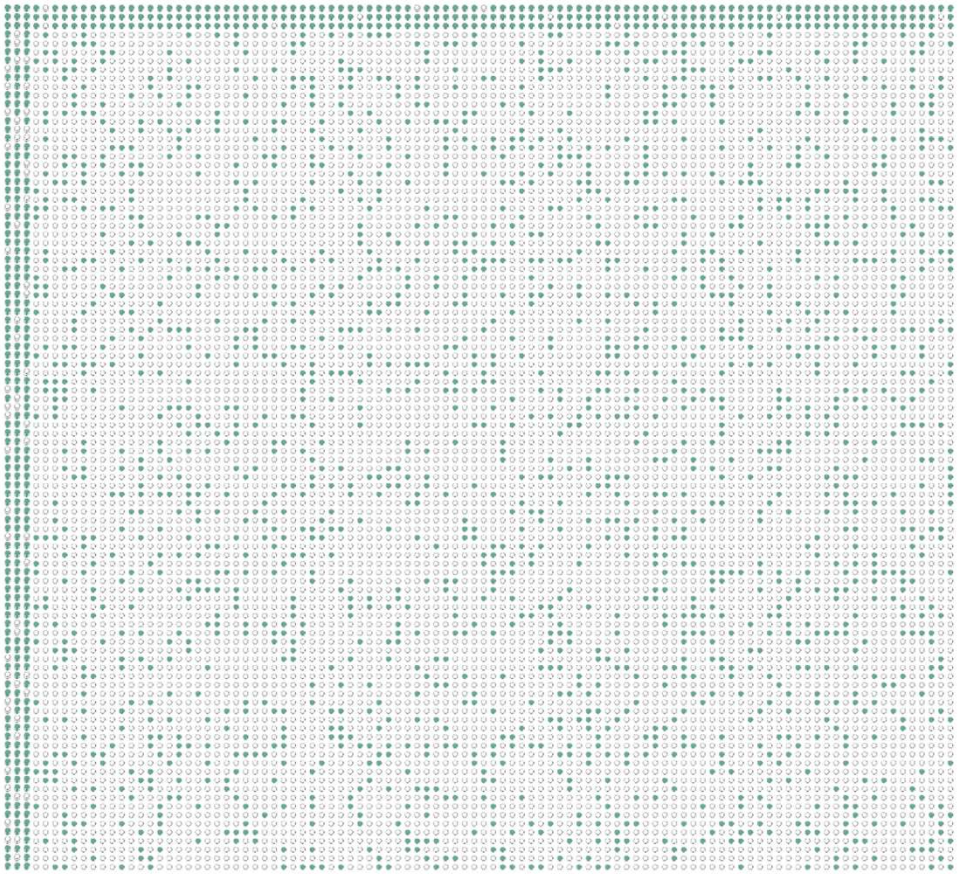


Radiotherapy and Oncology, 81, pp.S391-S392. Radiother Oncol. 2014; 112: 37–43 / J Clin Oncol 2010;28:4268

Data landscape

Data elements

Patients



- Clinical research
  - 3% of patients
  - 100% of features
  - 5% missing
  - 285 data points
- Clinical registries
  - 100% of patients
  - 3% of features
  - 20% missing
  - 240 data points
- Clinical routine
  - 100% of patients
  - 100% of features
  - 80% missing
  - **2000** data points

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

Barriers to sharing data

[..] the problem is not really technical […]. Rather, the problems are **ethical**, **political**, and **administrative**.
(Lancet Oncol 2011;12:933)

- Administrative (I don't have the resources)
- Political (I don't want to)
- Ethical (I am not allowed to)

- Technical (I can't)

A different approach

- If sharing is the problem: Don't share the data

- If you can't bring the data to the research
- You have to bring the research to the data

- Challenges
  - The research application has to be distributed (trains & track)
  - The data has to be understandable by an application (i.e. not a human) -> FAIR data stations

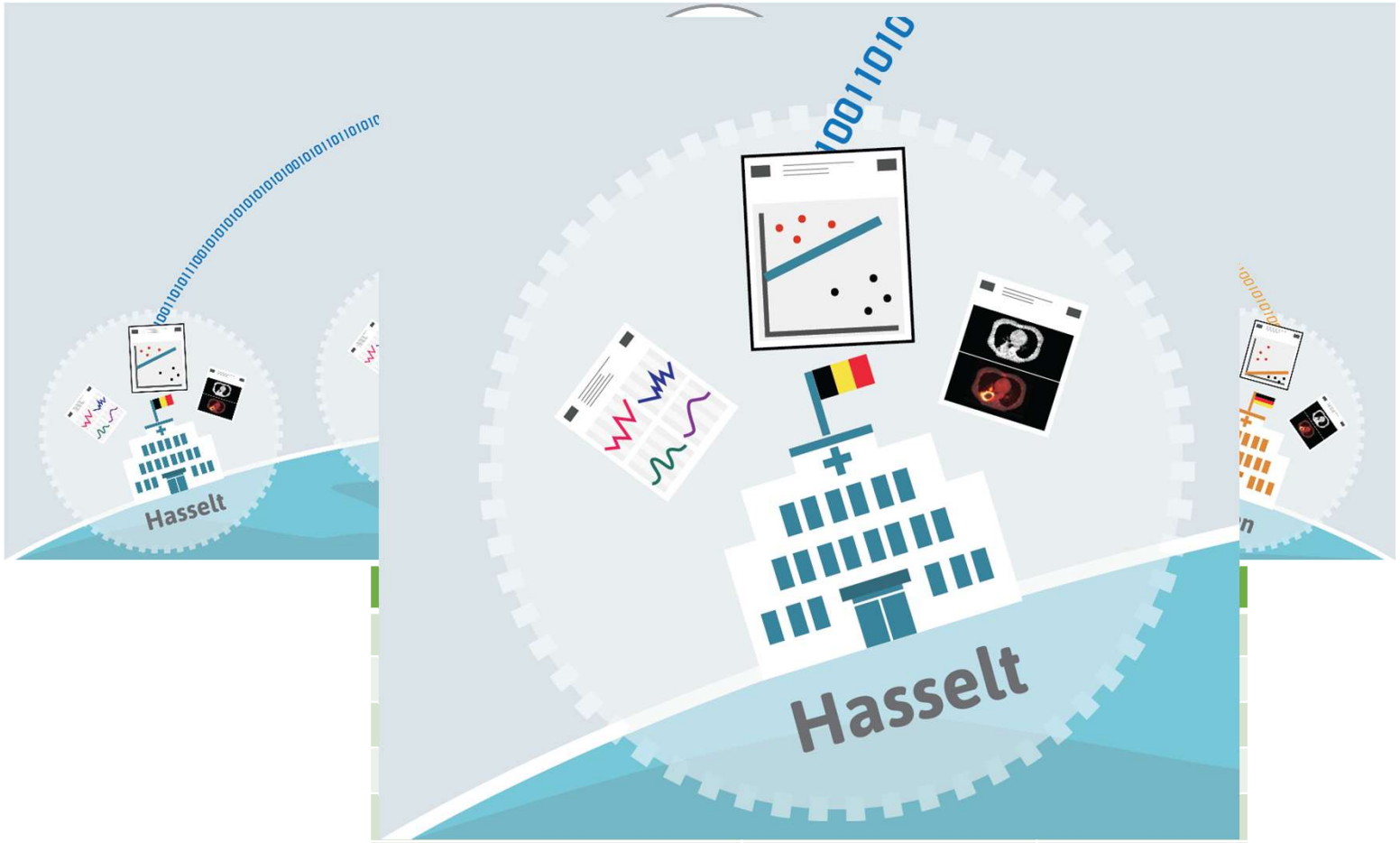
- The goals of Federated Learning (from RWD) vs. FAIR Data

- *2007: The "Computer Assisted Theragnostics" (CAT) project will use heterogeneous data from distributed databases in multiple clinical centres to develop and validate patient specific prediction models".*

PERSONAL HEALTH TRAIN

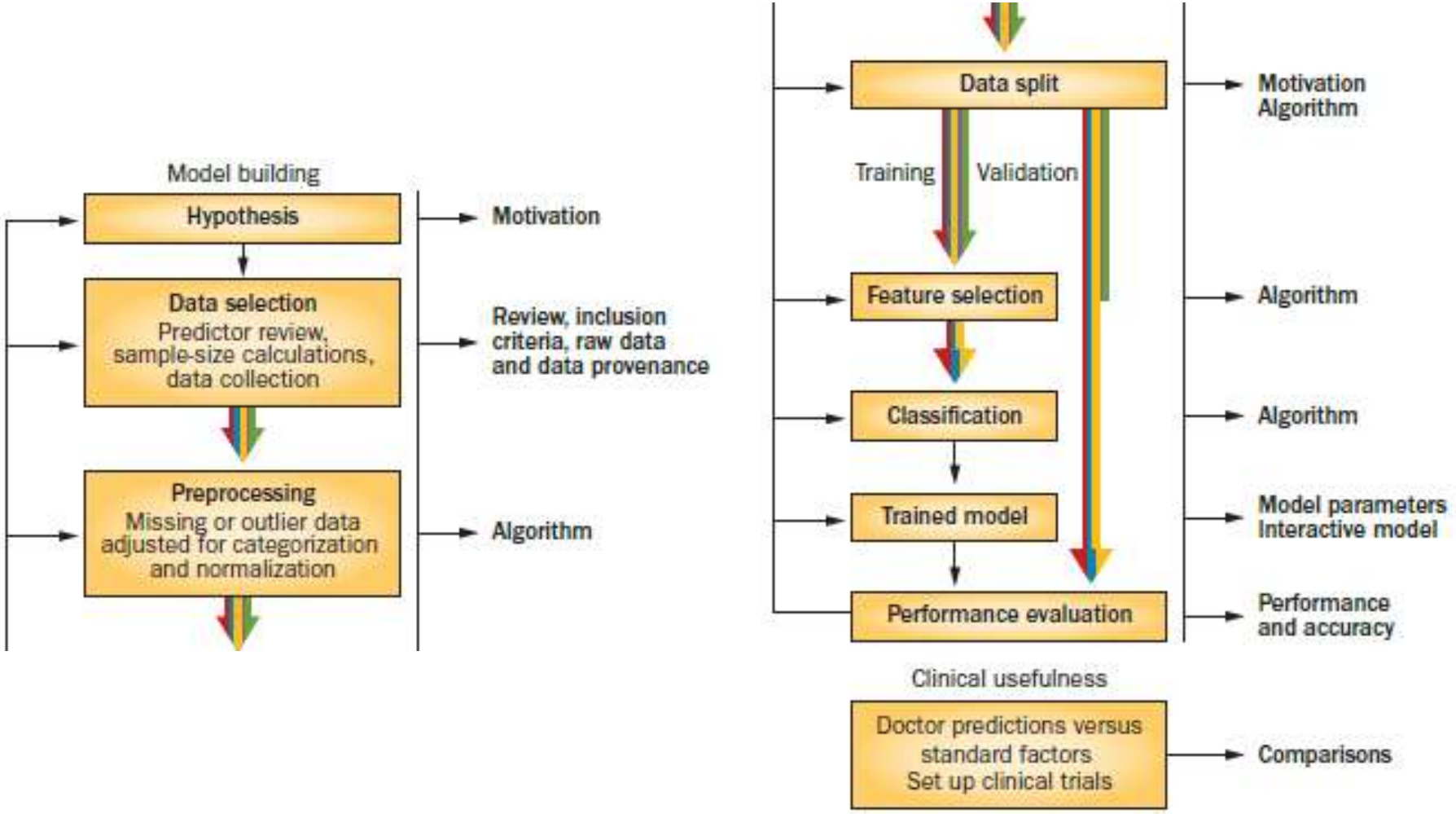
Wilkinson, DOI: 10.1038/sdata.2016.18

euroCAT example

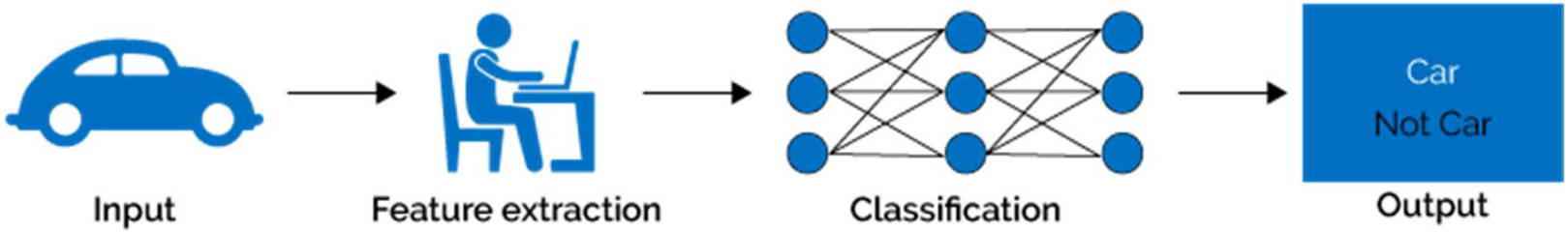


All 5 together (n=322)	World (n=*inf*)	?

General machine learning / AI approach

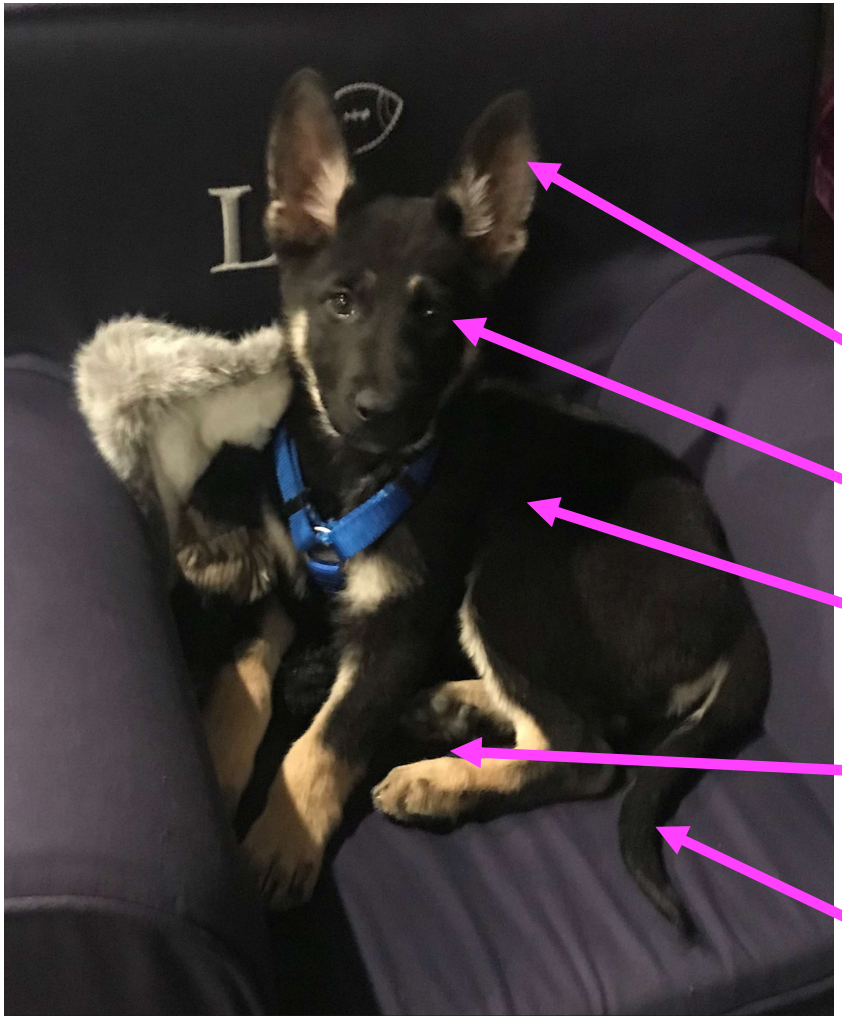


Machine Learning



Slide courtesy of Kristy Brock

Machine Learning

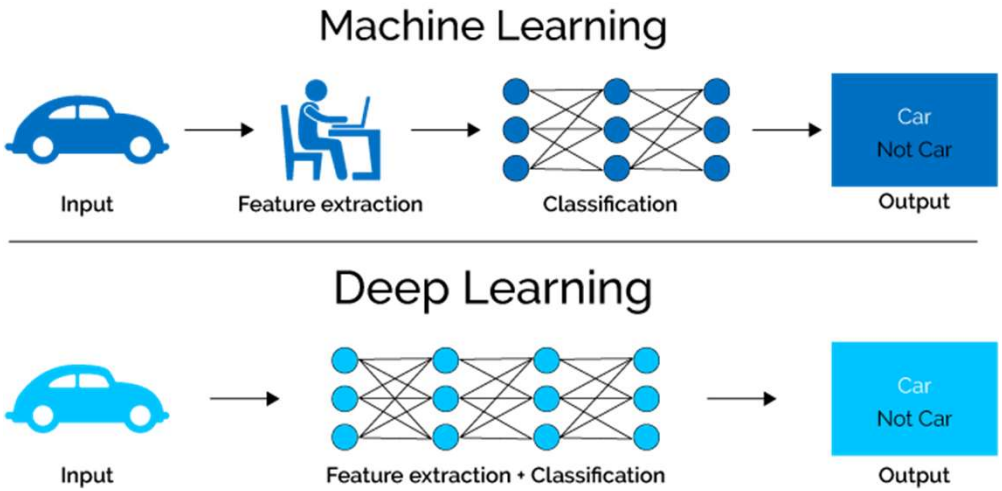


NOT A CAT!

Cat: 97%

Slide (and dog) courtesy of Kristy Brock

Deep Learning – Human out of the loop



Slide courtesy of Kristy Brock

Balance between Human and AI centricity

Table 3 | Comparisons between human evaluations and different types of AI approaches

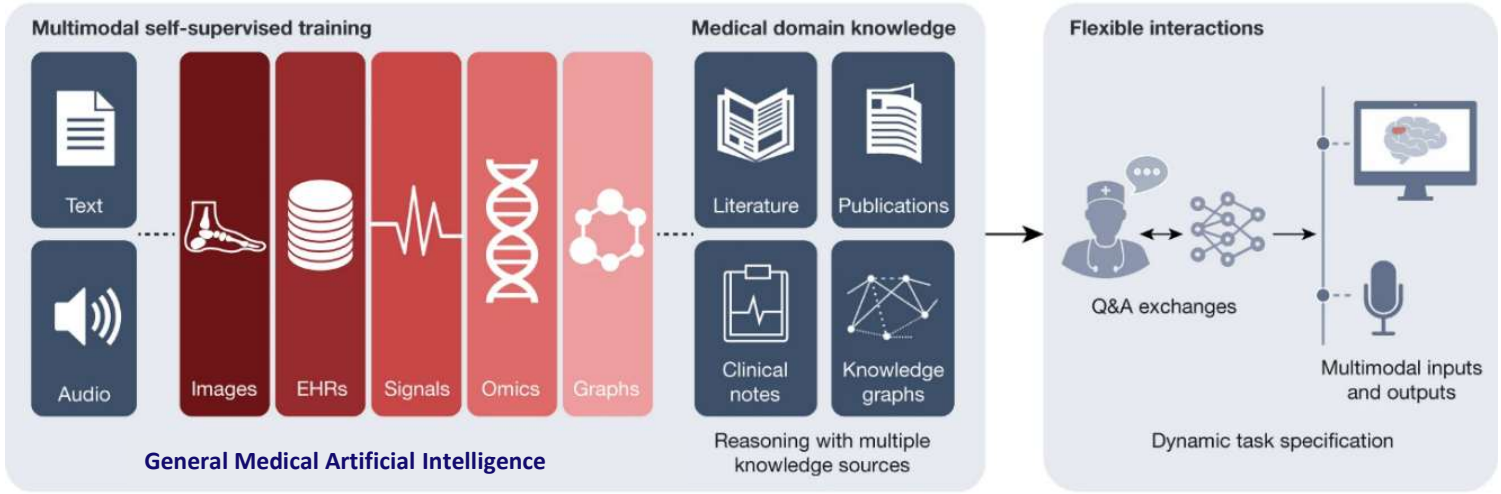
Approaches	Model comprehensibility	Performance	Reproducibility	Dependency on prior knowledge	Development and training costs[a]	Running costs	Around-the-clock availability	Update costs
Human evaluation	High	Moderate or high	Moderate	High	High	High	Low	High
Rule-based algorithms	High	Moderate or high	High	High	Moderate or high	Low	High	High
Feature-based machine-learning methods	Moderate or high	Moderate or high	High	Moderate[b]	Moderate	Low	High	Moderate[c]
Deep artificial neural networks	Low or moderate	High	High	Low	Moderate	Low	High	Low

Nat Biomed Eng 2018;2:719

Clinical Data Science

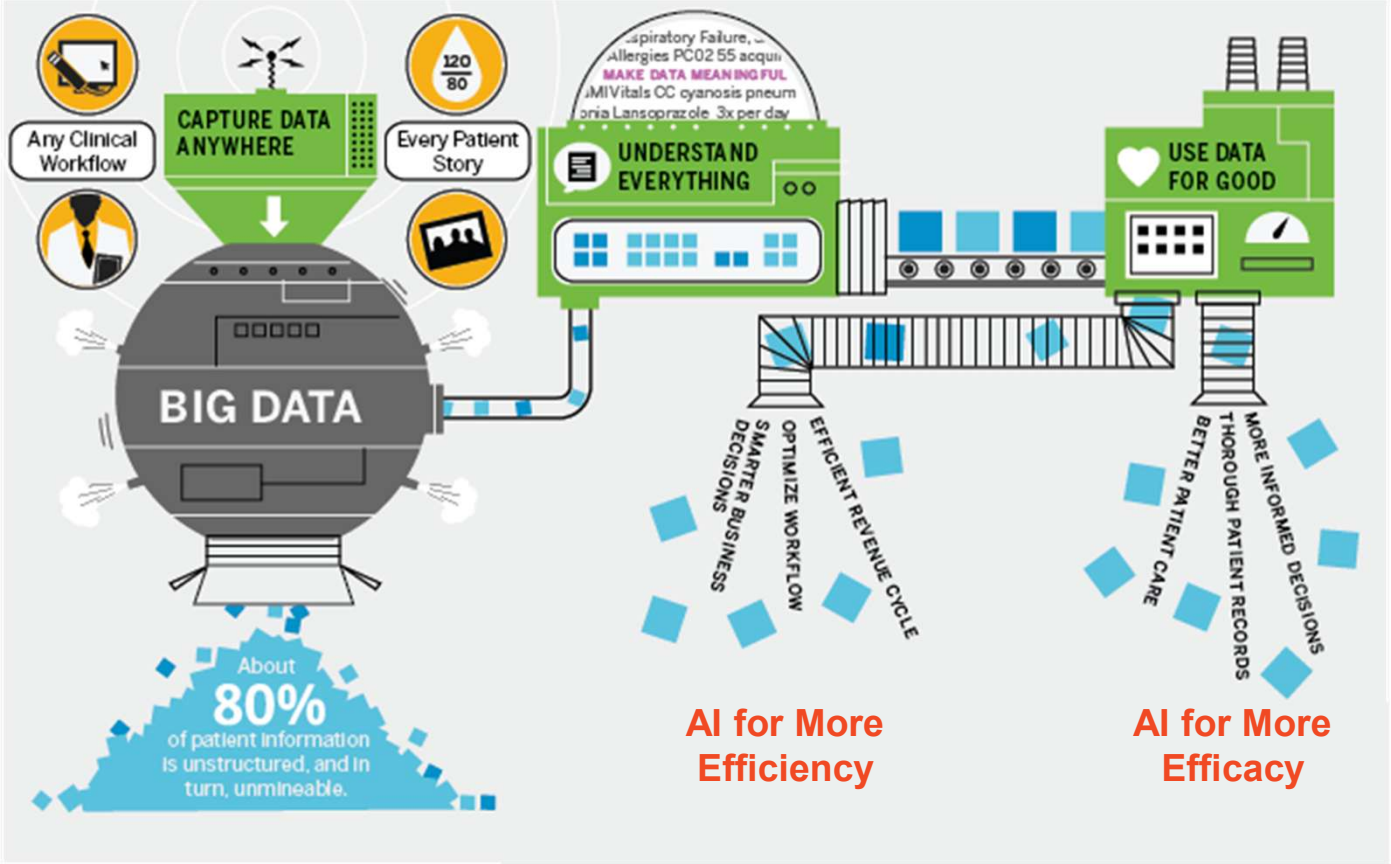
Deep Learning and Beyond
Self Supervised & Generalist & Multimodal & Interactive

Maastro
Maastricht University
Maastricht UMC+



Nature volume 616, pages 259–265 (2023)

Learning AI for Better Health Care

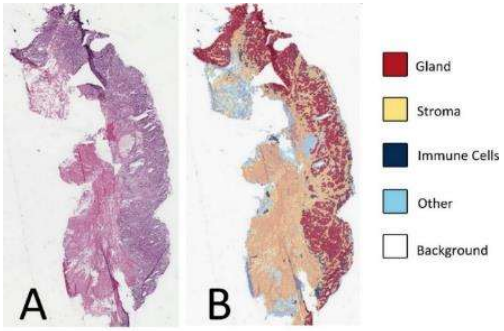


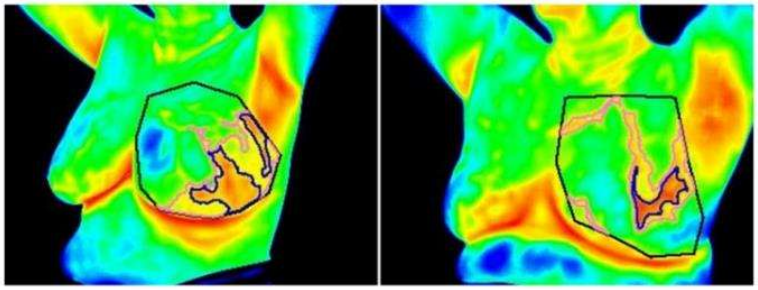
Source: nuance.com

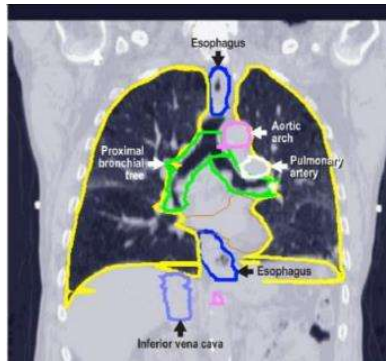
Clinical Data Science

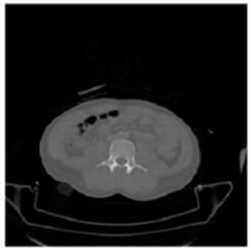
AI for Higher Efficiency

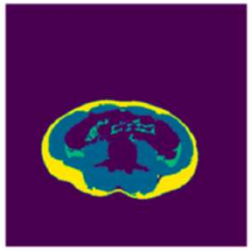
Maastro
Maastricht University
Maastricht UMC+

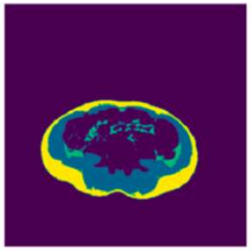


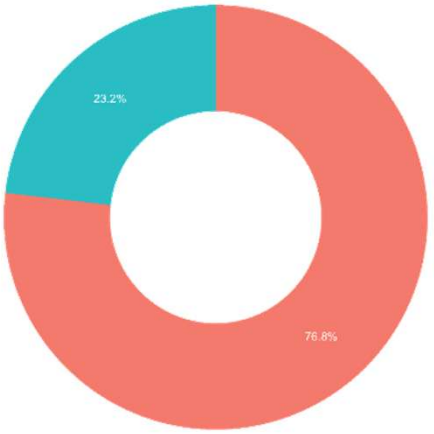












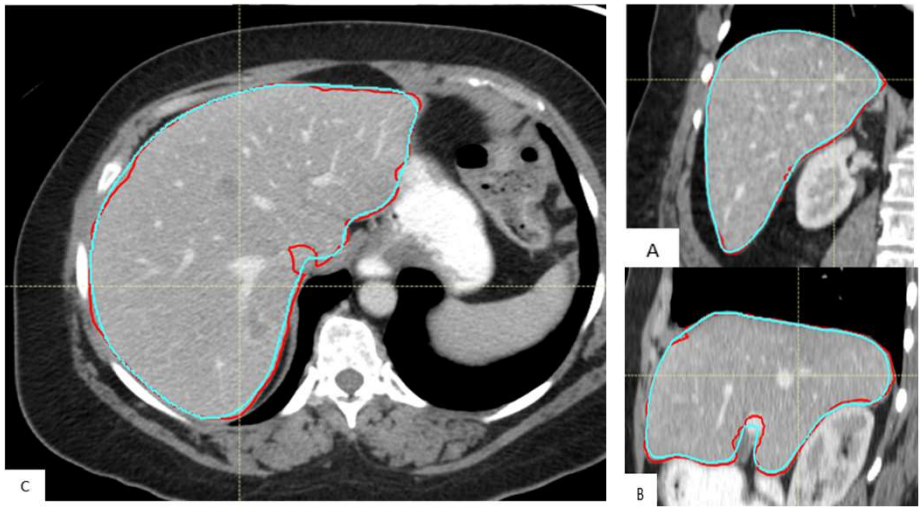
https://doi.org/10.3390/electronics13030498

Clinical Data Science

Artificial Intelligence is different from Human Intelligence

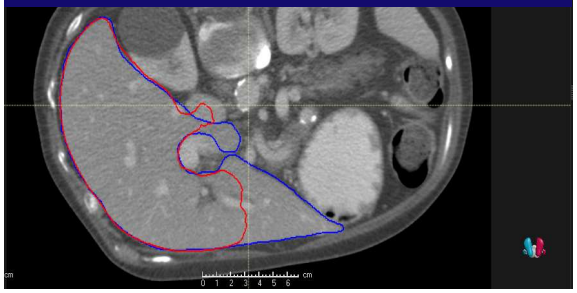
Maastro
Maastricht University
Maastricht UMC+

AI algorithms cannot be expected to perform at a higher level than their training data, but should deliver the same standard of performance consistently for data within the training space.

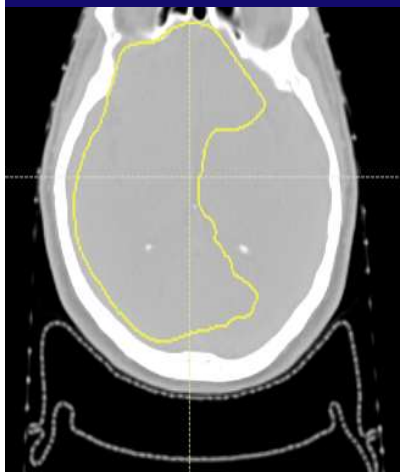


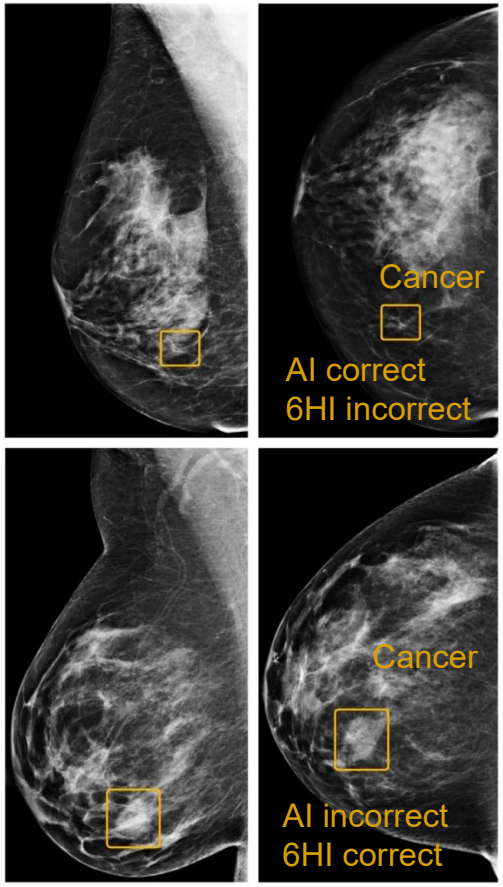
Images courtesy of MD Anderson / Brian M. Anderson Nature 2020, 577:89 | JASON, Artificial Intelligence for Health and Health Care, Dec 2017, JSR-17-Task-002

CT prone
All training data was supine



AI will find you a liver anywhere…





Learning from failure

MUMC wil Chinees algoritme inzetten bij coronabehandeling

© De Limburger

Het ziekenhuis van Maastricht (MUMC) wil een Chinees algoritme gebruiken bij te nemen beslissingen over coronapatiënten. Het algoritme is ontwikkeld op basis van patiënten in China en kan helpen om snelle keuzes te maken bij de behandeling.

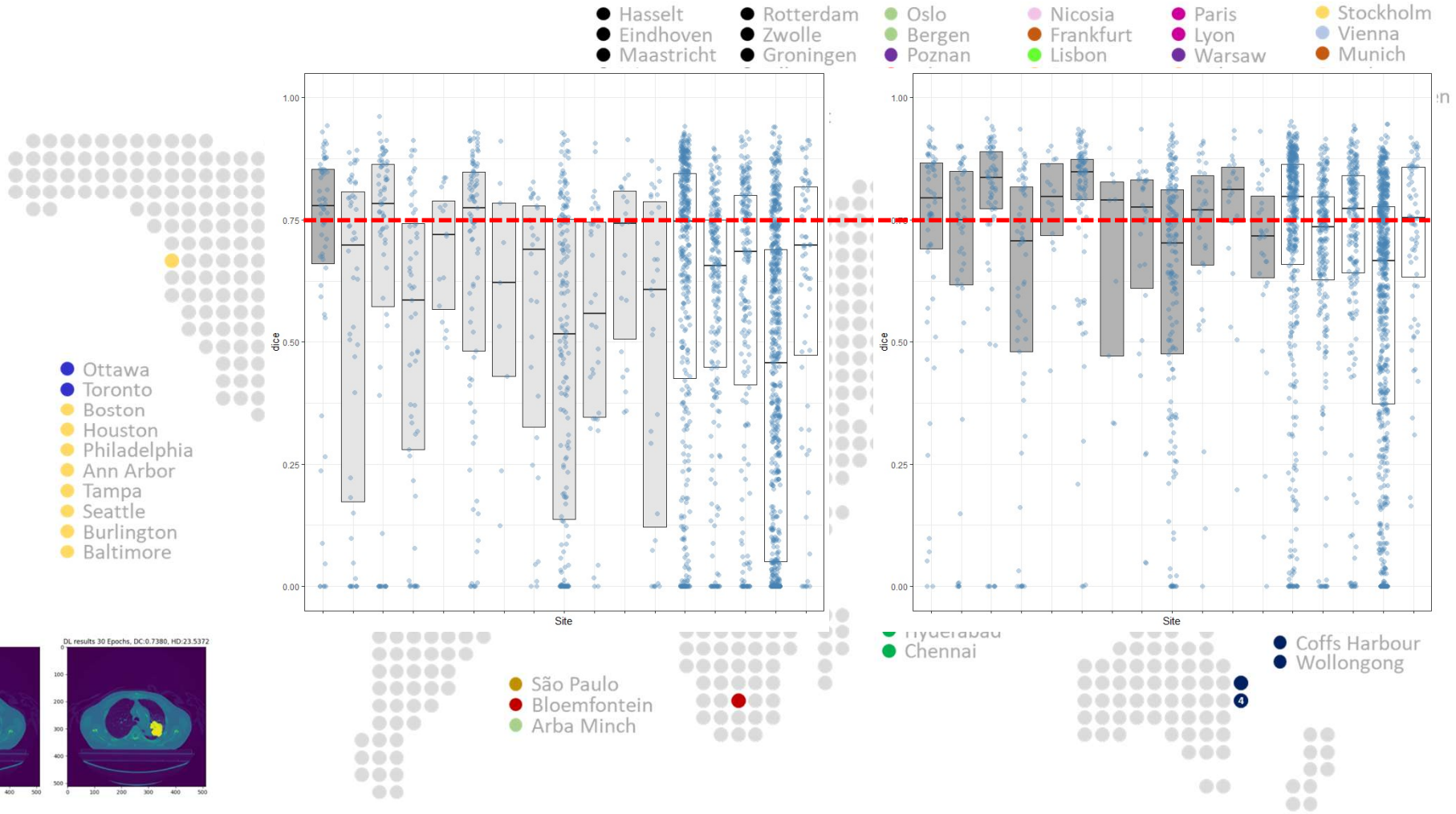
Redactie

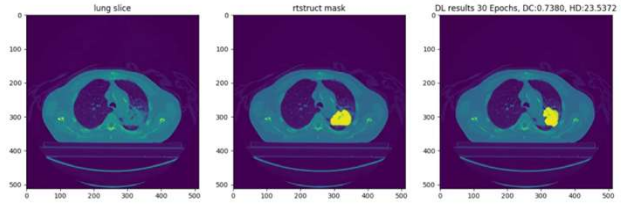
Donderdag 26 maart 2020 om 08:14

Clinical Data Science

Going global – Worldwide Federated Learning

Maastro
Maastricht University
Maastricht UMC+





Clinical Data Science

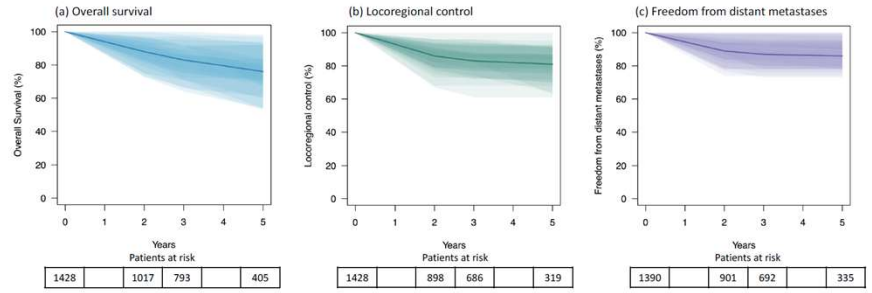
Outcomes in rare cancers

Maastro
Maastricht University
Maastricht UMC+

FEDERATED LEARNING WITH REAL-WORLD DATA: AN INTERNATIONAL MULTI-CENTRE STUDY TO DEVELOP AND VALIDATE PROGNOSTIC MODELS FOR ANAL CANCER

AUTHOR NAMES AND AFFILIATIONS

THEOPHANOUS S*,[1,2], LØNNE PI[3], CHOUDHURY A[4], BERBEE M[4], DEIJEN CL[5], DEKKER A[4], FIELD M[6,7,8,9], GAMBACORTA MA[10], GILBERT A[1,2], GUREN MG[11,12], JADON R[13], KOCHHAR R[14], MARTIN D[15,16,17], MOHAMED AA[18], MUIRHEAD R[19], PARÉS O[20], RASZEWSKI L[21], ROY R[22], SCARSBROOK A[1,2], SEBAG-MONTEFIORE D[2], SPEZI E[23], VASSILIOU V[24], MALINEN E+,[25], WEE L+,[4], APPELT A+,[1,2], ON BEHALF OF THE ATOMCAT CONSORTIUM



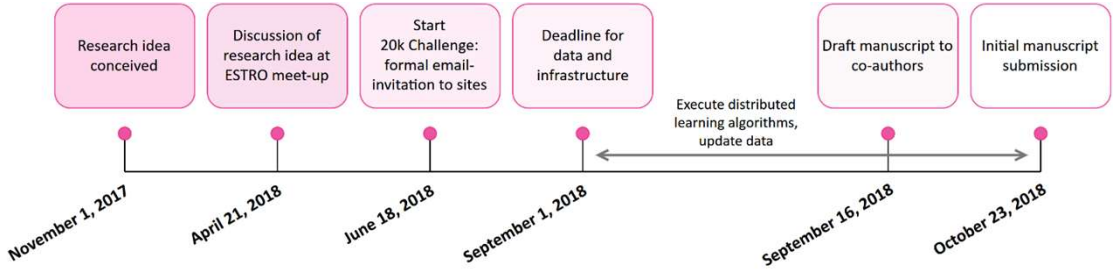
atomCAT2, 3->16 cancer centers

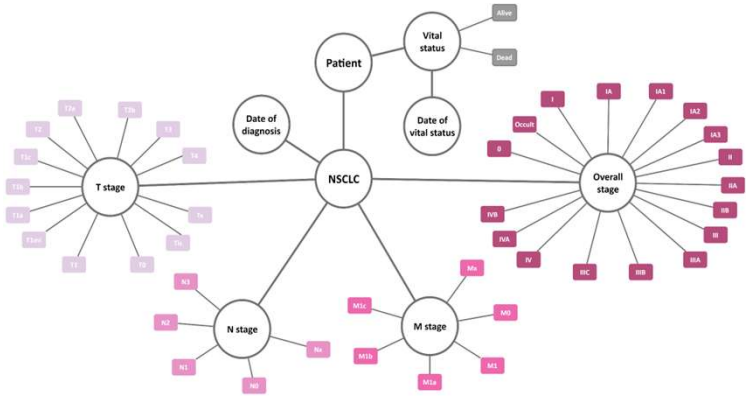
- Leeds, UK
- Oslo, Norway
- Maastricht, Netherlands
- Hull, UK
- Amsterdam , Netherlands
- Nicosia, Cyprus
- Cardiff, UK
- Lisbon, Portugal
- Rome, Italy
- Poznan, Poland
- Manchester, UK
- Oxford, UK
- Aachen, Germany
- Cambridge, UK
- Liverpool, Australia

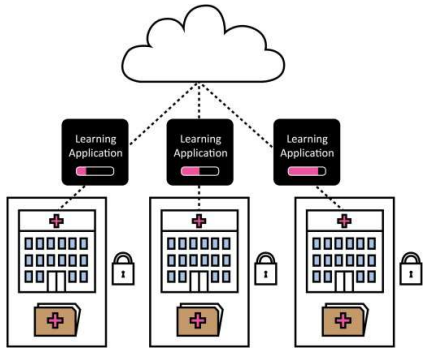
281 -> 1605 patients

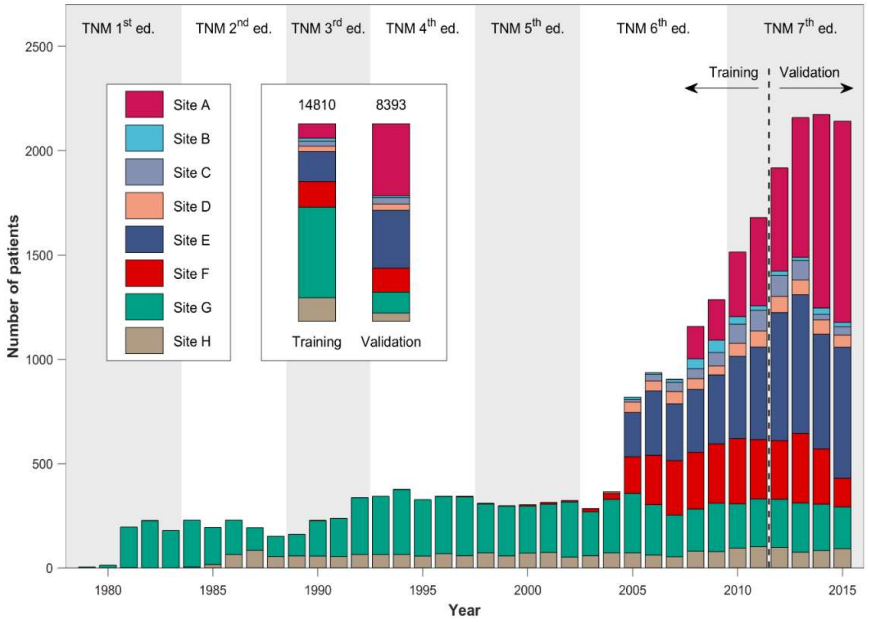
20k challenge

- Maastricht, Amsterdam, Cardiff, Nijmegen, Manchester, Rome, Rotterdam, Shanghai
- 37090 NSCLC patients







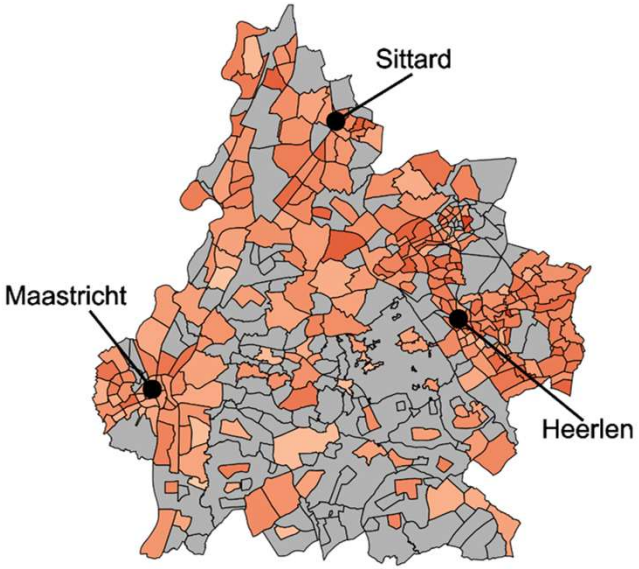


Deist, T. M. *et al.* Distributed learning on 20 000+ lung cancer patients – The Personal Health Train. *Radiotherapy and Oncology* **144**, 189–200 (2020).

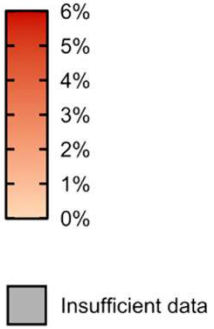
Clinical Data Science

Holistic Models – Vertical Partitions

Maastro
Maastricht University
Maastricht UMC+



5-year atherosclerotic cardiovascular disease (ASCVD) event rate



Standardised predictor	Subdistribution hazard ratio
male	2.12 (95% CI: 2.03 – 2.21)
age	1.80 (95% CI: 1.76 – 1.84)
wealth	0.96 (95% CI: 0.93 – 0.99)
education: low	1.04 (95% CI: 0.99 – 1.10)
education: high	0.90 (95% CI: 0.85 – 0.94)
particulate matter 2.5	1.13 (95% CI: 1.08 – 1.18)

AUC~0.68

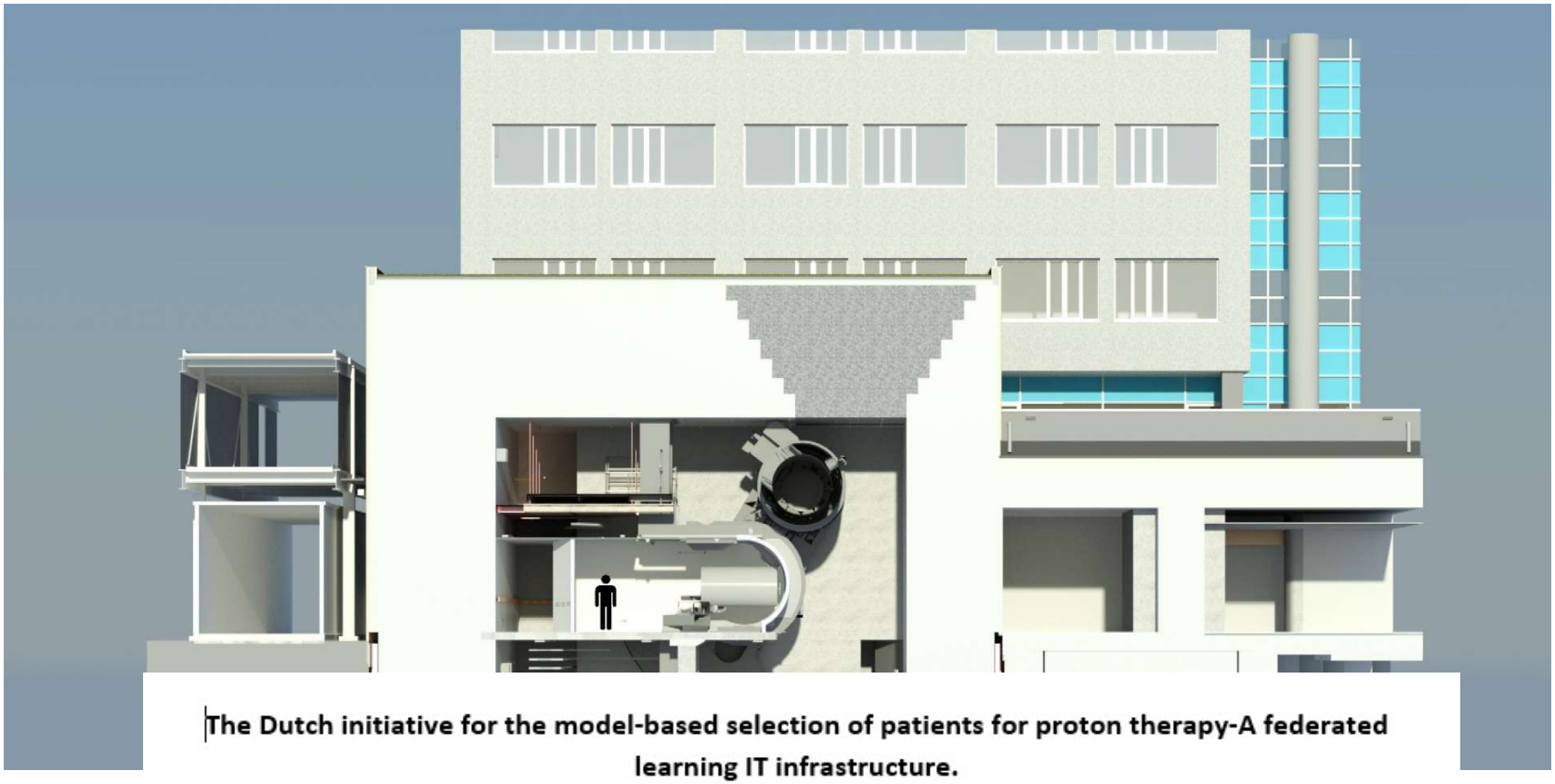
Statistics NL + Dutch Hospital Data + RIVM (environment)
(N=307,189)

CARRIER
Big data for prevention of cardiovascular disease

Clinical Data Science

Model Based Indications

Maastro
Maastricht University
Maastricht UMC+



The Dutch initiative for the model-based selection of patients for proton therapy-A federated learning IT infrastructure.

Petros Kalendralis[1], Matthijs Sloep[1], Jasper Snel[1,2], Nibin Moni George[1], Johannes A. Langendijk[3], Martijn Veening[3], Andre Dekker[1,2], Johan Van Soest[2,1], Rianne Fijten[1]

Key Message & Open issues

- We need AI for higher efficiency and higher efficacy in health care
- Federated Learning is quickly becoming the standard

Challenges

- Making data FAIR at the source
  - What? When? How?
- FAIR “trains”
- Legal/governance still taking long
- Solving privacy made political barriers evident
- Trusted trains / Federated learning library
- Different federated learning implementations

Coming up – Ongoing work

- Networks of hospitals start federated learning (FHIN, DHD, Levilo, Health Alliance Limburg) + GP networks
- Large health data infrastructure initiatives for secondary use (Health-RI, DIGIONE) – EHDS driven
- Federated learning of AI initiatives – Interoperability of infrastructure (Health-AI / LIFES)
- Specific federated learning initiatives
  - Cancer, Dementia, Epilepsy, Heart failure, Mental disease, Rheumatoid Arthritis, Bowel Disease, Dermatology, Radiology, Ophthalmology, Orthopedics, Pharmacy, Crime (~20 FL projects)
  - Vision transformers / foundation models
  - Pathology, genomics, free text

Clinical Data Science

You can’t do this alone

Maastro
Maastricht University
Maastricht UMC+

Netherlands
- Maastro Clinic, Maastricht, Netherlands
- Maastricht UMC+, 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
- AZ Delta, Roeselare, Belgium
- Cardiff University & Velindre CC, Cardiff, UK
- CHU Liege, Belgium
- Grand Hôpital de Charleroi, 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
- Arba Minch University, Arba Minch, Ethiopia
- University of Nairobi, Nairobi, Kenya

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
- Guangdong Provincial People’s Hospital, Guangzhou, China
- Zhejiang Cancer Hospital, Hangzhou, China

North America
- RTOG, 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
- IQVIA, London, UK