

Federated learning: Experience from The Netherlands

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

ATHENA symposium: Data driven Innovation in Personalized Medicine and Care Leuven | November 23, 2023 | 16:10-16:30 (15+5)







Disclosures

Research collaborations incl. funding, consultancy and speaker honoraria

- Pharma: Roche, Janssen, Bristol-Myers Squibb
- MedTech/Data: Varian Siemens, Philips, Sohard, 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







Clinical Data Science @ Maastricht

- Develop global FAIR federated health data sharing infrastructures
- Use Artificial Intelligence methods to learn from health data
- Apply Artificial Intelligence to improve health

Cancer, Alzheimer's, Heart disease, Diabetes, Parkinson's, Irritable Bowel Disease, Rheumatoid Arthritis, COPD, COVID, Pediatric Surgery



https://www.clinicaldatascience.nl/







Prediction of survival

Less likely to survive

Maastricht UMC+



Actual survival

Maastro



4

Learning health care system



963 poster

MODELS COMBINING CLINICAL DATA WITH MEDICAL KNOWLED-GE ARE MORE ACCURATE THAN DOCTORS IN PREDICTING NSCLC AFTER (CHEMO-)RADIOTHERAPY



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







The main problem of health data is that it is fragmented





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







Barriers to sharing data

[..] the problem is not really technical [...]. Rather, the problems are **ethical**, **political**, **and administrative**.

Lancet Oncol 2011;12:933

- 1. Administrative (I don't have the resources)
- 2. Political (I don't want to)
- 3. Ethical (I am not allowed to)
- 4. 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

PERSONAL

HEALTH TRAIN

Maastro



Wilkinson, DOI: 10.1038/sdata.2016.18

Maastricht University





A short history of the PHT

- 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".
- 2008: euroCAT project (with Siemens)
- 2015: Leiden FAIR data, Movie & "Personal Health Train" + KNAW Pitch -> Health RI, NFU Data4LifeSciences
 2018: VWS letter, KIA Topsector, PHT Community
 2019: Health-RI roadmap & funding (NFU)
 2020: NL AI Coalition, Groeifonds, Computable award

2023: Part of the VWS Missions

aastricht UMC+



aastro







euroCAT example









100 -

12

T2 T3 T4 T1

Track: Open Source

Resources V Documentation V Legal V

Nodes:

MAASTRO Private - Flyover Node MAASTRO Public - Flyover Node MD Anderson - Flyover Node D Montreal - Flyover Node Z Toronto - Flyover Node

Choose an option: T-stage







VANTAGE





Stations: 20k challenge and beyond

Table 2

Overview of tools used to make data FAIR. EMR: electronic medical records.

Provider	Amsterdam (NL)	Cardiff (WAL)	Maastricht (NL)	Manchester (ENG)	Nijmegen (NL)	Rome (IT)	Rotterdam (NL)	Shanghai (CN)
Source systems	NKI-AVL Tumour registry	Canisc (Cancer Network Information System Cymru, NHS Wales Information Services)	HiX (Chipsoft, Netherlands), municipality population registry (survival data)	Clinical Web Portal (in house e-records system). Mosaiq radiotherapy oncology information system. Medway Sigma BI patient administration system.	Radiotherapieweb (in-house EMR), municipality population registry (survival data)	BOA [12] and Speed RO	OpenClinica, Microsoft Access	Chinese EMR
ETL tools	MS SSIS	MATLAB	SAP Business Objects, MATLAB	Pentaho data integration, SQL, Java, Python, R	PHP, SQL, MATLAB	SQL	MATLAB	In-house software
Data warehouse Mapping Graph database	MS SQL Server	MS SQL Server	SAP Business Objects	PostgreSQL D2R Blazegr	SQL Server Q raph	SQL Server		None





Maastro



14

Current Federated Learning projects

- 1. Decisions support systems for <u>head&neck cancer</u> patients (TRAIN)
- 2. Al delineation of a lung cancer (ARGOS)
- 3. Predict outcome in <u>anal cancer</u> (atomCAT)
- 4. See that <u>Alzheimer's</u> starts (NCDC)
- 5. Predict prognosis for brain metastases (AMICUS)
- 6. Predict treatment sensitivity of <u>lung cancers</u> (Health-AI)
- 7. Personalized recommendation to prevent <u>cardiac disease</u> (CARRIER)
- 8. Count the number of <u>appendicitis</u> patients currently seen by all general practitioners (GRIP3)
- 9. Predict prognosis of <u>Alzheimer's (EPND)</u>
- 10. Predict if proton therapy works in **head & neck cancer** (PROTRAIT)
- 11. Predict outcomes in adolescents and **young adults with cancer** (STRONG-AYA)
- 12. Predict lymphoedema in breast cancer radiotherapy (PRE-ACT)
- 13. Use Radiomics to predict outcomes in <u>lung cancer</u> (TOTALRADIOMICS)
- 14. Predict cardiovascular risk in Chronic Myeloid Leukemia patients (CMYGUIDELINE)
- 15. See if <u>COVID</u> changed cancer care in Europe (DIGIONE)
- 16. Predict outcomes in difficult to treat <u>Rheumatoid Arthritis</u> (STRATA-FIT)
- 17. Predict if an artwork is used by <u>criminals</u> to pay each other (PRICELESS)
- 18. Combine data from various municipalities to detect **poverty** (ELSA-LAB)





aastro



Medical Data Works





Rare cancers

Predicting outcomes in anal cancer patients using multi-centre data and distributed learning (N=281)^{*}

Training nodes	MAASTRO Oslo	Leeds Oslo	Leeds MAASTRO
Validation node	Leeds	MAASTRO	Oslo
High risk disease (compared to low risk disease)	2.52 (0.93-6.78)	1.96 (0.68–5.67)	1.85 (0.71–4.86)
Male sex (compared to female sex)	3.59 (1.55–8.33)	3.83 (1.57–9.37)	2.12 (0.92–4.90)
Age at the start of RT	1.10 (0.74–1.64)	1.47 (0.99–2.17)	1.48 (1.05–2.10)
Primary tumour GTV	1.04 (1.00–1.08)	1.08 (1.03–1.13)	1.07 (1.03–1.11)
Primary tumour dose (EQD2)	0.97 (0.46–2.04)	0.35 (0.14–0.87)	0.97 (0.59–1.59)
Validation c-index	0.70	0.73	0.68
Cox Model trained on all data :	0.72	0.74	0.70

* Radiotherapy and Oncology (2021) v159 p183-189, https://doi.org/10.1016/j.radonc.2021.03.013

Aaastricht UMC+





- Leeds, UK
- Oslo, Norway
- Maastricht, Netherlands
- Hull, UK
- Amsterdam , Netherlands
- Nicosia, Cyprus
- Cardiff, UK

Leeds no

- Lisbon, Portugal
- Rome, Italy
- Poznan, Poland
- Manchester, UK
- Oxford, UK
- Aachen, Germany
- Cambridge, UK
- Liverpool, Australia 1605 patients



Model Based Indication – Federated Learning



Petros Kalendralis¹, Matthijs Sloep¹, Jasper Snel^{1,2}, Nibin Moni George¹, Johannes A. Langendijk³, Martijn Veening³, Andre Dekker^{1,2}, Johan Van Soest^{2,1}, Rianne Fijten¹







Open issues

- Making data FAIR (->working on better tools)
- Legal/governance still taking long (->we have standardized a lot)
- Trusted trains / Federated learning library (->working on a this library)
- Different federated learning implementations
- Paying for data, data market, de-risk data sharing, forcing data sharing (-> commercial model in development with IQVIA)
- Data trace
- Vertical partitions with large N
- MPC with complex machine learning
- Consent, automated ethics (-> working on it in EU project)
- N=1 stations, rare diseases

Waastricht UMC+







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, University of Michigan, Ann Arbor, USA 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,

Africa

Norway

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 Siemens, Malvern, PA, USA Cancer Hospital of Shantou University, Shantou, China

North America RTOG, Philadelphia, PA, USA MGH, BWH, Harvard, Boston, MA, 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 Roche, Woerden, NL



· () () () 😑 🕘 🗐 🕲 😄 🥌 🎯 () () 🗁 🚍 🕲 🚍 🕲 💭

Clinical Data Science research aims

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

www.clinicaldatascience.nl









Thank you for your attention





