The TRUSTroke Project

How AI and a CERN federated learning platform can assist clinician in the management of stroke patients

HORIZON-HLTH-2022-STAYHLTH-01-04

Pietro CALIANDRO – Policlinico Gemelli (Italy)

Luigi SERIO – CERN (Switzerland)

29.09.2023

trustroke

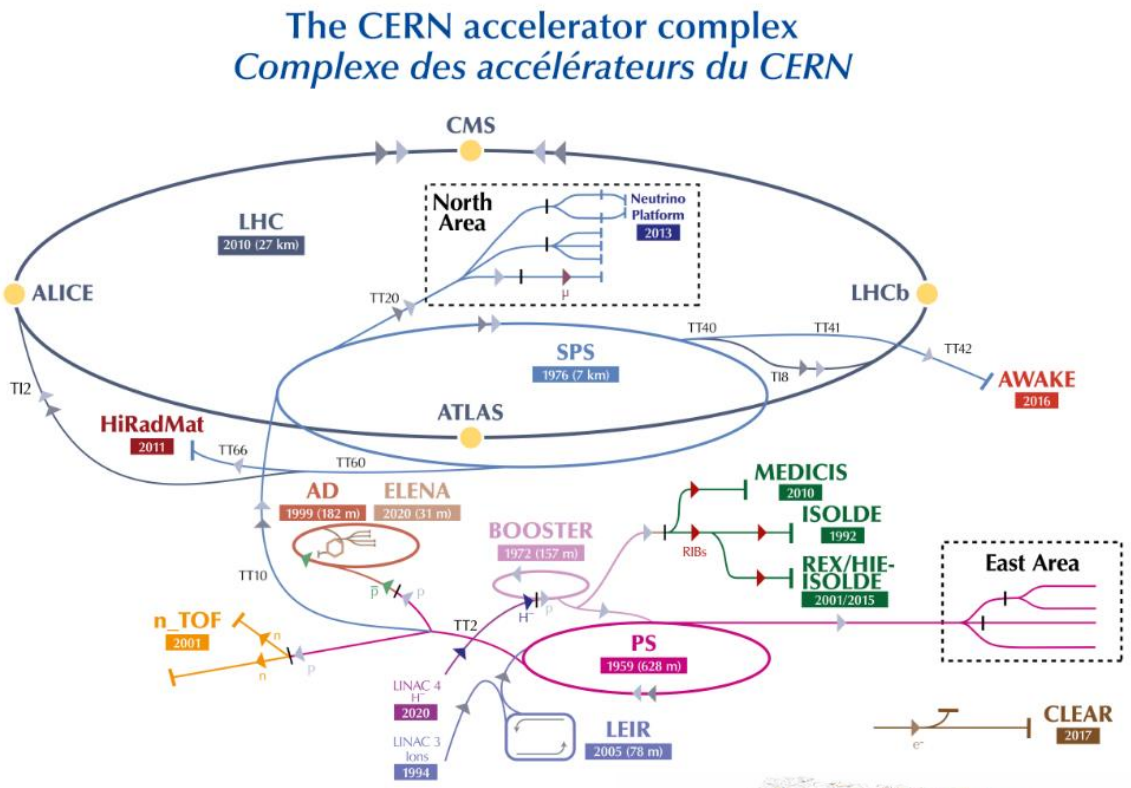
Table of Contents

Table of Contents

CERN - the world’s biggest laboratory for particle physics

International Organization established on 1 July 1953 - “Science for Peace”

CERN accelerator complex



Complex system of systems

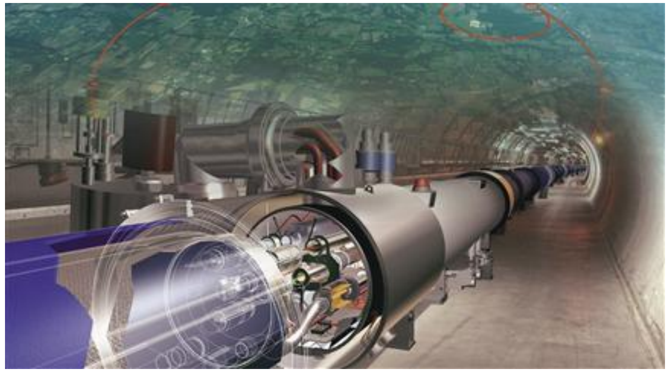
extending over 27-km circumference, 100-m underground,
equipment generating huge amounts of data

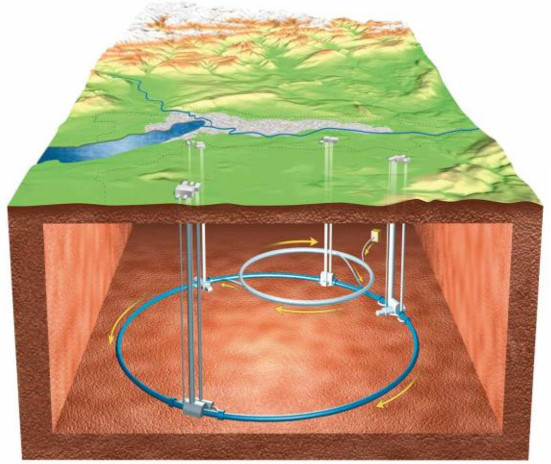
Optimisation, reliability and sustainability of the operation and diagnosis of a complex systems of systems

Critical review of centralised versus decentralised computing and data storage

New paradigm: keep data on edge devices or locally, distribute models parameters and knowledge

elegant and cost efficient,
optimise computing,
minimise communication & storage,
guarante privacy, data protection and model robustness





A laboratory for people around the world
In 1954 CERN had 12 Member States (23 in 2023)

Geographical & cultural diversity
Users of **110 nationalities**
19.4% women

CERN's annual budget is ~1200 MCHF (equivalent to a medium-sized European university)

As of 31 December 2021
Employees:
2676 staff, **783** fellows

Associates:
11 175 users, **1556** others

Contractor's employees: ~2000

23 Member States
Austria – Belgium – Bulgaria – Czech Republic
Denmark – Finland – France – Germany – Greece
Hungary – Israel – Italy – Netherlands – Norway
Poland – Portugal – Romania – Serbia – Slovakia Spain –
Sweden – Switzerland – United Kingdom

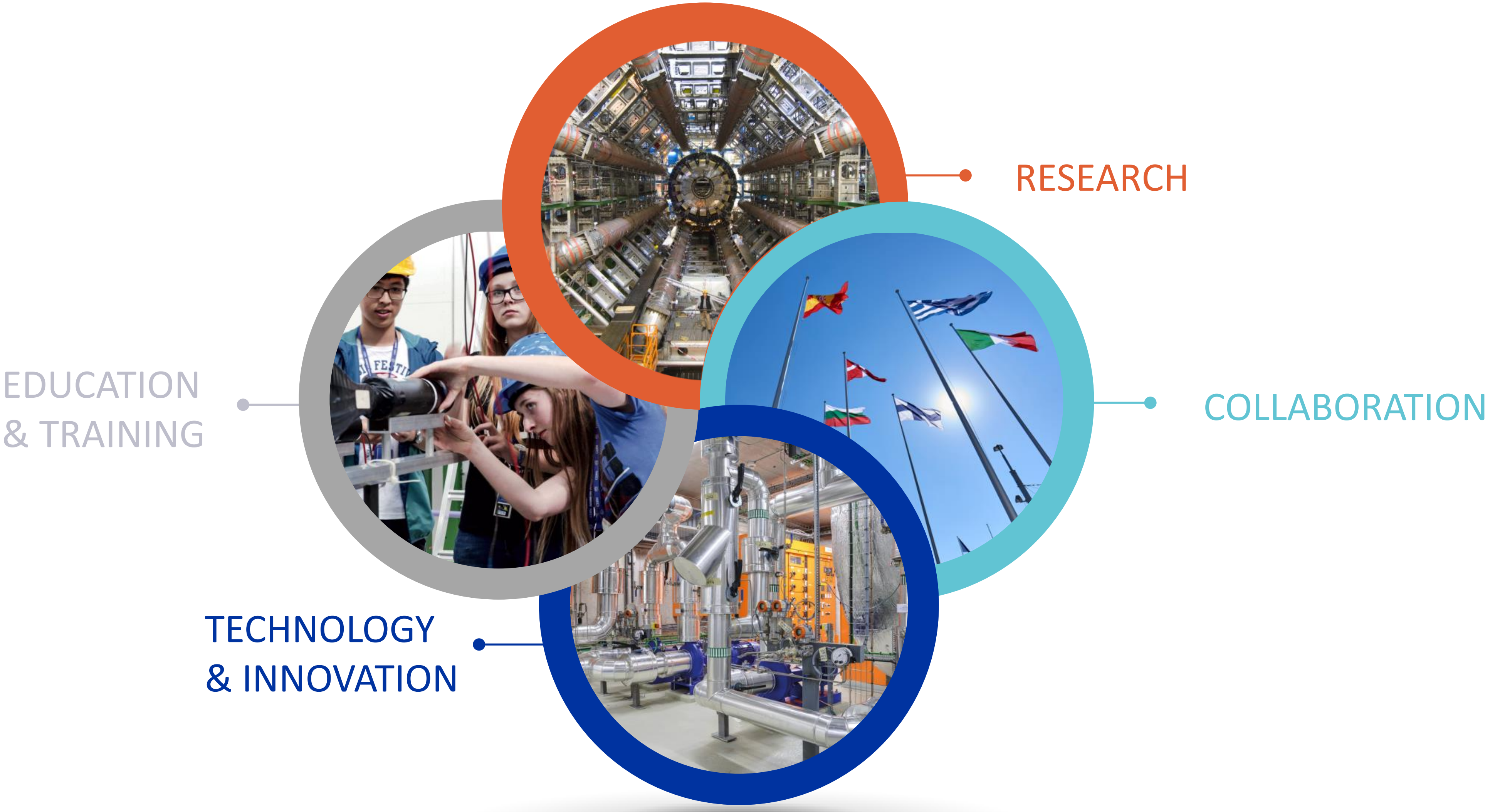
3 Associate Member States
in the pre-stage to membership
Cyprus – Estonia – Slovenia

7 Associate Member States
Croatia – India – Latvia – Lithuania – Pakistan
Türkiye – Ukraine

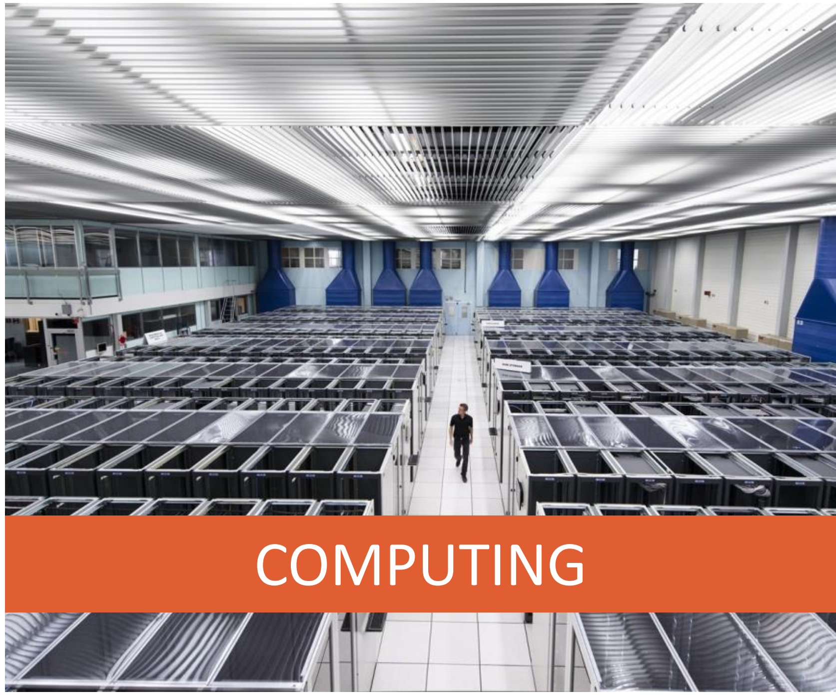
6 Observers
Japan – Russia (suspended) – USA
European Union – JINR (suspended) – UNESCO

Around 50 Cooperation Agreements with non-Member States and Territories
Albania – Algeria – Argentina – Armenia – Australia – Azerbaijan – Bangladesh – Belarus – Bolivia -Bosnia and Herzegovina - Brazil – Canada – Chile – Colombia
Costa Rica – Ecuador – Egypt – Georgia – Honduras - Iceland – Iran – Jordan – Kazakhstan - Lebanon – Malta – Mexico – Mongolia – Montenegro – Morocco
Nepal - New Zealand – North Macedonia – Palestine - People's Republic of China – Peru – Philippines – Qatar – Republic of Korea – Saudi Arabia – Sri Lanka
South Africa – Thailand - Tunisia – United Arab Emirates – Vietnam

Four pillars underpin CERN’s mission



CERN develops technologies in three key areas



CERN’s technological innovations have applications in many fields

CERN is the birthplace of the World Wide Web

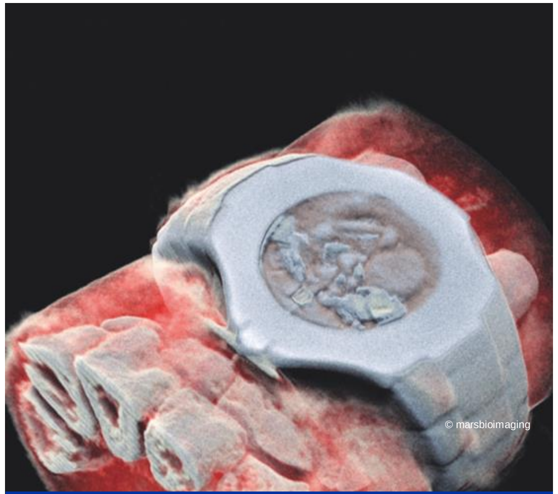
And there are many more examples

Medical imaging, cancer therapy, material science, cultural heritage, aerospace, automotive, environment, health & safety, industrial processes.

CERN's technological innovations have important applications in medicine & healthcare

Accelerator technologies are applied in cancer radiotherapy with protons, ions and electrons

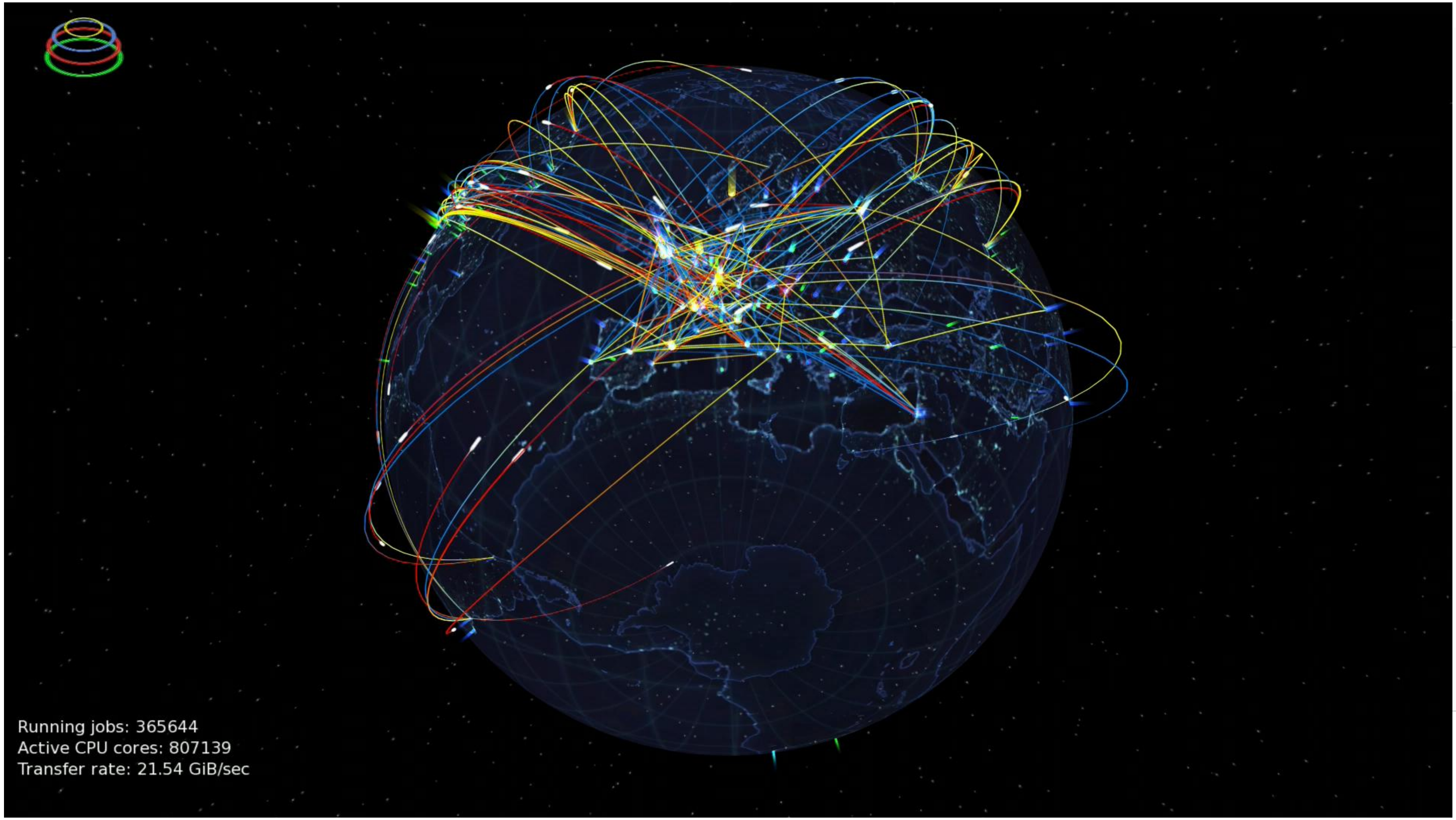
Technologies applied at CERN are also used in PET, for medical imaging and diagnostics



Pixel detector technologies are used for high resolution 3D colour X-ray imaging

CERN produces innovative radioisotopes for nuclear medicine research

WORLDWIDE LHC COMPUTING GRID



A key tool for physics

The most sophisticated data-taking & analysis system ever built for science, providing near real-time access to LHC data.

Seamless access

Computing resources which include data storage capacity, processing power, sensors, visualization tools and more.

Global collaboration

42 countries
170 computing centres
Over 1 million computer cores
2 exabytes of storage

Enabling discovery

WLCG computing enabled physicists to announce the discovery of the Higgs Boson on 4 July 2012.

A novel AI-based tool based on the integration of clinical and patient data over a Federated Learning infrastructure developed, tested, validated and hosted at CERN

“CAFEIN originates in the application of a CERN technology used to identify defects in LHC”

CERN CAFEIN

Table of Contents

CERN CAFEIN — Computer-Aided deFEcts and anomalies detection, Identification and classificatioN system for digital images and data

CERN CASO — Computer-Aided Support to Operation

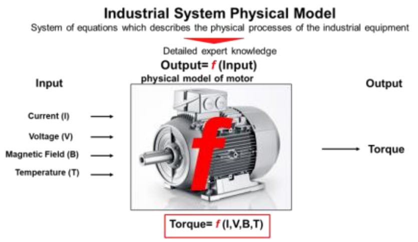
Part 1 – L. SERIO - CERN

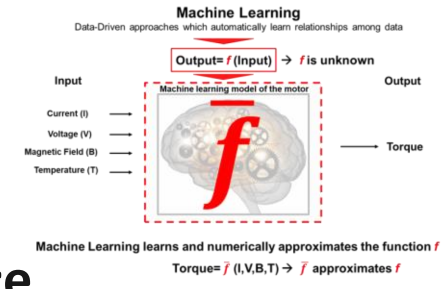
- CERN and its impact on society
- CERN developments in the field of AI
- Knowledge Transfer of AI developments to healthcare
- The TRUSTroke Project and CERN's contribution

Part 2 – P. CALIANDRO – Policlinico Gemelli

- Pathway/Standard of care in the acute phase (stroke code, reperfusion/treatments)
- Standard of care in the chronic phase (outcomes, follow-up, adherence to treatments, recurrence
- Crucial Clinical End Points (CEPs)

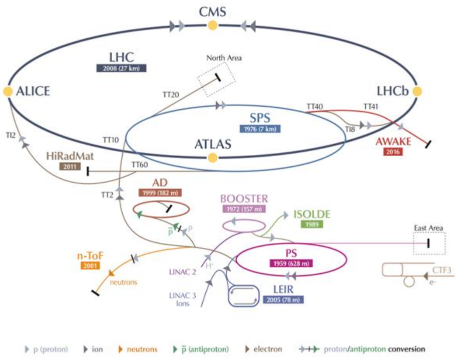
Artificial Intelligence and CERN

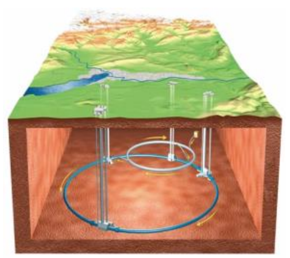




- **Operation support, early identification of failures and prescriptive maintenance**

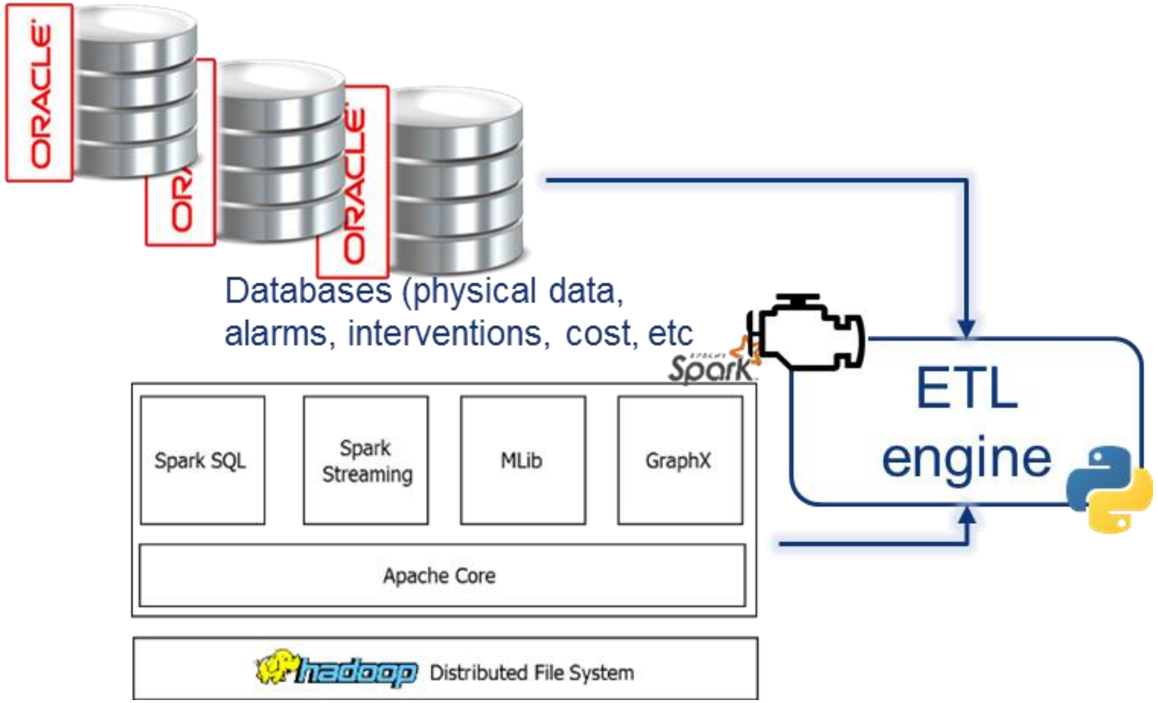
- **Big Data analytics and machine learning techniques to extract descriptive and predictive models**
  - Suitable for **complex systems** and **variable conditions**
  - Efficient when **difficult** to develop **physical modelling**
  - Allows to identify **patterns** in signals, **anomalies** or failures
  - Allows to discover “hidden” **dependencies**
  - Reveals new information from available data (a.k.a. **data-mining**)

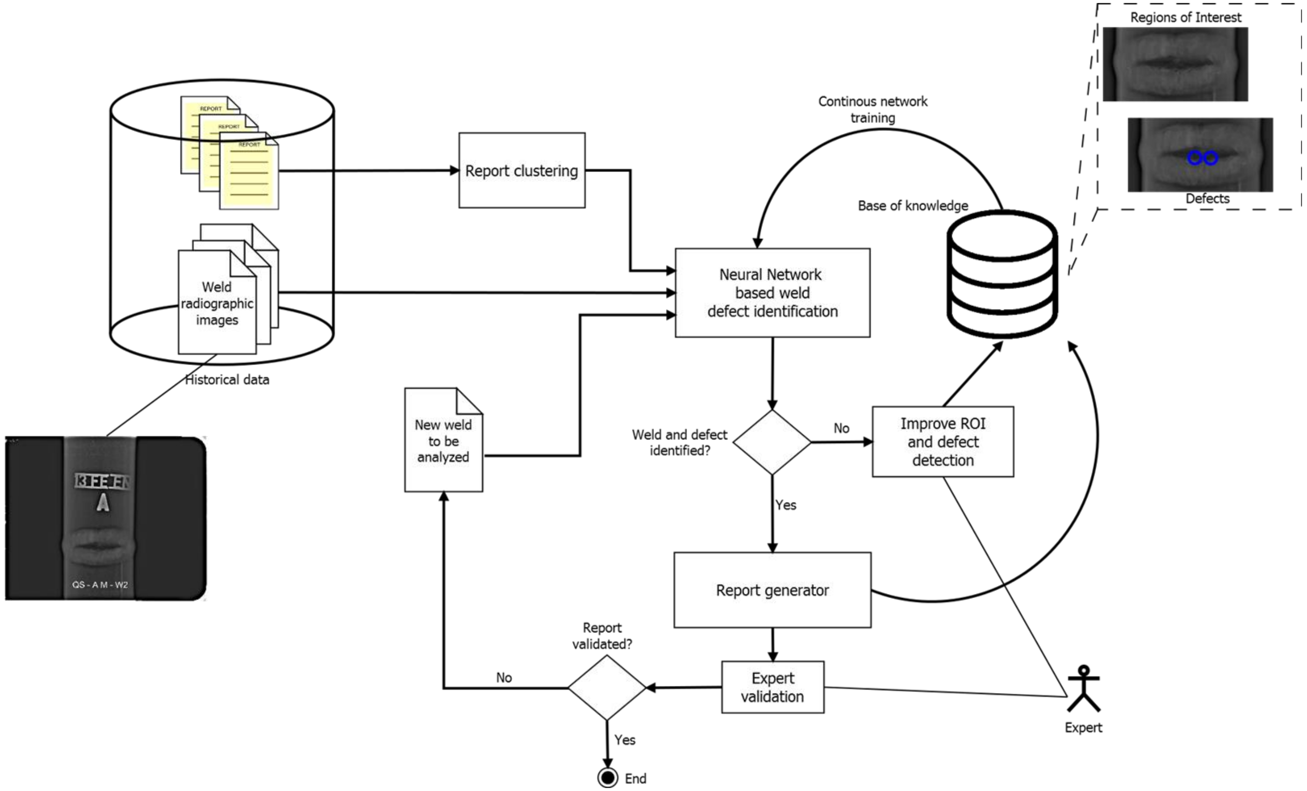




- **Complex system of systems with very stringent requirements in terms of availability and reliability**
  - Requiring **tools** for: quality control, faults analysis, prevention, prediction and mitigation
  - Providing a complete **test bed** for: complex fault trees, systems dependencies, risks and failures propagation, data and images analysis and interpretation

Brief history of the initial work 1/2



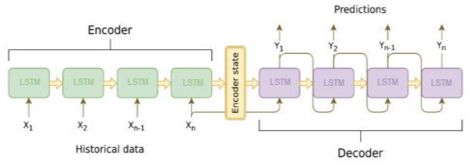


Framework to automatically Extract, Transform, Load heterogeneous data for analysis

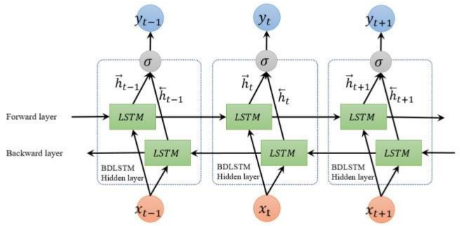
Framework to automatically Extract, Transform, Load, Cluster heterogeneous data and images for analysis

CERN CASO

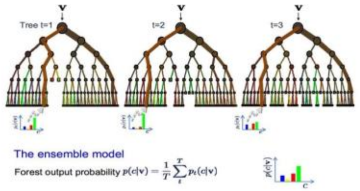
Brief history of the initial work 2/2



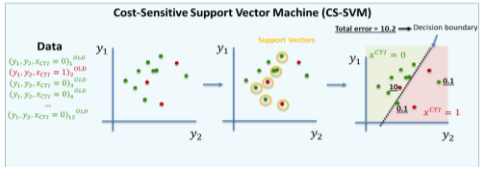
Collimators – **detection of the temperature sensor failure** – LSTM



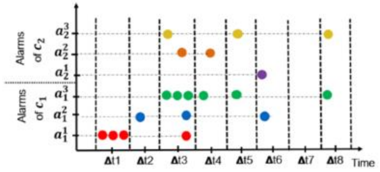
UPSs – **detection of battery ageing** – bi-directional LSTM

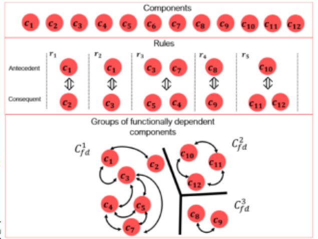


Transformers – **fault detection/RUL**– autoencoders and random forest



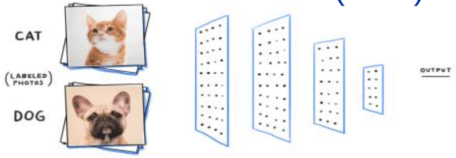
Electrical devices vs. **Beam Dumps** – RELIEF and Wrapper (Genetic + CS-SVM)

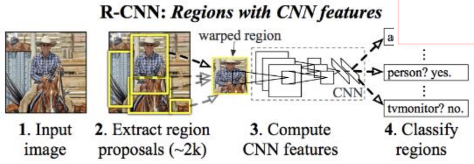




Mining Dependencies of Systems and Components from Alarms Cascade – APRIORI

Conv. Neural Net. (CNN)





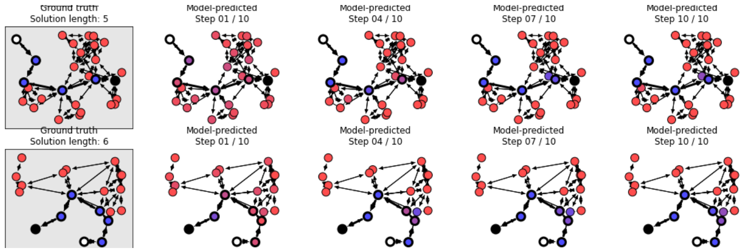
Automatic **detection and classification** of Welds – R-CNN

Complex systems modelling: Graph Neural Networks

Neural Network that operates on **graph data (GNN)**

A graph network takes a graph as input and returns a graph as output.

The output graph has the same structure, but updated attributes.

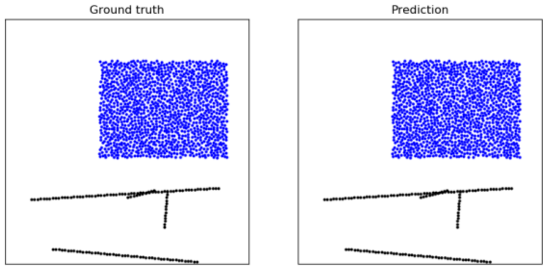


Proved to be able to learn interactions (and physics) **just by "observing" the real world**

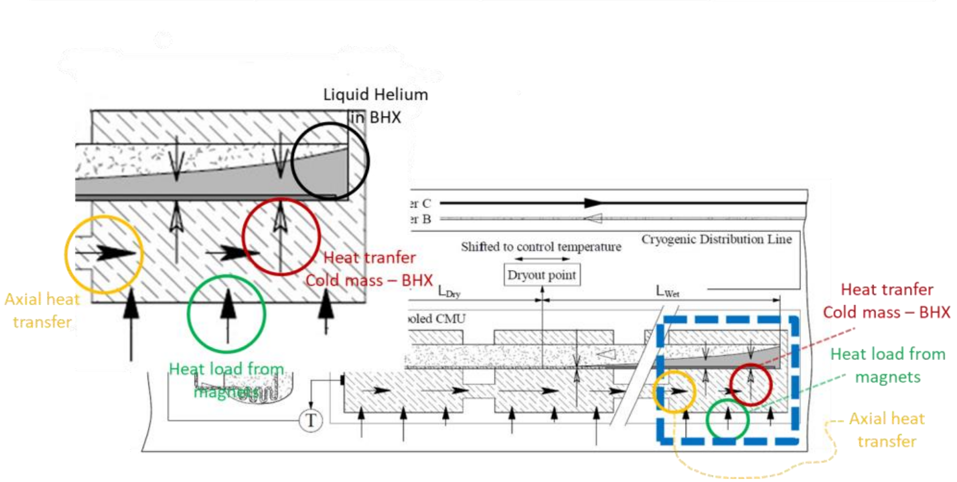
Many real-world objects and phenomena can be represented as graph problem. *e.g.: Simulations* *[Sanchez-Gonzalez at al. 2020]*

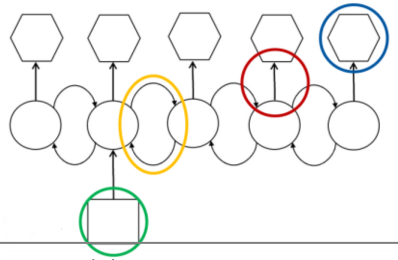
Useful in complex simulations, can speed-up and optimise computation

They can learn relations ("physics") just based on observations => They have the potential to find new *(yet unknown)* relations



Modelling of the LHC cryogenics Superfluid Helium Flow in Bayonet Heat Exchanger Tubes using Graph Neural Networks

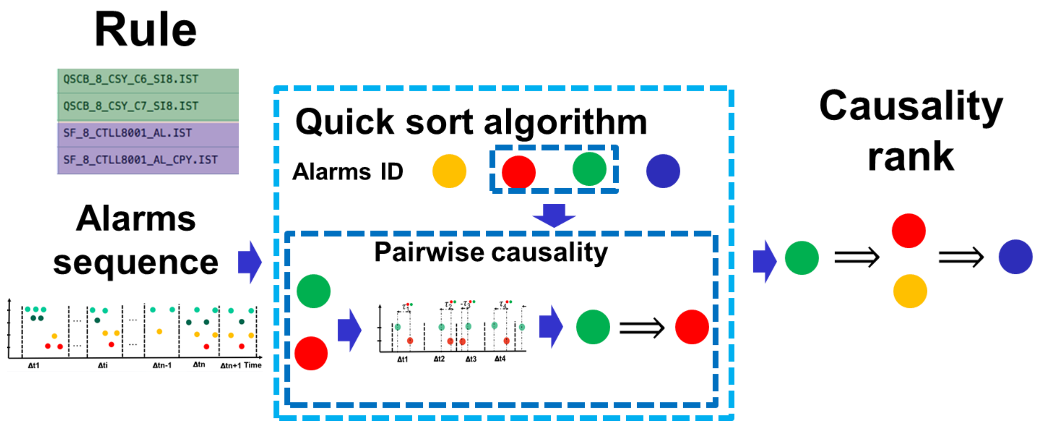




Ref. Investigation and perspectives of using Graph Neural Networks to model complex systems: the simulation of the helium II bayonet heat exchanger in the LHC – R. Stoklasa, N. Calabrese, L. Serio – CEC2023 – Hawaii USA

AI assisted diagnosis and prescriptive maintenance for critical infrastructures operation

Data-driven tools capable of discovering dependencies and abnormal behaviours



Capable of inferring and interpreting data from different and heterogeneous sources and systems

Augmented Evolutionary and APRIORI Algorithm for the Identification of Functional Dependencies

Ref. A Data-Mining Framework for Functional Analysis of Complex Technical Infrastructures – F. Antonello – Politecnico di Milano – PhD - 2017/2020

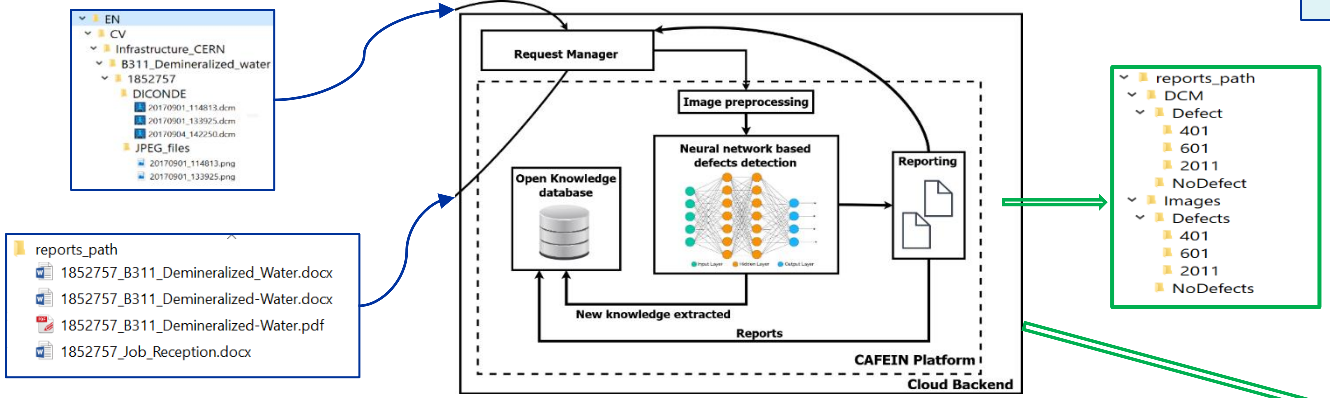
Ref.: A Niching Augmented Evolutionary Algorithm for the Identification of Functional Dependencies in Complex Technical Infrastructures from Alarm Data, F. Antonello, P.Baraldi, L. Serio, E. Zio, IEEE SYSTEMS JOURNAL, 10.1109/JSYST.2022.3146014

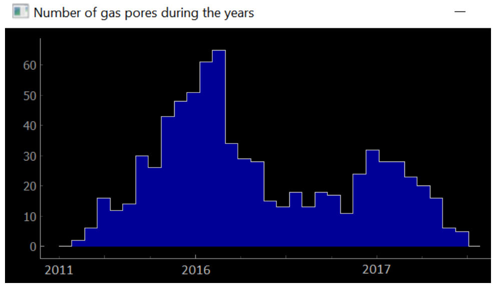
AI assisted X-ray image analysis for quality control of LHC welds

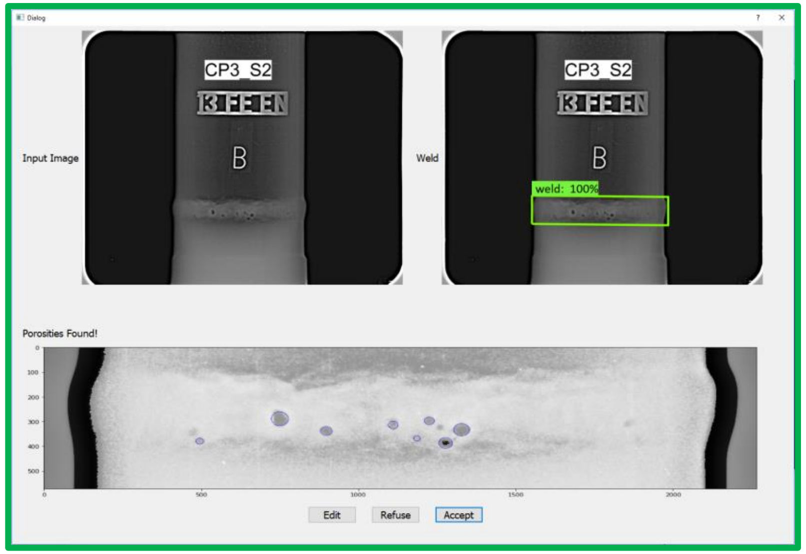
CERN CAFEIN

Deep Learning – Convolutional Neural-Networks

Image Classification, Object Detection, Semantic Segmentation, etc…







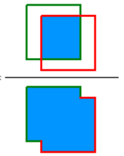
The developed CNN-based detector (by category):

- AP* with IoU* of 50 % for **weld** detection = 97 %
- AP** with IoU* of 50 % for **pore** detection = 79 %

The evaluation set considered is composed of 450 images

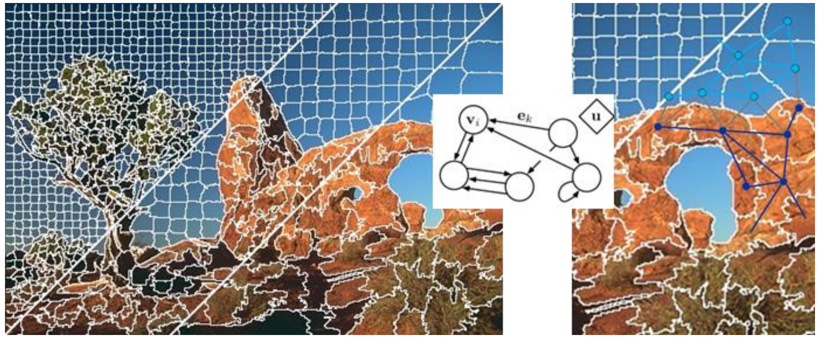
*Average Precision
**Intersection over Union

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} =$$



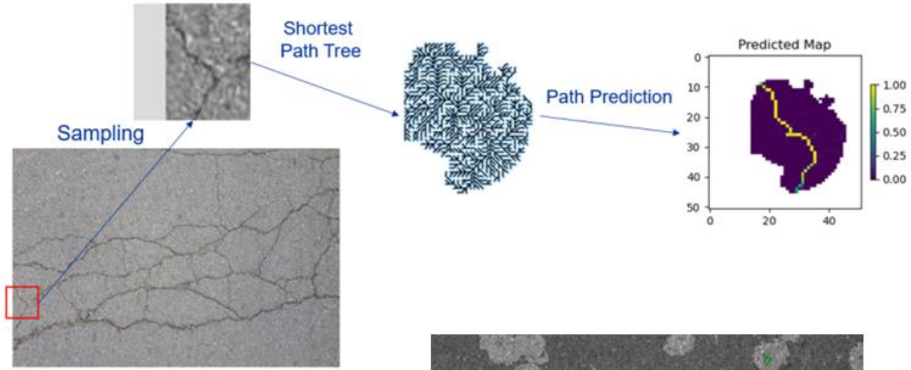
Ref. A deep-Learning based Method for the Detection of Defects from Images of Industrial Equipment A. Perin - Politecnico di Milano – MSc thesis 2020

Image segmentation using superpixels and GNNs



GNNs are versatile and easily scalable

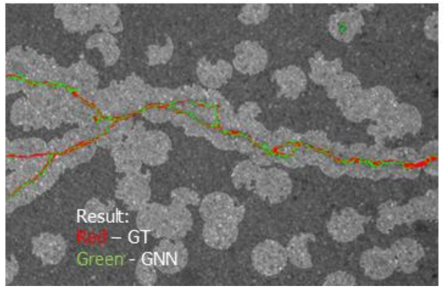
Have been successfully tested and implemented at CERN to segment cracks in images



Precision : 0.937

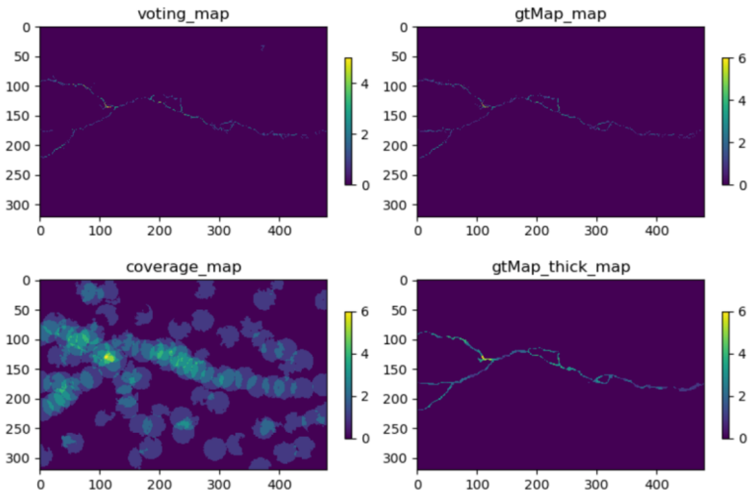
Recall : 0.846

F1-score : 0.889



Reconstructed Image 115 from samples, for threshold 0.61

Reconstructed predictions



Location of samples used for reconstruction

Reconstructed original image segmentation

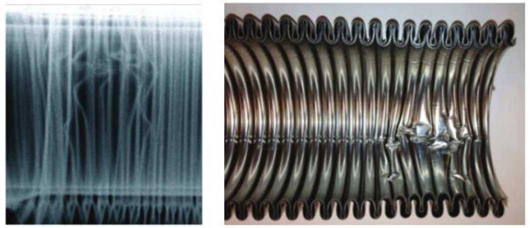


Fig. 2. (a) X-ray of first leaky bellow; (b) photo of the first leaky extracted bellow.

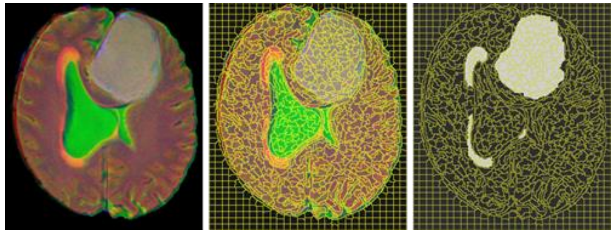


Table of Contents

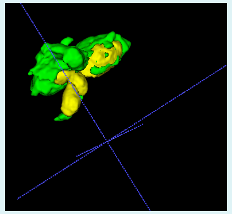
CERN Budget for Knowledge Transfer to Medical applications

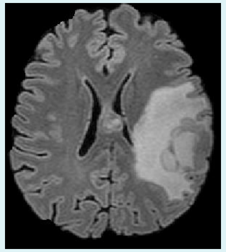
CERN CAFEIN

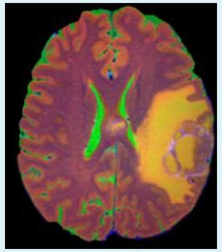
Extension to the medical field of AI tools for

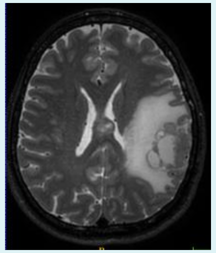
- functional and dependency analysis of complex critical infrastructure

- digital imaging for radiography autonomous defects detection









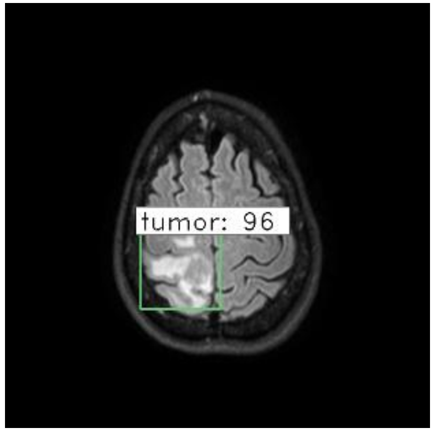
- **Field of application:**
  - Semi-automated **analysis** and **modelling** of **medical data** and **images**
  - **Diagnosis** and treatments based on multiple features and data **beyond human perception**
  - **Federated learning and distributed computing** to **ensure privacy** for a wide and safe international **collaboration** as well as access to **diagnostic models in remote areas**
- **Competing technologies:**
  - No other technologies **ready to use** in the field, **tailored to clinical needs** and **privacy preserving**
- **Medical application:**
  - **Brain pathologies** detection, analysis and segmentation based on CNN applied to MRI images

Medical application proof of concept: detection and segmentation

Training on initial 230 multimodal MRI (3 T) multimodal images

- Detection accuracy: 84 %
- Classification accuracy: above 93 % (image based)
- Classification accuracy: 85 % (radiomics based)

model's Output
detection & classification



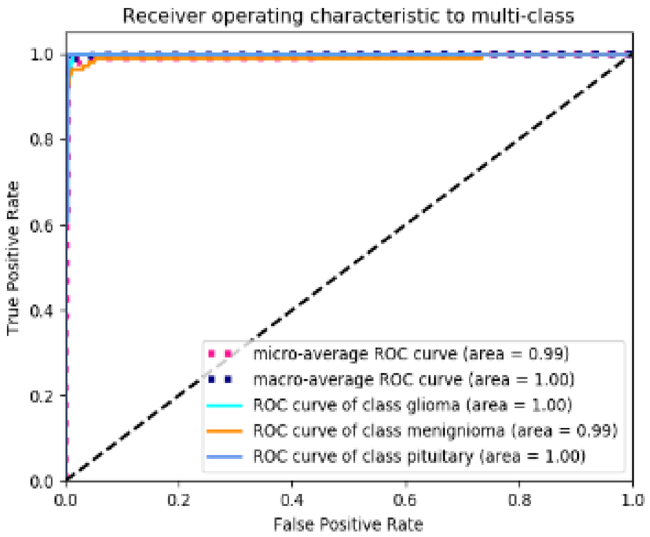
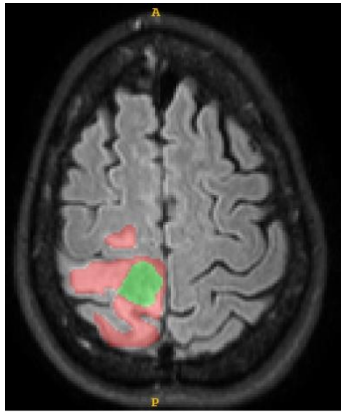
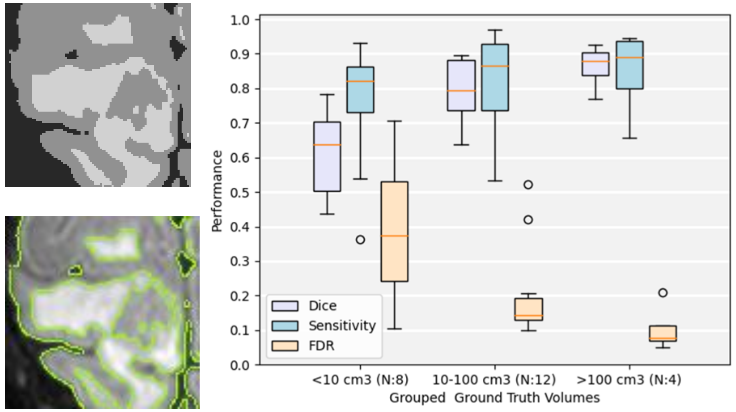


image labeled by the radiologist

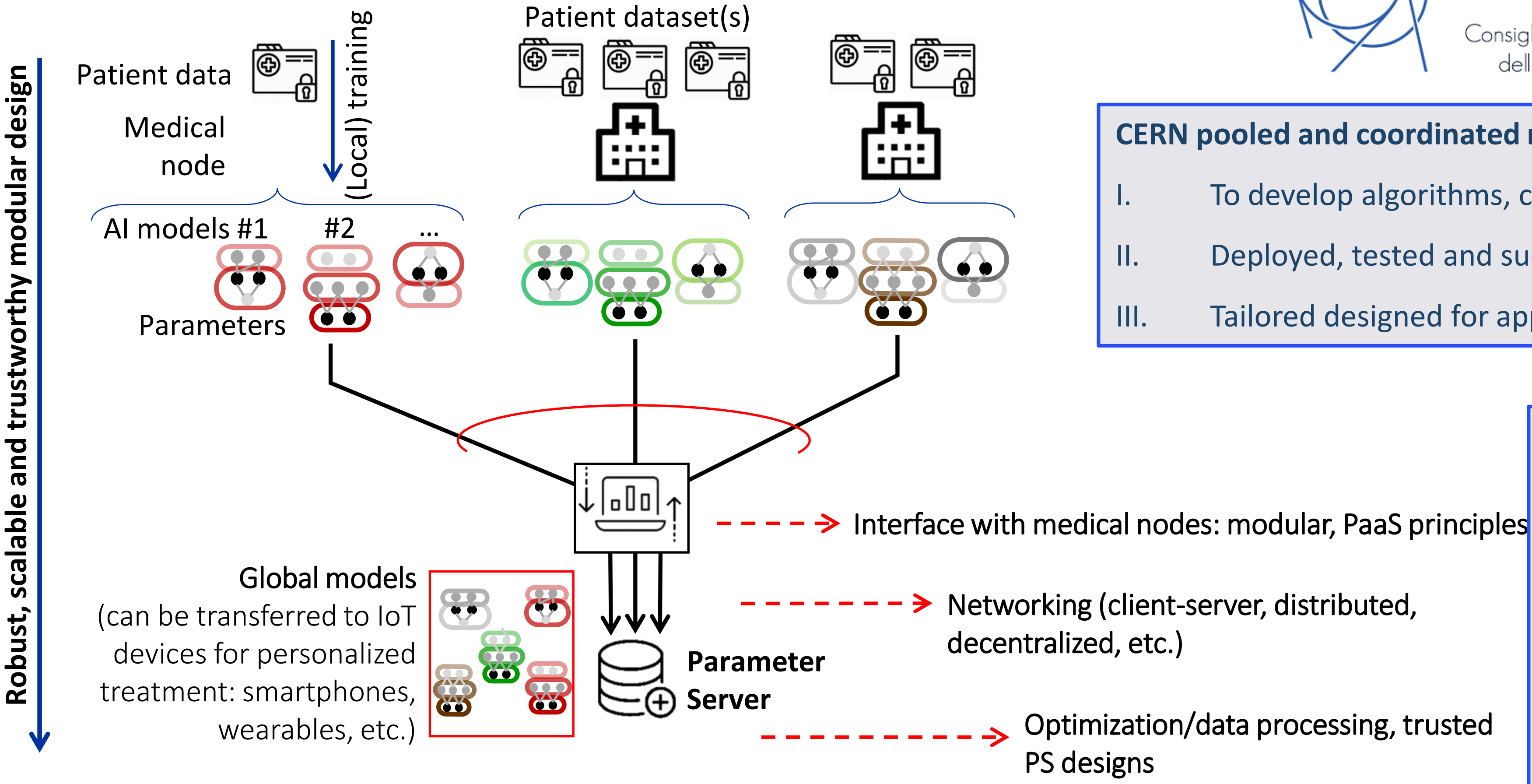




Ref.: Development and use of deep learning algorithms for brain tumor diagnosis and classification based on MR images - I. Stathopoulos – PhD – 2020/2022

Federated Learning and Distributed Computing

required step to extend to clinical datasets (privacy and anonymization)



CERN pooled and coordinated resources and know-how from academia

I. To develop algorithms, communications and network design

II. Deployed, tested and successfully validated a POC

III. Tailored designed for applications for healthcare institutions

The federated platform is composed of:

I. a modular interface to instantiate, monitor and control FL processes

II. a network infrastructure that supports advanced security methods

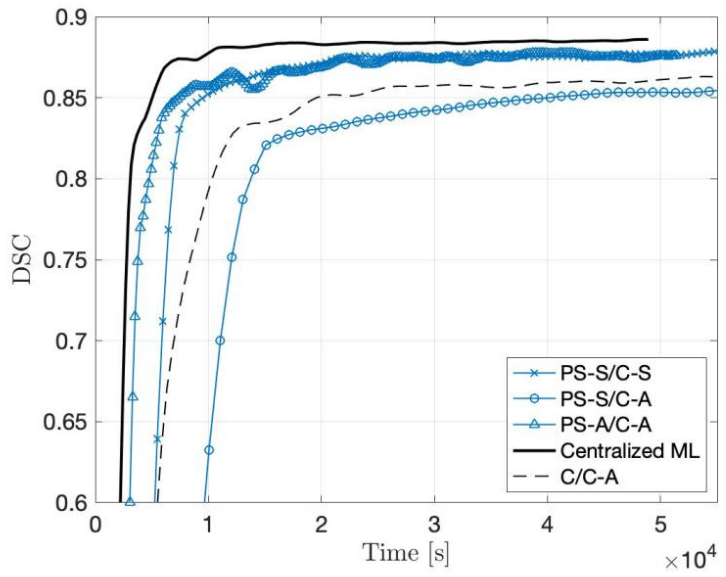
III. a parameter server designed against security attacks (data poisoning and model inversion)

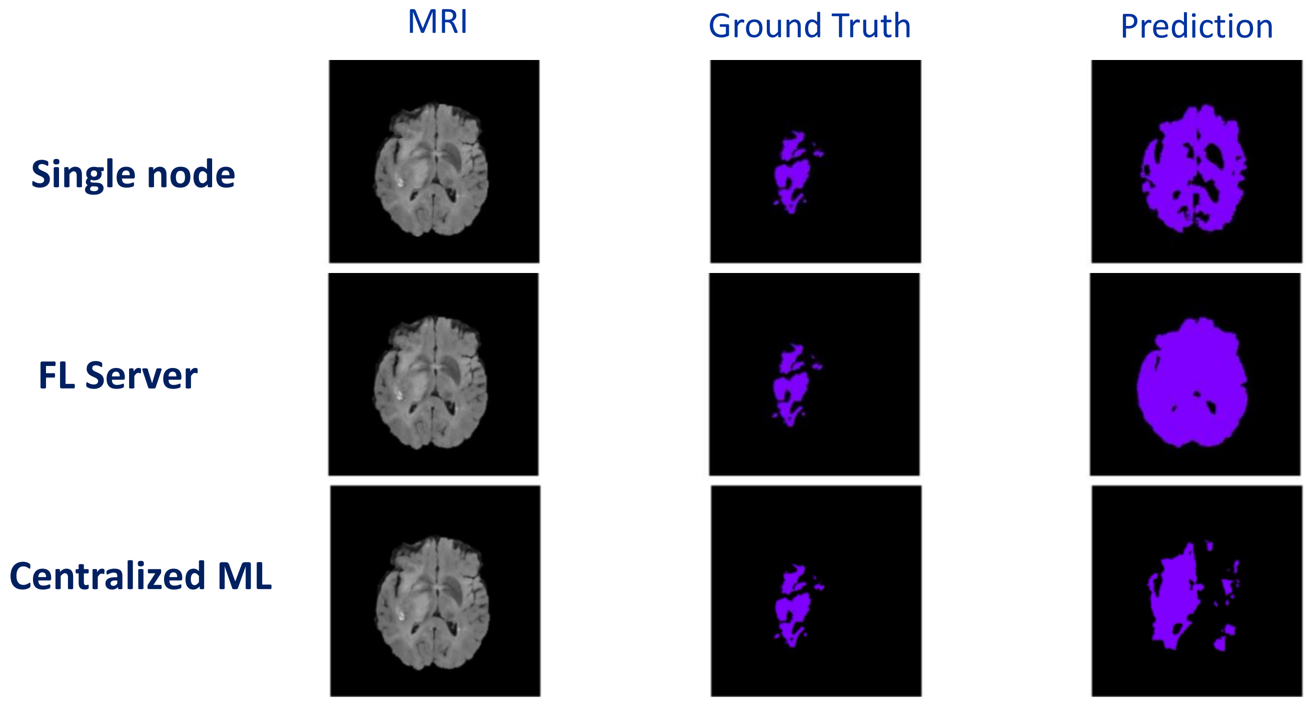
Ref.: Federated learning architectures and algorithms for diagnostic imaging in healthcare networks – B. Tedeschini – MSc thesis – 2021

Platform POC test and validation

Federated approach impact:

- **Robustness** of global models v. local models
- **Privacy** and **confidentiality** of data
- Communication optimization and sustainability

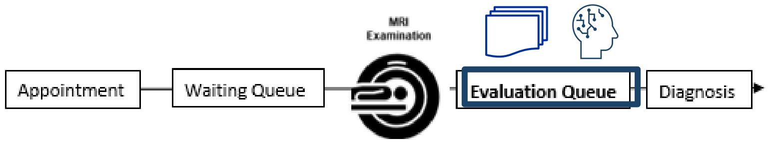




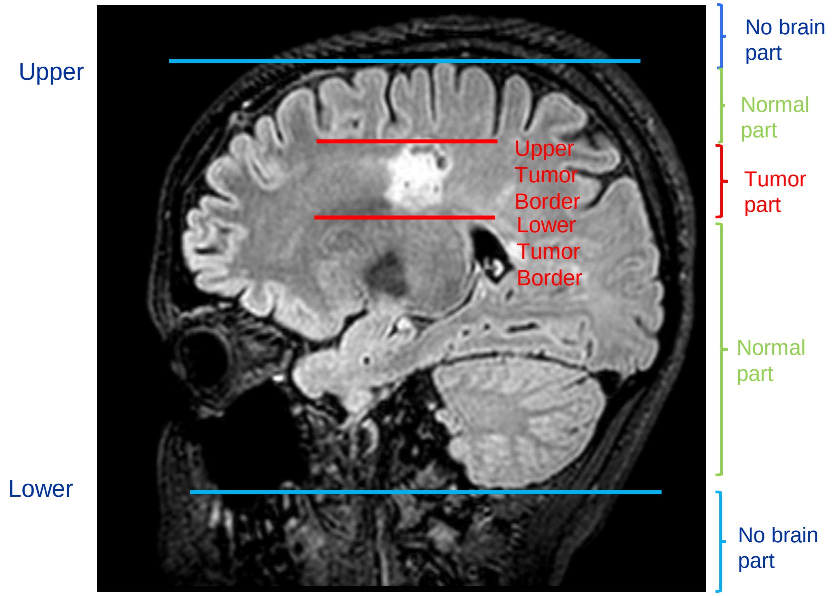
Ref.: Decentralized Federated Learning for Healthcare Networks: A Case Study on Tumor Segmentation, B. Camajori Tedeschini, S. Savazzi, R. Stoklasa, L. Barbieri, I. Stathopoulos, M. Nicoli, L. Serio, January 2022, in IEEE Access, 10.1109/ACCESS.2017.DOI

CERN CAFEIN

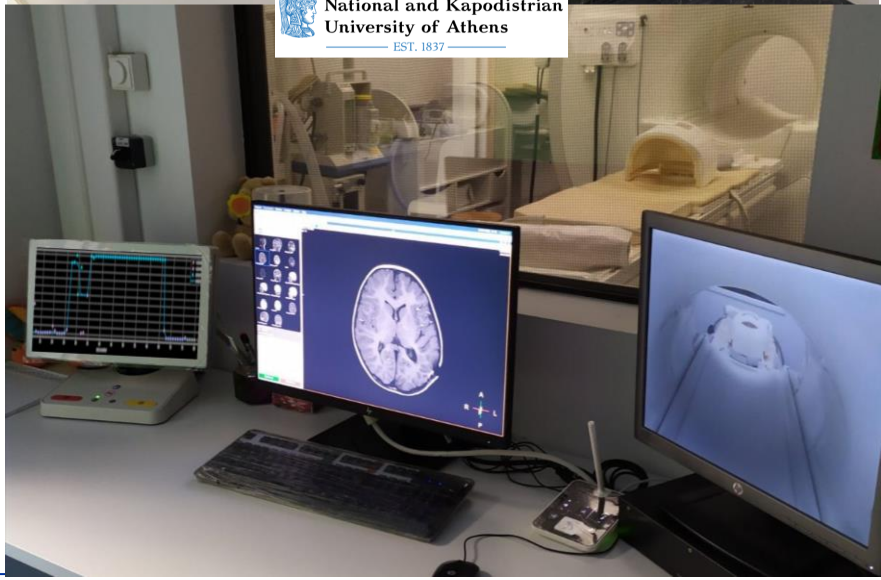
Deployment of the first clinical application:
brain pathologies screening tool



- Typical examination comes as consecutive 2D slices
- Only **a small amount** contains **useful information**
- Tool for **automated** and **optimized screening** to save time and use efficiently clinical resources
- 3D multi-modal U-Net

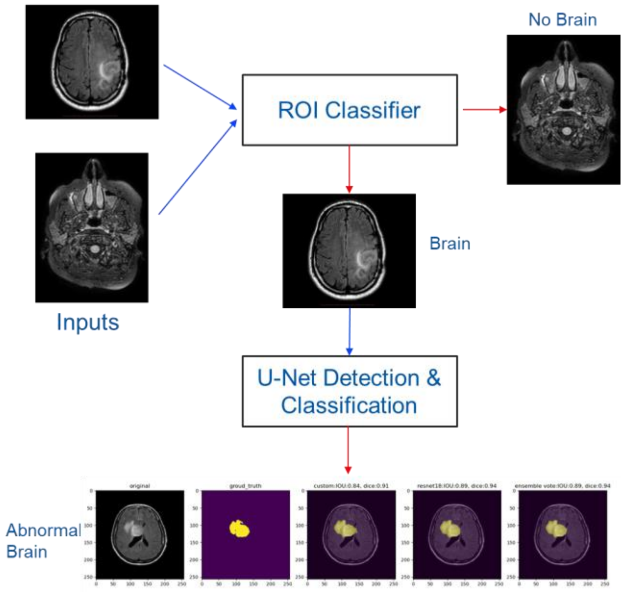


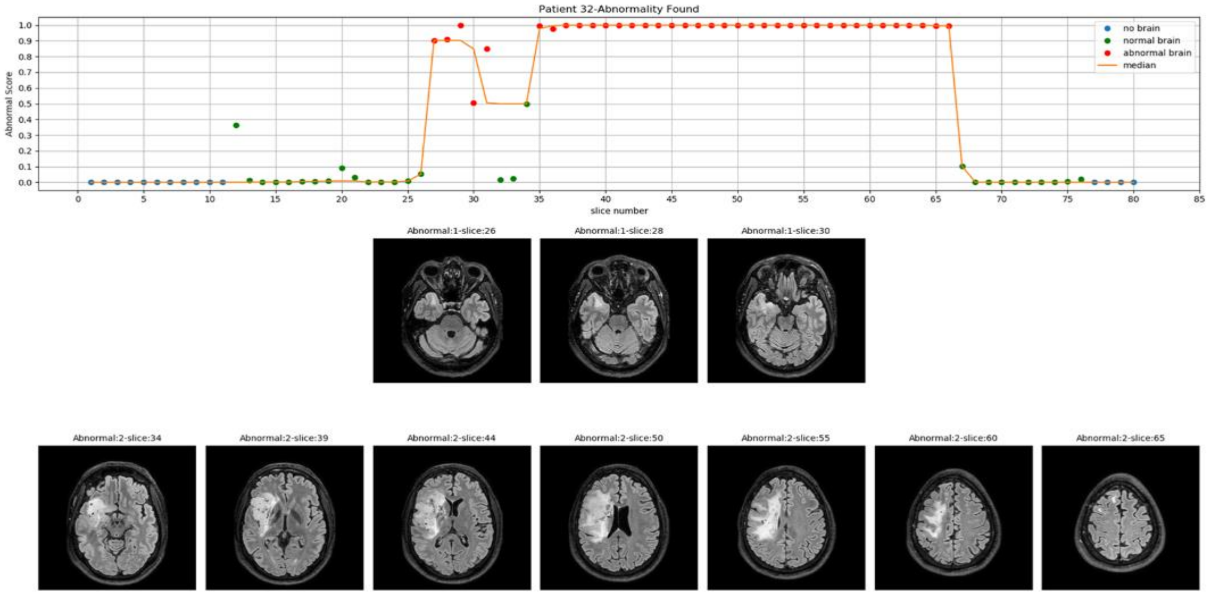
HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837



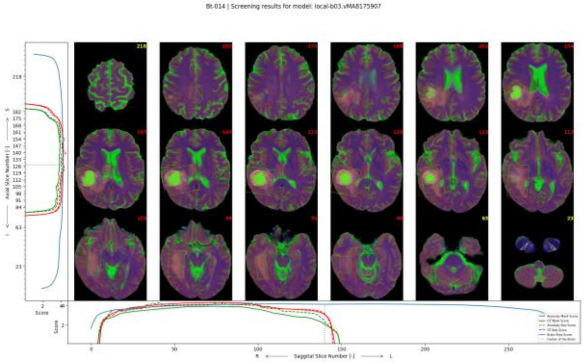
Ref.: Brain Lesions Screening Tool Based on Deep Learning - Stathopoulos I, Stoklasa R, Tsochatzis A, Velonakis G, Karavasilis E, Efstathopoulos E, Serio L 1st Panhellenic Conference of Medical Physics - Athens – September 2022 https://doi.org/10.1016/S1120-1797(22)03066-6

Results of the screening tool for MRI images
pathologies detection and selection of relevant images

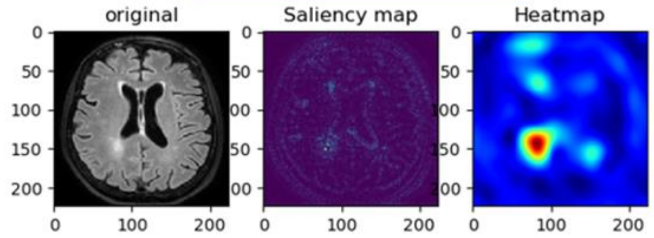


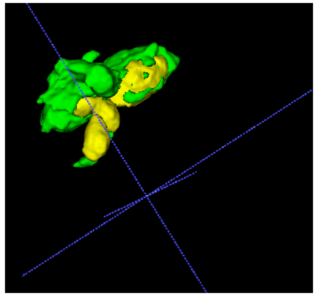


3D U-Net Detection & Segmentation

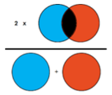


Interpretability





- **Per-slice accuracy above 97 %**
- **Average DSC: 0.89**
- **Parametric detection thresholding**



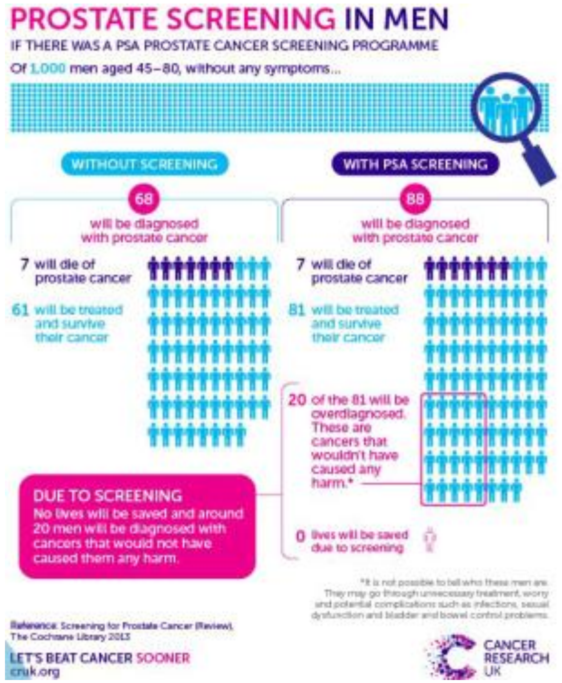
Ref.: Development and use of deep learning algorithms for brain tumor diagnosis and classification based on MR images - Ioannis Stathopoulos – PhD thesis - Athens University – 2023

Platform extension - screening

Risk based approach for the screening of breast and prostate cancer

“Does screening work” shall move to “for whom does screening work?” – ref IARC / WHO

- Avoid overdiagnosis and overtreatment
- Optimize clinical support
- Profit from available large datasets and national studies / follow-up



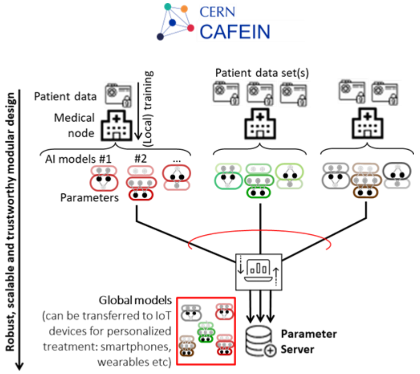
Breast and Prostate cancer

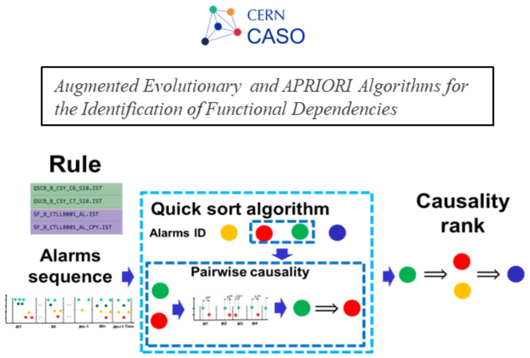
» Remaining top 5 cancers w/o risk-based approach

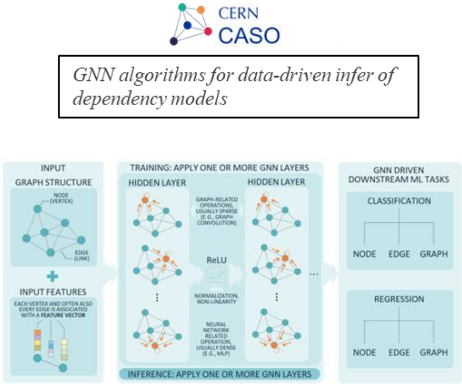
» Large nationally available datasets presently **not shared internationally**

» **EPIC study**: one of the largest cohort with more than **0.5 M participants**

- Diet, nutrition, lifestyle & environmental factors, the incidence of cancer







Perspectives and future work

CERN CAFEIN

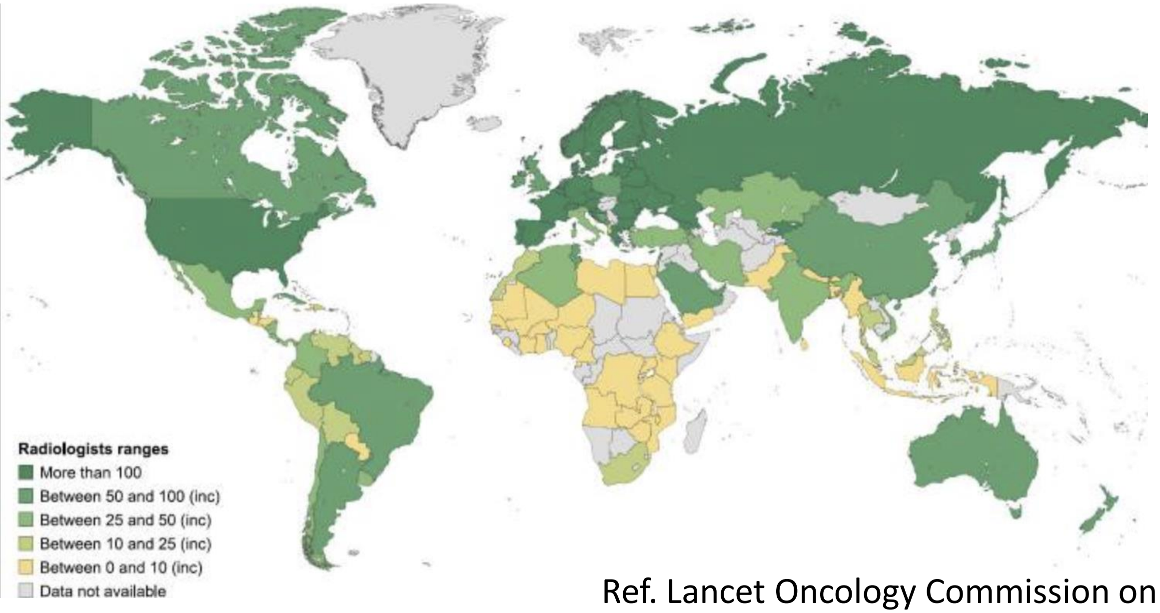
share knowledge without exchanging data

AI to support healthcare in areas were access to healthcare is limited due to lack of specialists in the field of technology, **diagnosis**, therapies and **monitoring**

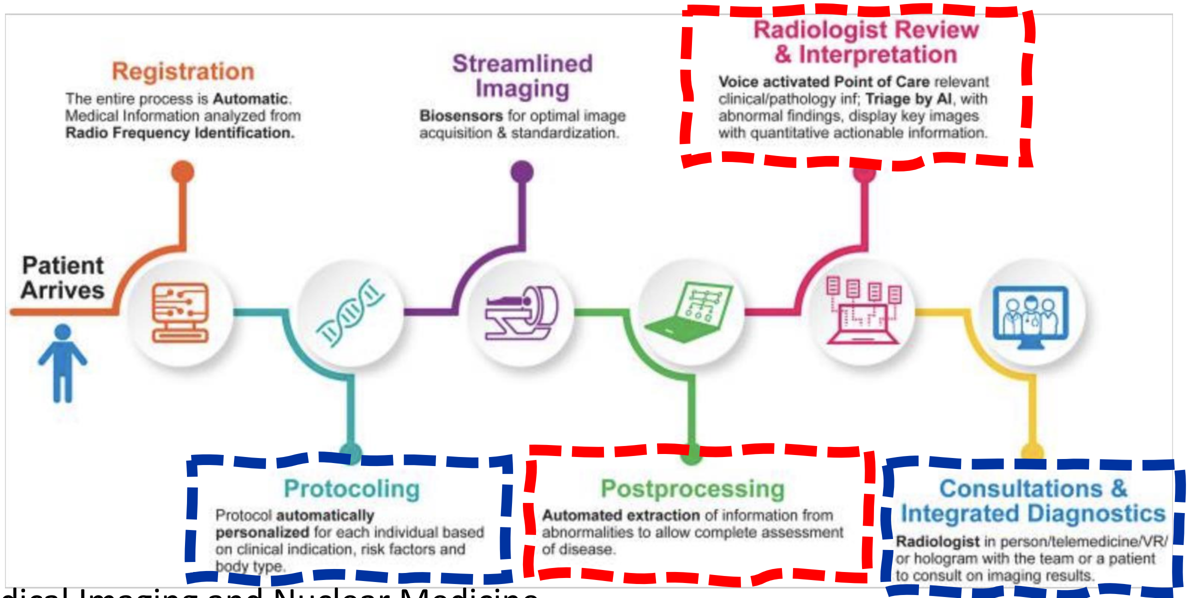
CERN CASO

CERN CAFEIN

Estimated number of radiologists per million inhabitants



Artificial intelligence-driven workflow for imaging in patients with cancer



Ref. Lancet Oncology Commission on Medical Imaging and Nuclear Medicine

AI algorithms with Federated Learning to provide efficiently **quality healthcare everywhere** and at the same time enhance the robustness of the models with **huge amounts of untapped data**

Perspectives and future work

CERN CAFEIN

share knowledge without exchanging data

AI to support healthcare in area were access to healthcare is limited due to lack of specialists in the field of **technology**, diagnosis, **therapies** and monitoring

Much of the world has limited or no access to cancer treatment – especially radiation therapy (RT)

For nearly 60% of cancers, RT is most useful tool for cancer cure or palliation; **inadequate supply of RT linear accelerators** (LINACs).

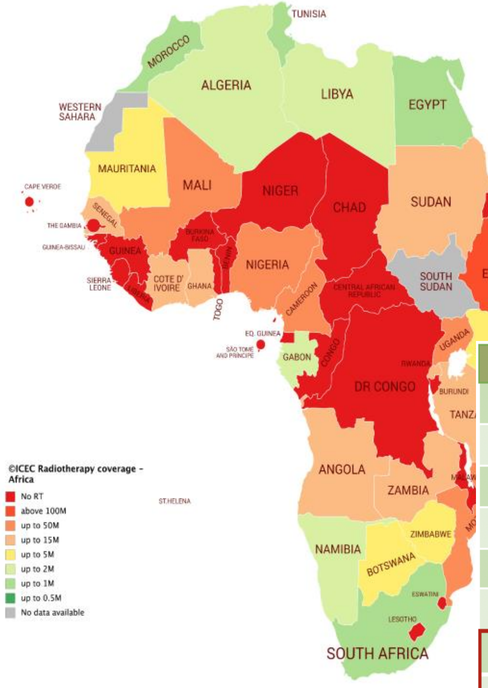
- Gap greatest in low-middle income countries (LMICs)

27.5 million new cancer diagnoses and 16.3 million projected cancer-related deaths worldwide in 2040. (WHO)

- **70% of these deaths will occur in LMICs**

Current **LINAC technology is complex, labor intensive, and high cost to acquire, install, operate and service.**[1]

[1] Jaffray, D. A., Knaul, F., Atun, R., Adams, C., Barton, M. B., Baumann, M., ... Gospodarowicz, M. (2015). Global Task Force on Radiotherapy for Cancer Control. *The Lancet Oncology*, *16*(10), 1144-1146. [237]. https://doi.org/10.1016/S1470-2045(15)00285-5



CERN CAFEIN

CERN CASO

Country	LINACs	Population	People per LINAC
Ethiopia	1	115 M	115,000,000
Nigeria	7	206 M	29,400,000
Tanzania	5	59.7 M	11,900,000
Kenya	11	53.8 M	4,890,000
Morocco	42	36.9	880,000
South Africa	97	59M	608,000
UK	72	67.M	195,000
Switzerland	540	8.6 M	119,000
US	3827	331 M	87,000

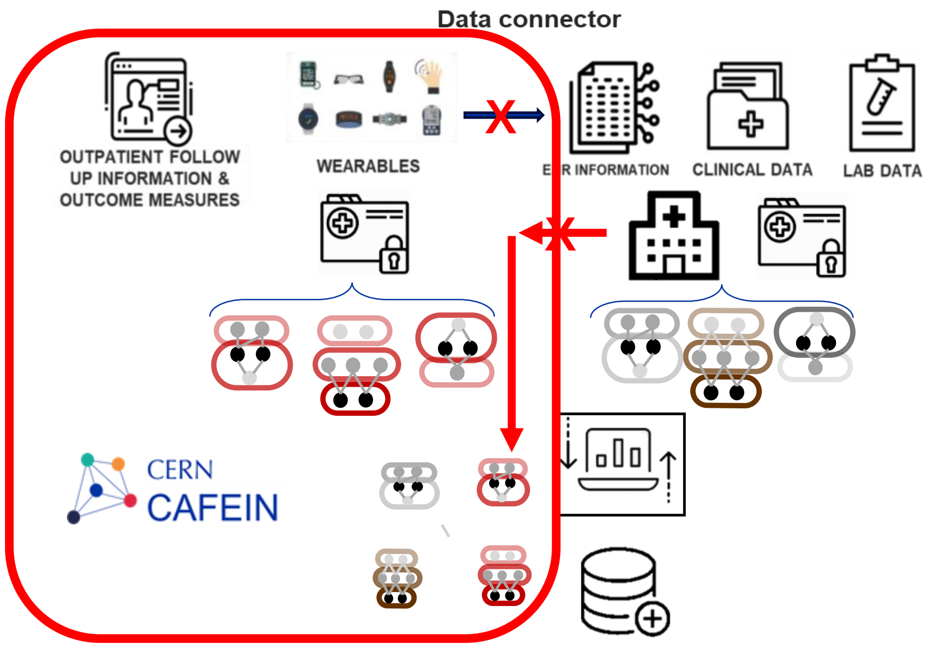
AI algorithms with Federated Learning to provide efficiently **support to design, operation and maintenance** of high technology devices in countries with limited access to specialists and ressources

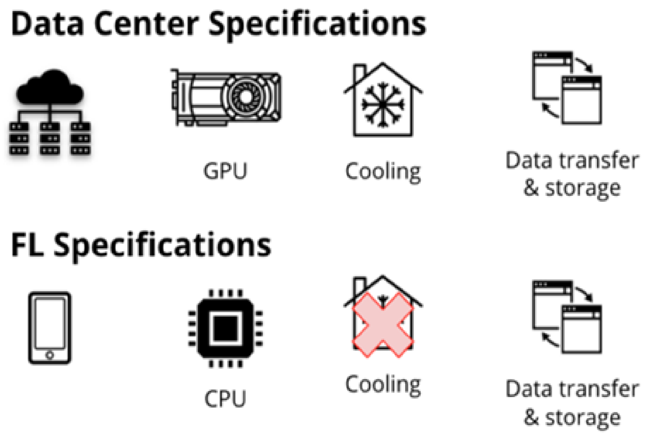
Perspectives and future work

share knowledge without exchanging data

Algorithms and federated learning in **edge devices** wearable by the patient

Empowering the patient with his data, their control and privacy, the knowledge of the global models





Federated learning to **support sustainability** optimising the learning process while **minimising data exchange, storage and communication overheads**.

Table of Contents

the TRUSTroke project

novel AI-based tool to assist in the management of stroke

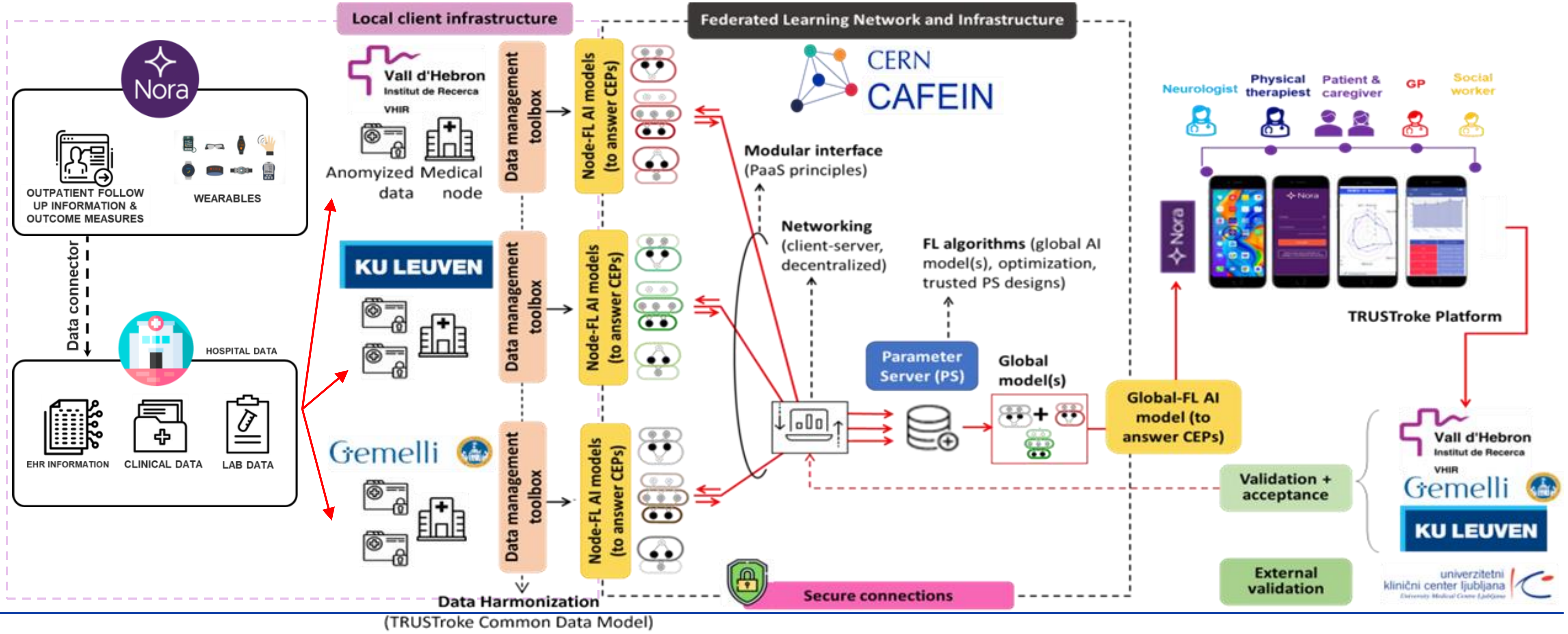
AI-tool based on the integration of clinical and patient reported data

Trustworthy assessment of **disease progression** and **risk of recurrence**

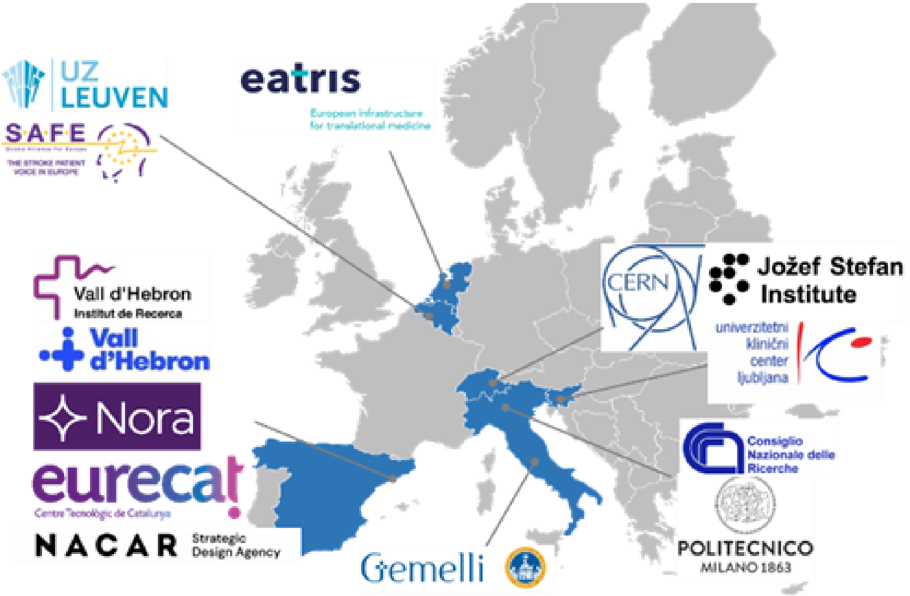
Almost **10’000 enrolled patients’** data will train algorithms over **CERN federated learning platform**

Stroke is the leading cause of severe disability worldwide

- 1.1 m strokes/y in EU
- 0.5 m deaths/y in EU
- 9.5 m stroke survivors



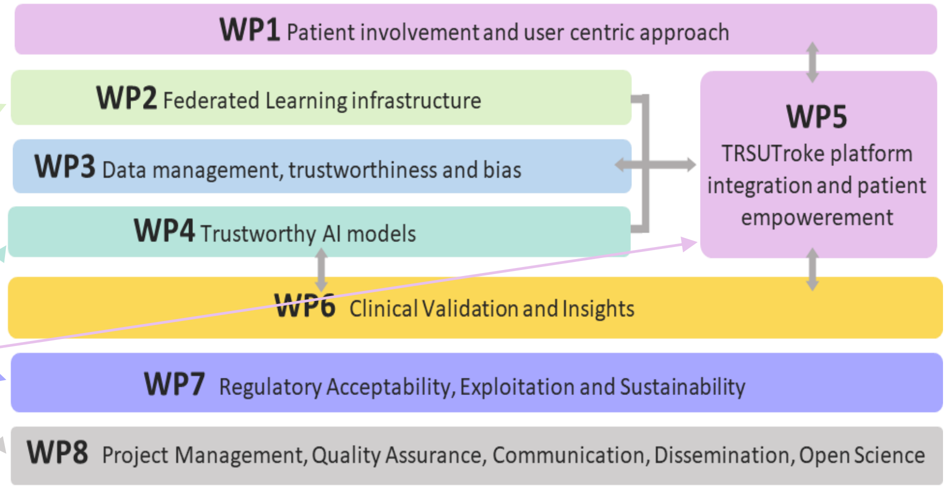
TRUSTroke Project Organization



	Coordination	Trustworthy AI	AI models	Software Integration	Data Interoperability	Patient Engagement & Communication	Sustainability	Clinical Validation	FL development	Network Security	UX experience
Vall D'Hebron Hospital	●	●				●		●			
Fondazione Policlinico Gemelli		●				●		●			
KU Leuven						●		●			
UKC Ljubljana						●		●			
EATRIS							●				
CERN		●	●						●	●	
Eurecat		●	●	●							
Josef Stefan Institute		●	●		●						
Nora Health				●	●	●	●				●
Stroke Alliance for Europe						●			●		
Politecnico di Milano		●							●	●	
Consiglio Nazionale delle Ricerche		●	●						●		
Nacar											●

***WP2**, **led by CERN**, is devoted to the design and development of the FL infrastructure, the implementation and validation of the federated system composed of different hospitals across Europe.*

***WP4, 5, 7 and 8**, CERN participation*



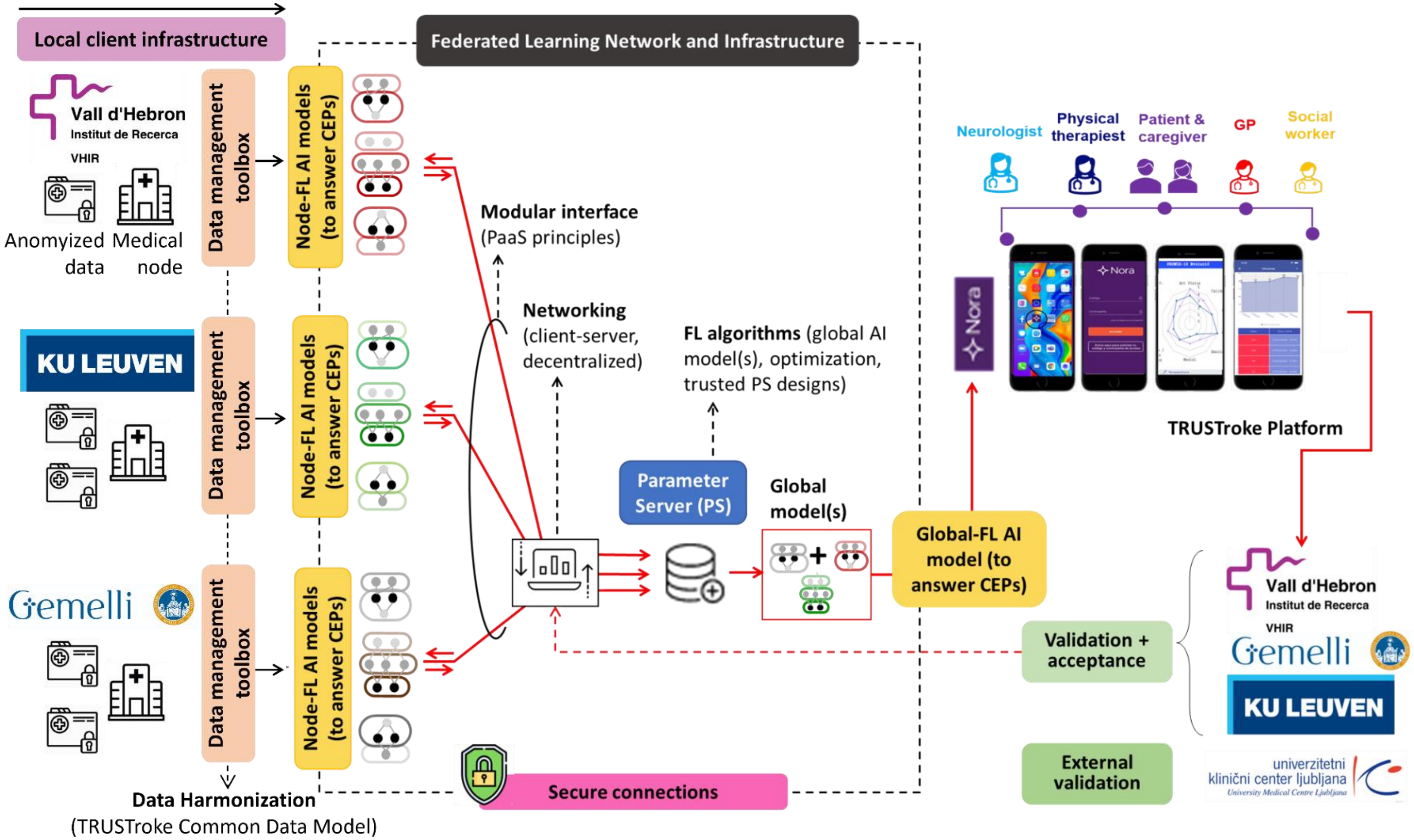
TRUSTroke WP2 Federated Infrastructure (hosted by CERN)

trustroke
CERN
Consiglio Nazionale delle Ricerche
POLITECNICO MILANO 1863

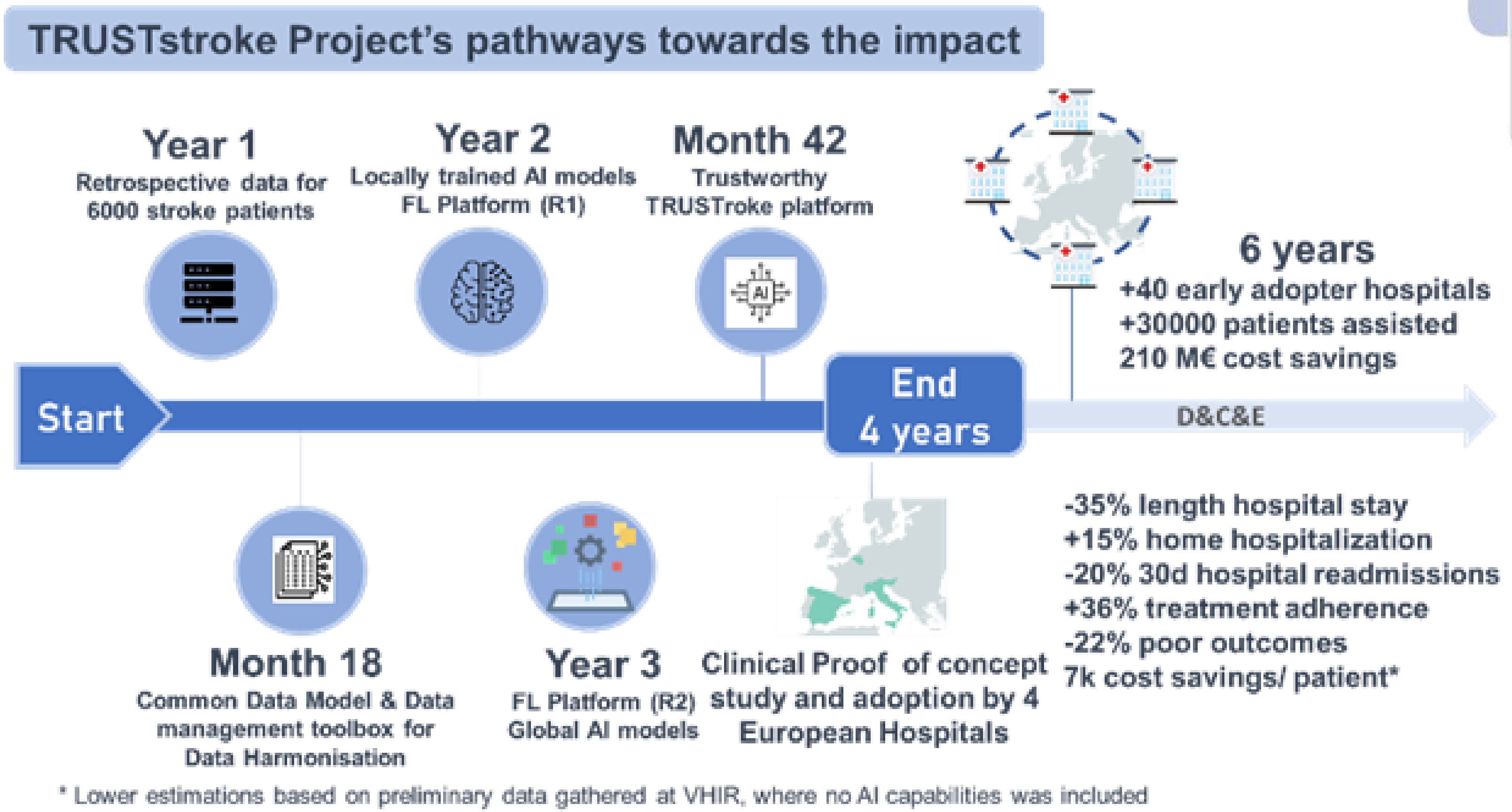
FL platform for secure, multicentric and privacy preserving AI training

Leader: CERN
Participants: CNR and POLIMI

Design, Develop and Deliver a secure and robust FL platform hosted and operated by CERN



TRUSTroke Project Planning



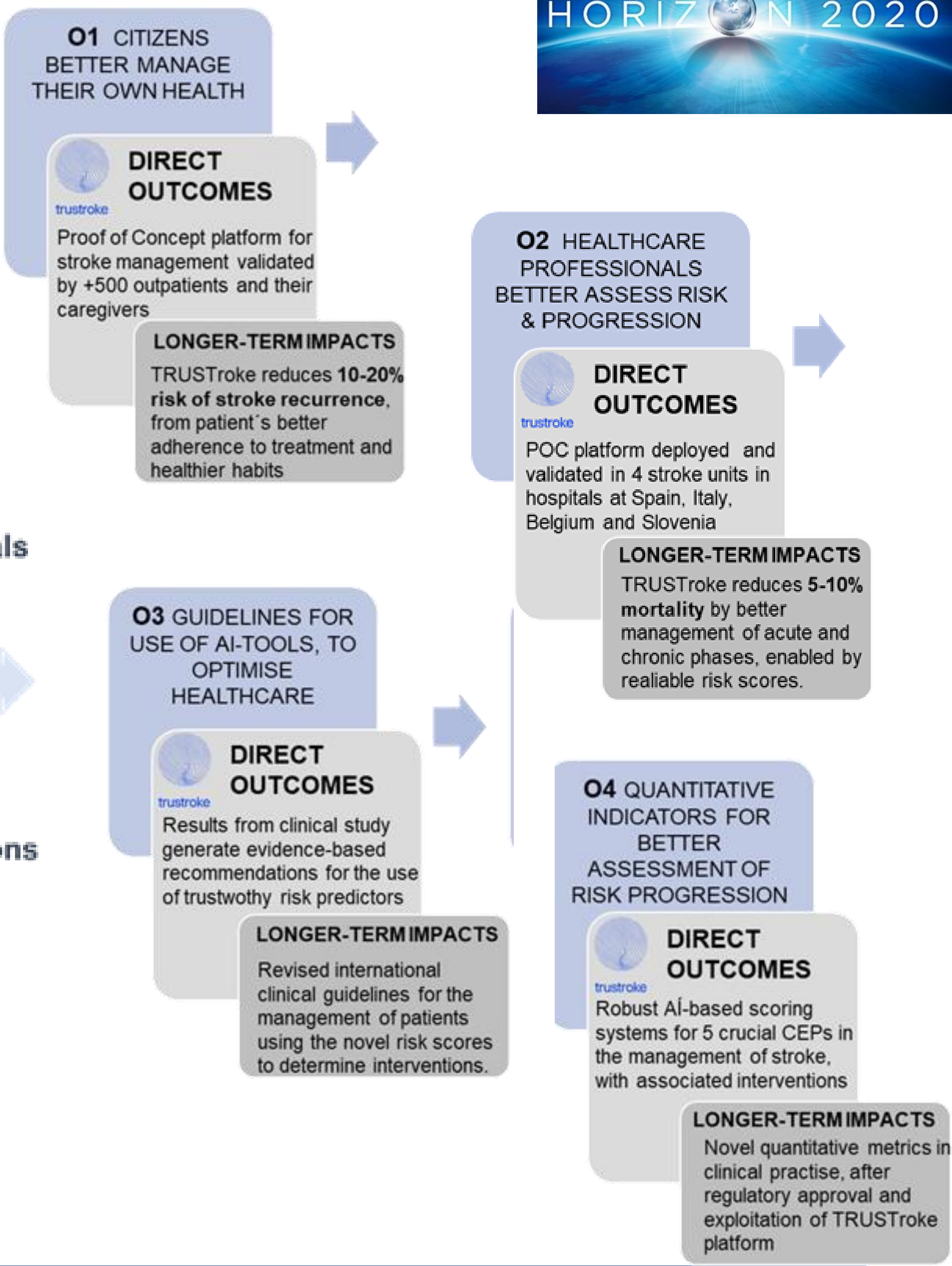
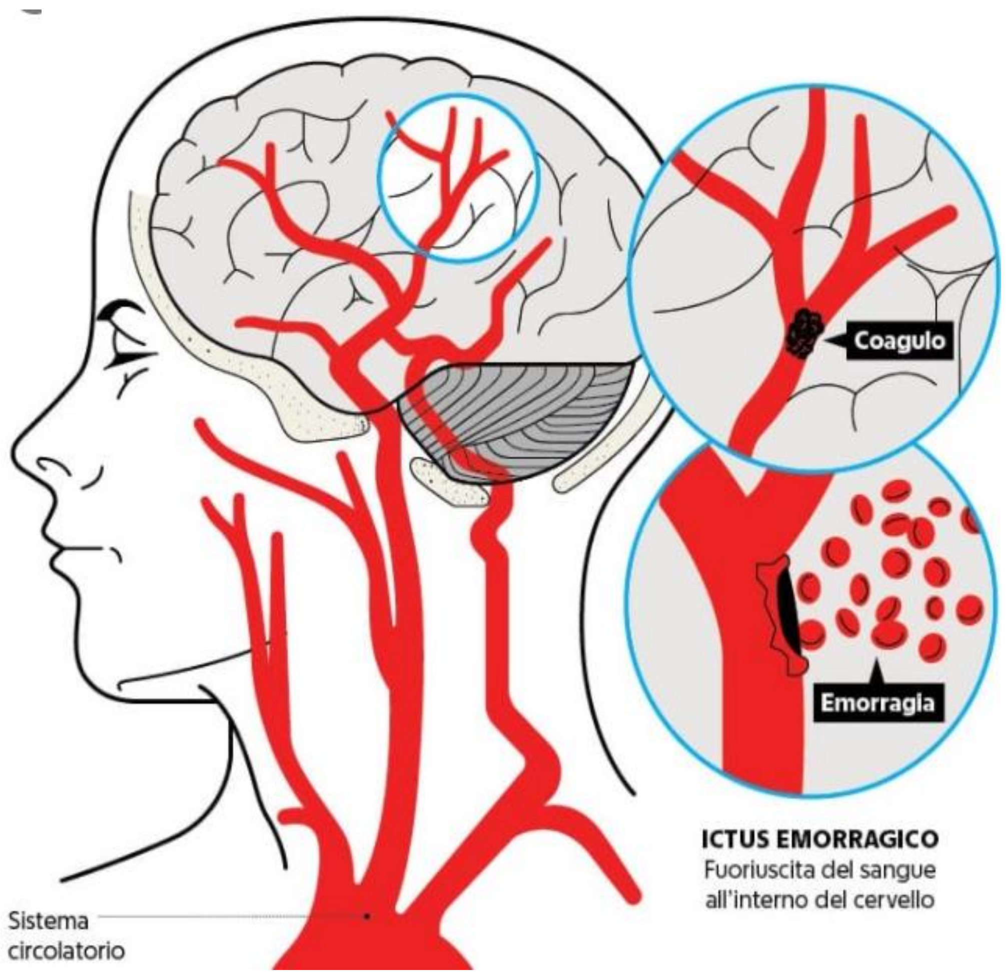


Table of Contents

Topics to cover



- Pathway/Standard of care in the acute phase (stroke code, reperfusion/treatments)

- Standard of care in the chronic phase (outcomes, follow-up, adherence to treatments, recurrence)

Emergency Department

Clinical Evaluation
Support of vital functions
Blood exams

Brain CT /angio-CT
Perfusion CT

Ischemia

hemorrhage

Fibrinolysis
+/-
Mecchanical Thrombectomy

TICI score

Table 1 **The proposed modified TICI score**

Score	Revised Thrombolysis in Cerebral Infarction Scale
0	No perfusion or anterograde flow beyond site of occlusion
1	Penetration but not perfusion. Contrast penetration exists past the initial obstruction but with minimal filling of the normal territory
2	Incomplete perfusion wherein the contrast passes the occlusion and opacifies the distal arterial bed but rate of entry or clearance from the bed is slower or incomplete when compared to non-involved territories
2A	Some perfusion with distal branch filling of <50% of territory visualized
2B	Substantial perfusion with distal branch filling of ≥ 50% of territory visualized
2C	Near complete perfusion except for slow flow in a few distal cortical vessels, or presence of small distal cortical emboli
3	Complete perfusion with normal filling of all distal branches

Acute Ischemic stroke

Emergency Department

Support of vital functions
Blood exams
Stroke team evaluation

CT cranio/angio-CT
Perfusion CT

hemorrhage

Ischemia

Fibrinolysis
+/-
Mecchanical Thrombectomy

TICI score

Stroke Unit

Multiparametric Monitoring

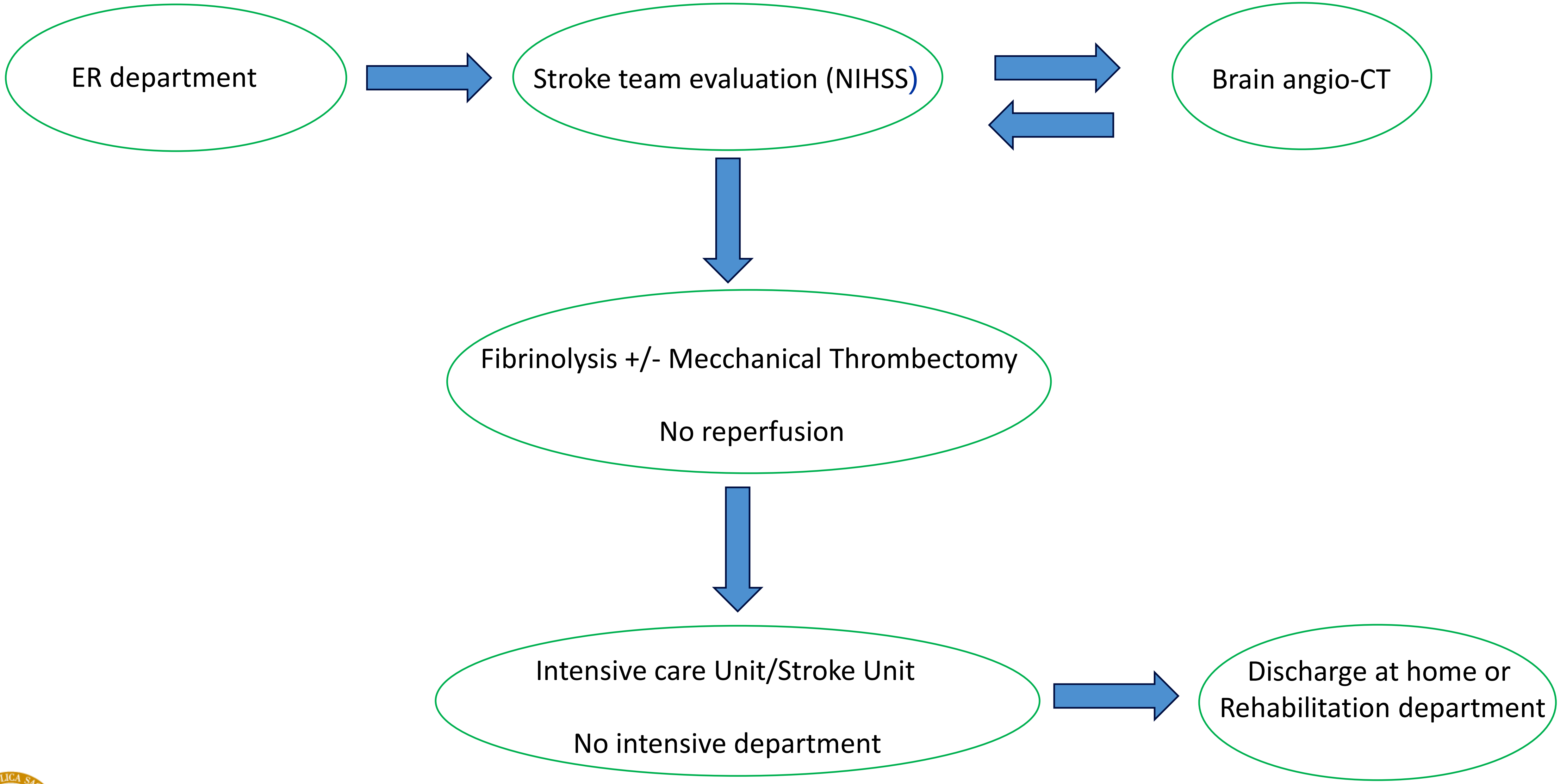
Therapies

Prevent neurological and not neurologicl complications

Search for causes of stroke
- Brain MRI
- Echocardiogram
- ECG Holter
- Blood Exams

RESEARCH

?


ER department
Stroke team evaluation (NIHSS)
Brain angio-CT
Fibrinolysis +/- Mecchanical Thrombectomy
No reperfusion
Intensive care Unit/Stroke Unit
No intensive department
Discharge at home or
Rehabilitation department

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Endovascular Therapy for Ischemic Stroke with Perfusion-Imaging Selection

CLASS (STRENGTH) OF RECOMMENDATION

CLASS I (STRONG) — Benefit >>> Risk

Suggested phrases for writing recommendations:
- Is recommended
- Is indicated/useful/effective/beneficial
- Should be performed/administered/other
- Comparative-Effectiveness Phrases†:
  - Treatment/strategy A is recommended/indicated in preference to treatment B
  - Treatment A should be chosen over treatment B

CLASS IIa (MODERATE) — Benefit >> Risk

Suggested phrases for writing recommendations:
- Is reasonable
- Can be useful/effective/beneficial
- Comparative-Effectiveness Phrases†:
  - Treatment/strategy A is probably recommended/indicated in preference to treatment B
  - It is reasonable to choose treatment A over treatment B

CLASS IIb (WEAK) — Benefit ≥ Risk

Suggested phrases for writing recommendations:
- May/might be reasonable
- May/might be considered
- Usefulness/effectiveness is unknown/unclear/uncertain or not well established

CLASS III: No Benefit (MODERATE) — Benefit = Risk
(Generally, LOE A or B use only)

Suggested phrases for writing recommendations:
- Is not recommended
- Is not indicated/useful/effective/beneficial
- Should not be performed/administered/other

CLASS III: Harm (STRONG) — Risk > Benefit

Suggested phrases for writing recommendations:
- Potentially harmful
- Causes harm
- Associated with excess morbidity/mortality
- Should not be performed/administered/other

LEVEL (QUALITY) OF EVIDENCE‡

LEVEL A
- High-quality evidence‡ from more than 1 RCT
- Meta-analyses of high-quality RCTs
- One or more RCTs corroborated by high-quality registry studies

LEVEL B-R (Randomized)
- Moderate-quality evidence‡ from 1 or more RCTs
- Meta-analyses of moderate-quality RCTs

LEVEL B-NR (Nonrandomized)
- Moderate-quality evidence‡ from 1 or more well-designed, well-executed nonrandomized studies, observational studies, or registry studies
- Meta-analyses of such studies

LEVEL C-LD (Limited Data)
- Randomized or nonrandomized observational or registry studies with limitations of design or execution
- Meta-analyses of such studies
- Physiological or mechanistic studies in human subjects

LEVEL C-EO (Expert Opinion)

Consensus of expert opinion based on clinical experience

COR and LOE are determined independently (any COR may be paired with any LOE).

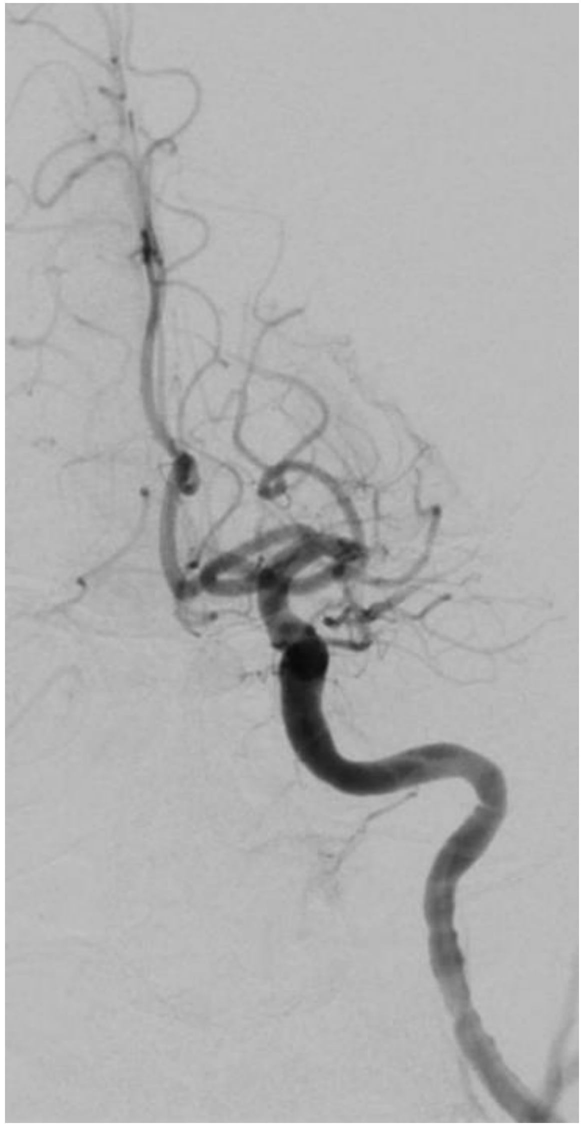
A recommendation with LOE C does not imply that the recommendation is weak. Many important clinical questions addressed in guidelines do not lend themselves to clinical trials. Although RCTs are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.

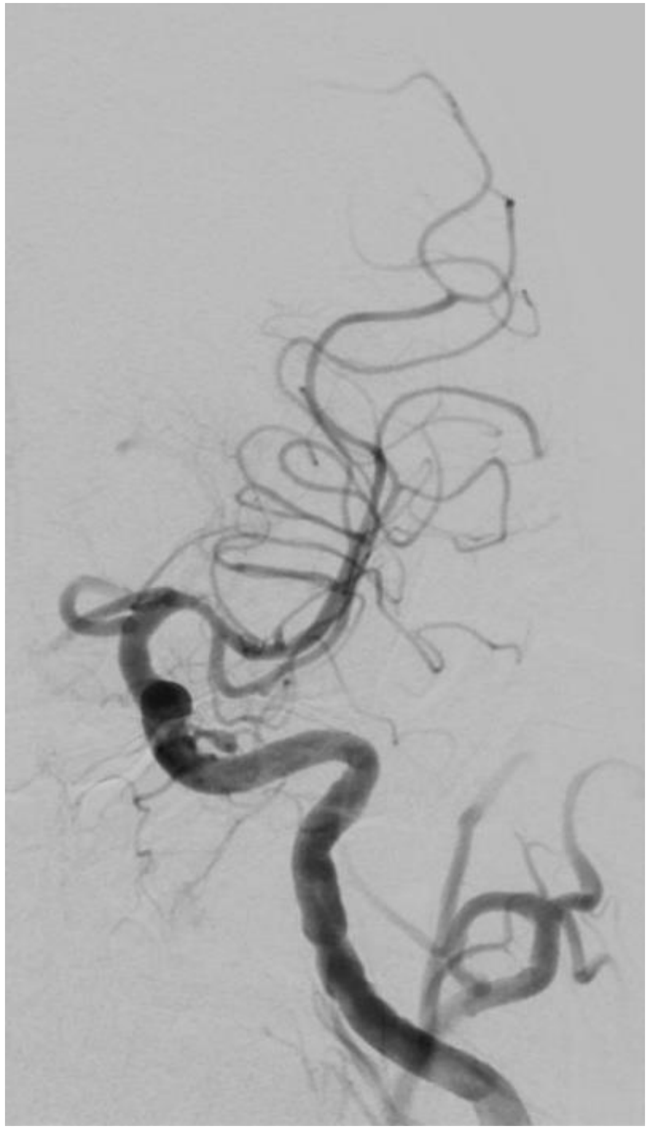
* The outcome or result of the intervention should be specified (an improved clinical outcome or increased diagnostic accuracy or incremental prognostic information).

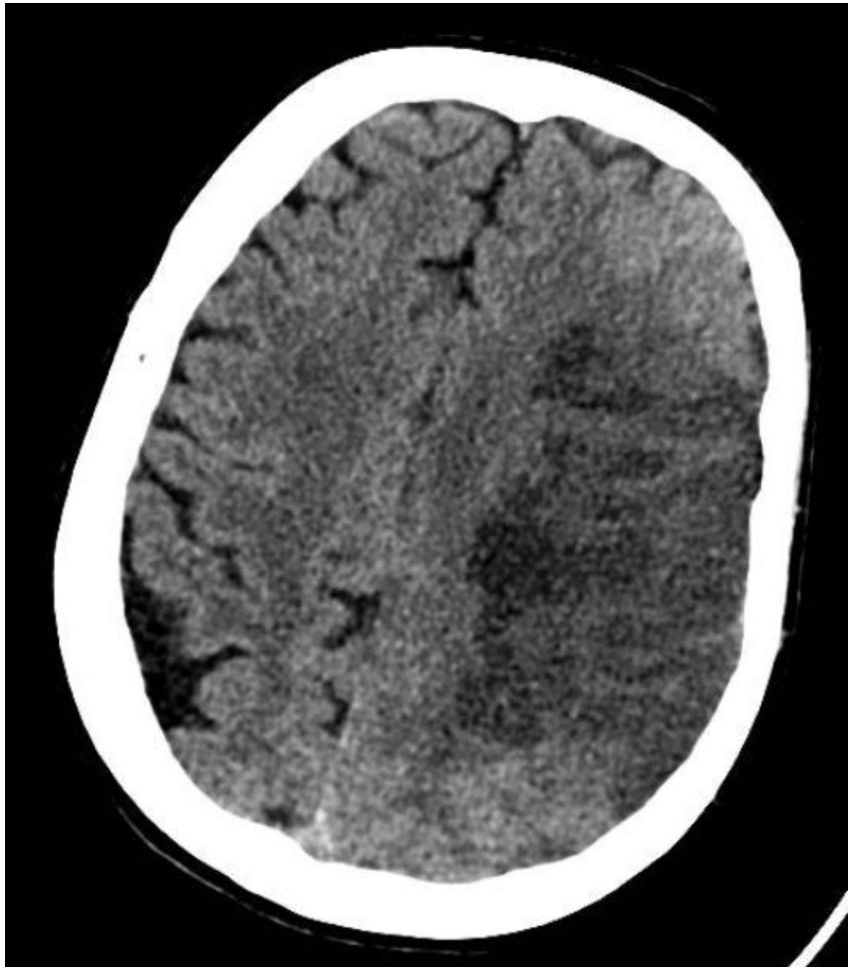
† For comparative-effectiveness recommendations (COR I and IIa; LOE A and B only), studies that support the use of comparator verbs should involve direct comparisons of the treatments or strategies being evaluated.

‡ The method of assessing quality is evolving, including the application of standardized, widely used, and preferably validated evidence grading tools; and for systematic reviews, the incorporation of an Evidence Review Committee.

COR indicates Class of Recommendation; EO, expert opinion; LD, limited data; LOE, Level of Evidence; NR, nonrandomized; R, randomized; and RCT, randomized controlled trial.

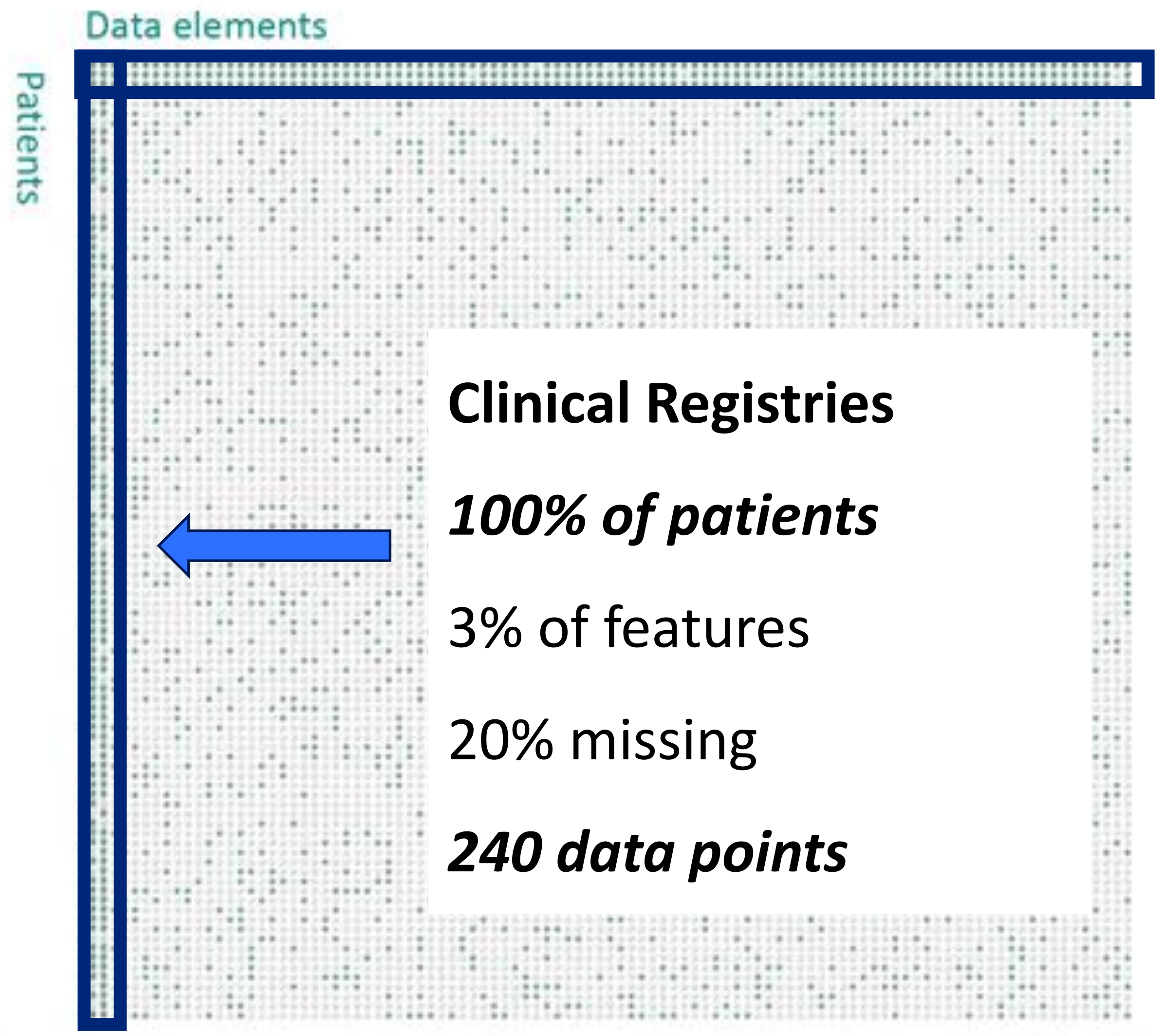






Right hemiplegia and global aphasia

AI & REAL WORLD DATA



Clinical Registries

100% of patients

3% of features

20% missing

240 data points

Clinical research

3% of patients

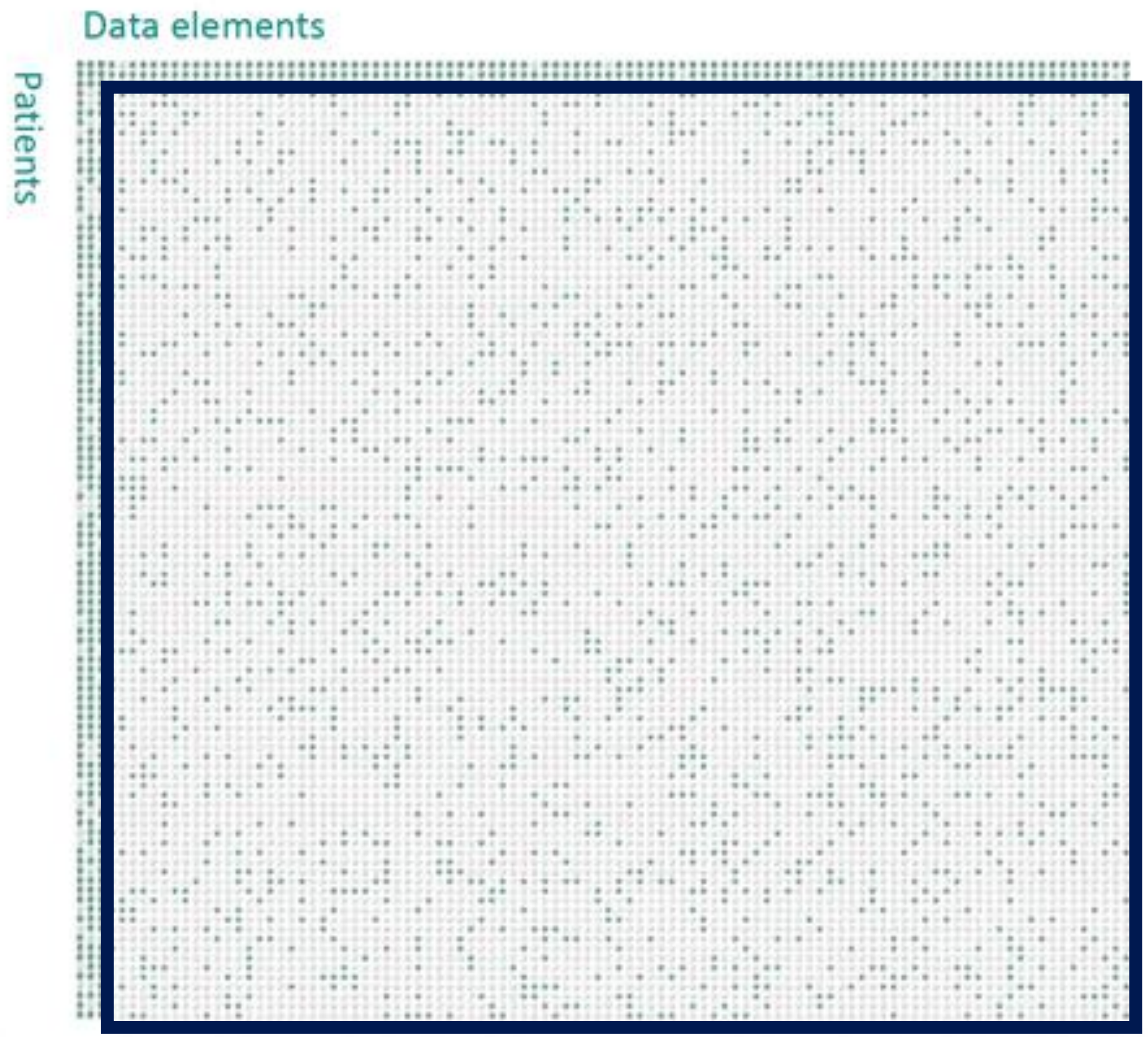
100% of features

5% missing

285 data points

Courtesy of A. Dekker – MAASTRO Clinic

AI & REAL WORLD DATA



Clinical routine

100% of patients

100% of features

80% missing

2000 data points

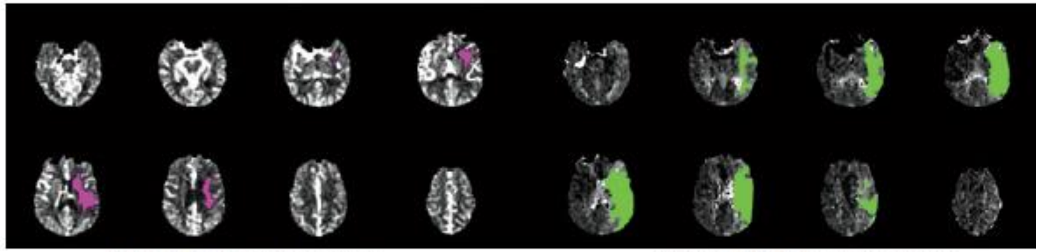
Courtesy of A. Dekker – MAASTRO Clinic

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

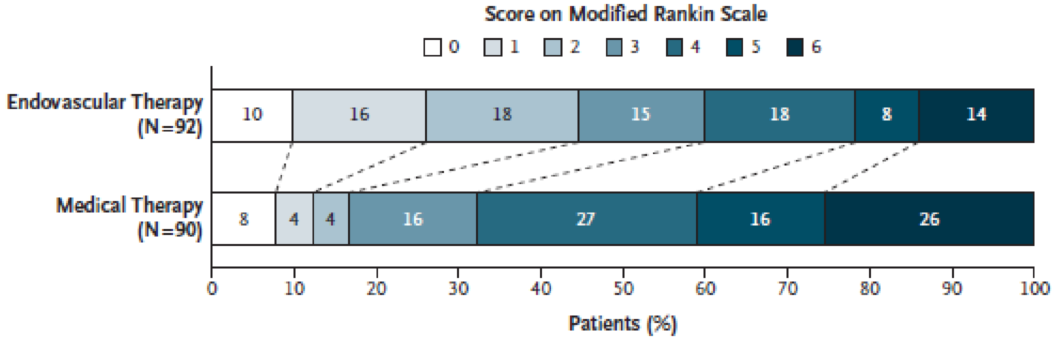
Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging

G.W. Albers, M.P. Marks, S. Kemp, S. Christensen, J.P. Tsai, S. Ortega-Gutierrez, R.A. McTaggart, M.T. Torbey, M. Kim-Tenser, T. Leslie-Mazwi, A. Sarraj, S.E. Kasner, S.A. Ansari, S.D. Yeatts, S. Hamilton, M. Mlynash, J.J. Heit, G. Zaharchuk, S. Kim, J. Carrozzella, Y.Y. Palesch, A.M. Demchuk, R. Bammer, P.W. Lavori, J.P. Broderick, and M.G. Lansberg, for the DEFUSE 3 Investigators*



Volume of Ischemic Core, 23 ml

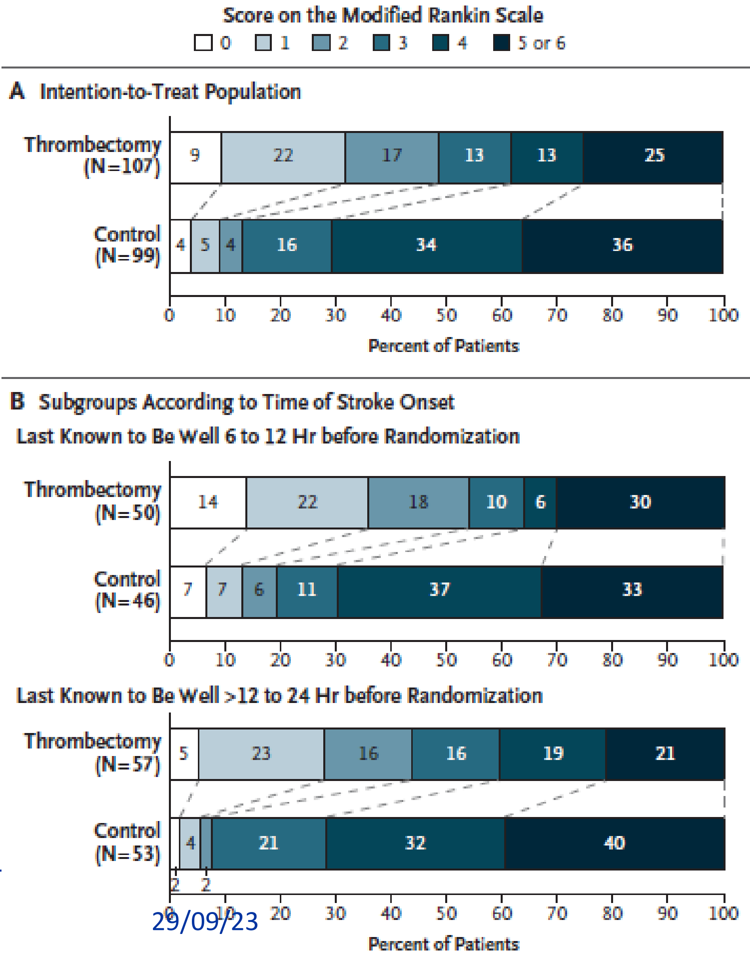
Volume of Perfusion Lesion, 128 ml



The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 JANUARY 4, 2018 VOL. 378 NO. 1

Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct

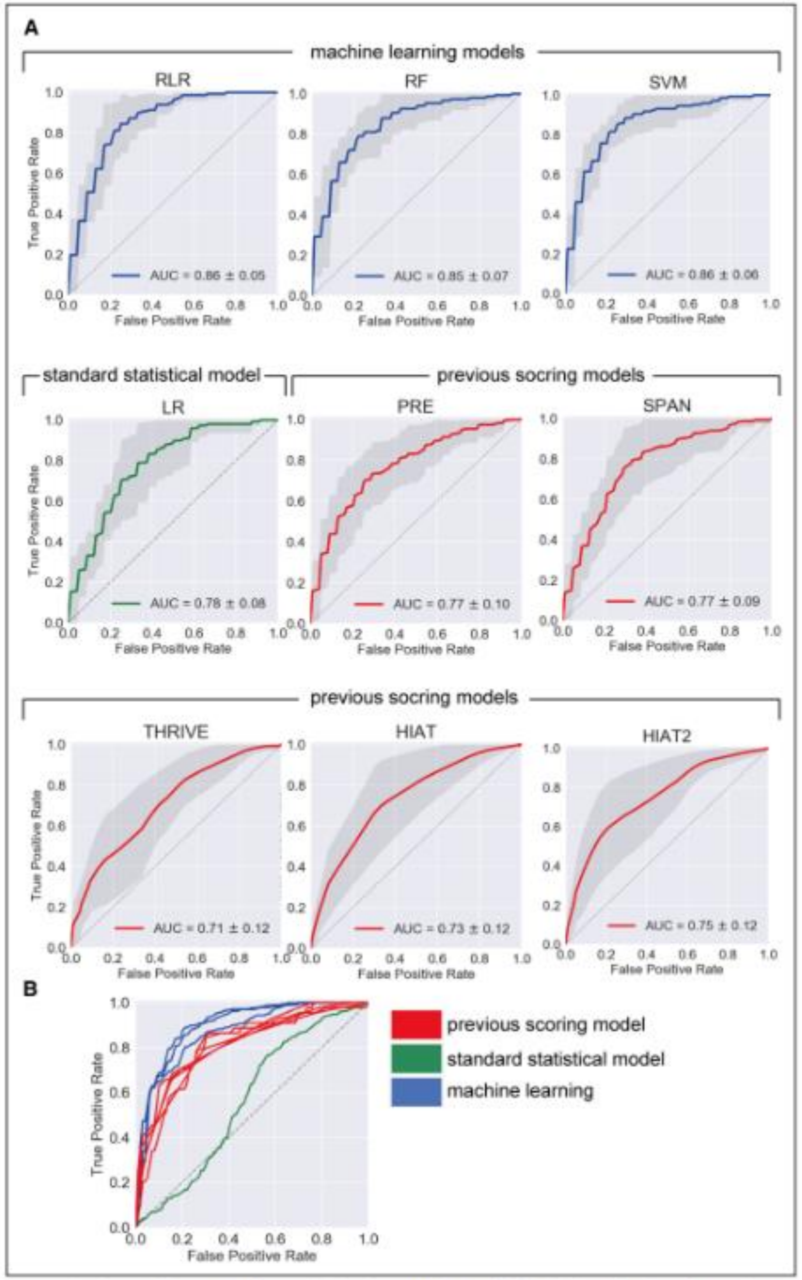
R.G. Nogueira, A.P. Jadhav, D.C. Haussen, A. Bonafe, R.F. Budzik, P. Bhuva, D.R. Yavagal, M. Ribo, C. Cognard, R.A. Hanel, C.A. Sila, A.E. Hassan, M. Millan, E.I. Levy, P. Mitchell, M. Chen, J.D. English, Q.A. Shah, F.L. Silver, V.M. Pereira, B.P. Mehta, B.W. Baxter, M.G. Abraham, P. Cardona, E. Veznedaroglu, F.R. Hellinger, L. Feng, J.F. Kirmani, D.K. Lopes, B.T. Jankowitz, M.R. Frankel, V. Costalat, N.A. Vora, A.J. Yoo, A.M. Malik, A.J. Furlan, M. Rubiera, A. Aghaebrahim, J.-M. Olivot, W.G. Tekle, R. Shields, T. Graves, R.J. Lewis, W.S. Smith, D.S. Liebeskind, J.L. Saver, and T.G. Jovin, for the DAWN Trial Investigators*

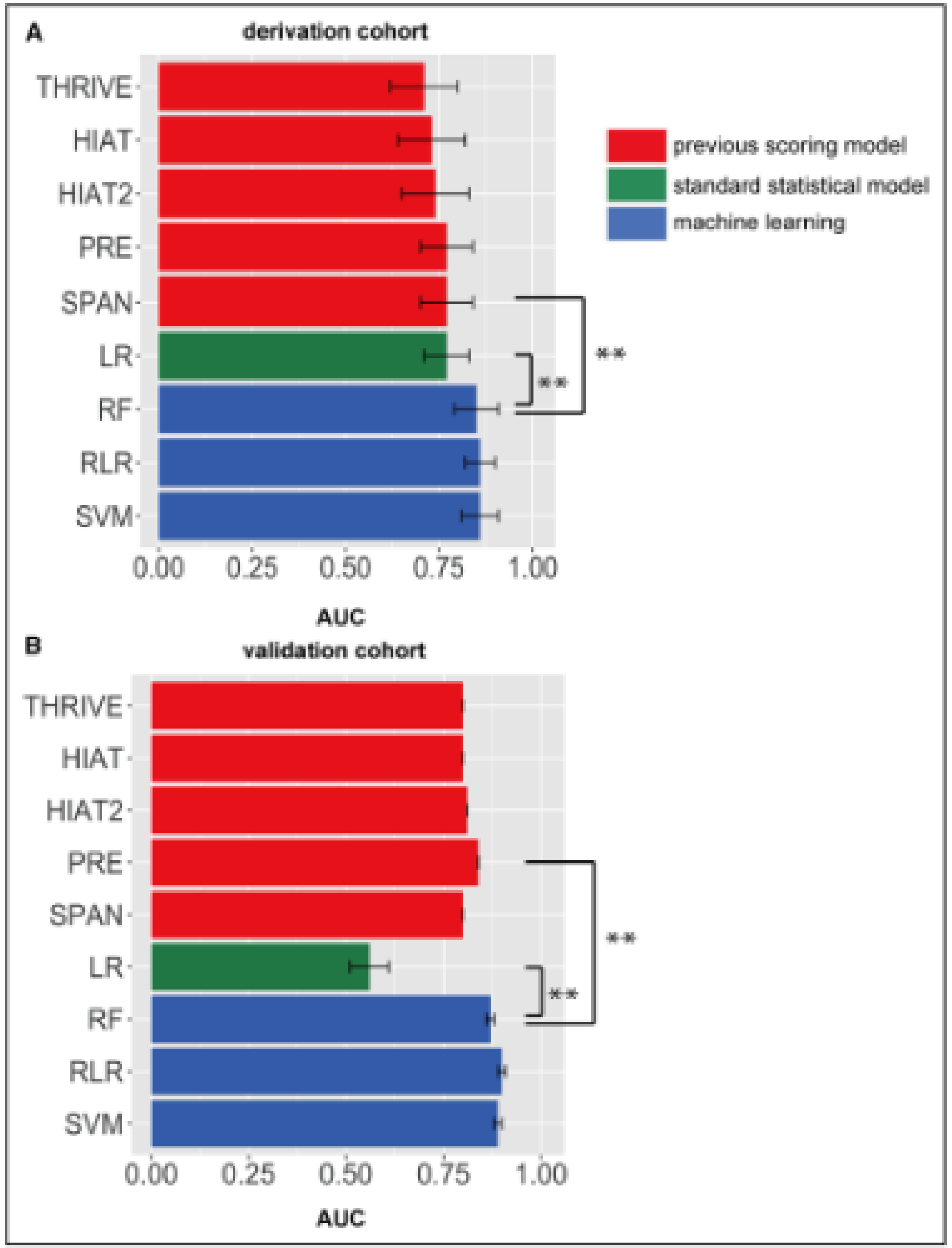


Predicting Clinical Outcomes of Large Vessel Occlusion Before Mechanical Thrombectomy Using Machine Learning

Hidehisa Nishi, MD; Naoya Oishi, MD, PhD; Akira Ishii, MD, PhD; Isao Ono, MD;
Takenori Ogura, MD, PhD; Tadashi Sunohara, MD; Hideo Chihara, MD, PhD;
Ryu Fukumitsu, MD, PhD; Masakazu Okawa, MD, PhD; Norikazu Yamana, MD, PhD;
Hirotoshi Imamura, MD, PhD; Nobutake Sadamasa, MD, PhD; Taketo Hatano, MD, PhD;
Ichiro Nakahara, MD, PhD; Nobuyuki Sakai, MD, PhD; Susumu Miyamoto, MD, PhD

The derivation cohort included 387 LVO patients, and the external validation cohort included 115 LVO patients
with anterior circulation who were treated with mechanical thrombectomy.





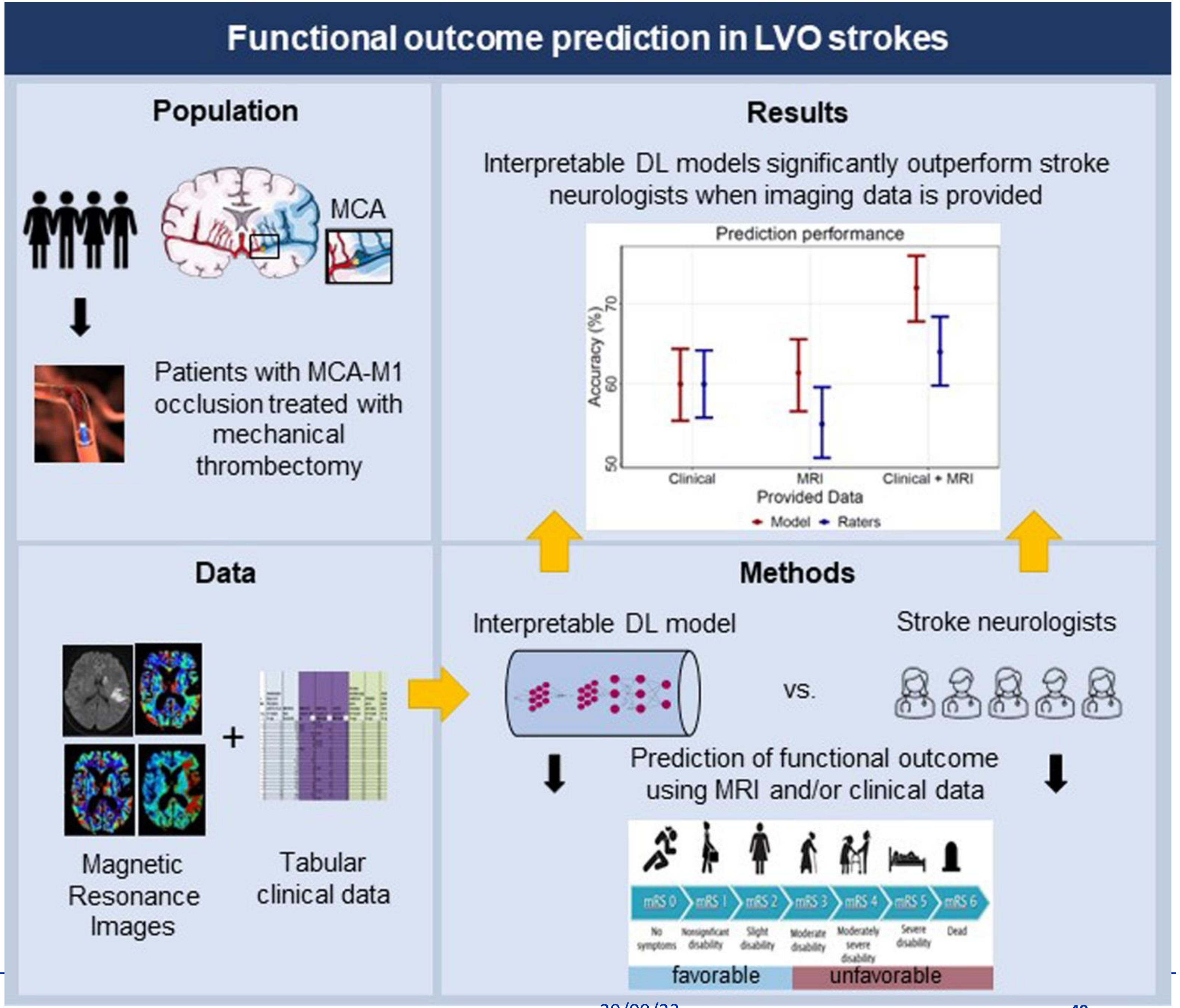
Stroke

CLINICAL AND POPULATION SCIENCES

Deep Learning Versus Neurologists: Functional Outcome Prediction in LVO Stroke Patients Undergoing Mechanical Thrombectomy

Lisa Herzog, Dr sc. nat.; Lucas Kook, Dr. sc. nat.; Janne Hamann, Dr med; Christoph Globas, PD Dr med; Mirjam R. Heldner, Dr med; David Seiffge, Dr med; Kateryna Antonenko, PD Dr med; Tomas Dobrocky, Dr med; Leonidas Panos, Dr med; Johannes Kaesmacher, PD Dr med; Urs Fischer, Prof Dr med; Jan Gralla, Dr med; Marcel Arnold, Dr med; Roland Wiest, Dr med; Andreas R. Luft, Dr med; Beate Sick, Dr; Susanne Wegener, Dr med

Collected data of 222 patients with middle cerebral artery M1 segment occlusion who received mechanical thrombectomy



frontiers in Neurology

ORIGINAL RESEARCH
published: 25 September 2018
doi: 10.3389/fneur.2018.00784

Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms

Hendrikus J. A. van Os[1*], Lucas A. Ramos[2,3], Adam Hilbert[2], Matthijs van Leeuwen[4], Marianne A. A. van Walderveen[5], Nyika D. Kruyt[1], Diederik W. J. Dippel[6], Ewout W. Steyerberg[7,8], Irene C. van der Schaaf[9], Hester F. Lingsma[8], Wouter J. Schonewille[10], Charles B. L. M. Majoie[11], Silvia D. Olabarriaga[3],

Characteristics	All patients ($n = 1,383$)
Mean age ± SD *(years)*	69.8 ± 14.4
Men, *n (%)*	738 (53.5)
NIHSS score, *median (IQR)**	16 (11–20)
Mean systolic blood pressure ± SD *(mm Hg)*	150 ± 25
MEDICAL HISTORY, *N (%)*	
Atrial fibrillation	411 (30.7)
Hypertension	697 (51.1)
Diabetes mellitus	235 (17.1)
Myocardial infarction	216 (15.9)
Peripheral artery disease	127 (9.4)
Ischaemic stroke	227 (16.5)
Hypercholesterolemia	411 (29.7)
Pre-stroke mRS > 2, *n (%)*	158 (11.6)
Smoking, *n (%)*	314 (22.9)
MEDICATION USE, *N (%)*	
DOAC**	35 (2.6)
Coumarine	179 (13.0)
Antiplatelet	461 (33.7)
Heparin	52 (3.8)
Blood pressure medication	707 (52.1)
Statin	490 (36.2)
Intravenous alteplase treatment, *n (%)*	1,054 (76.2)
ASPECTS, *median (IQR)*	9 (7–10)
Time from stroke onset to groin in minutes, *median (IQR)*	210 (160–270)
Collateral score ≥ 2	764 (55)

1,383 EVT patients included patients from the Multicenter Randomized Clinical Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands (MR CLEAN) Registry

TABLE 2 | Discrimination of machine learning algorithms and logistic regression models across the various prediction settings.

Models, AUC (95% CI)*	**Prediction setting (used variables: predicted outcome)**		
	Baseline: post-mTICI	**Baseline: mRS**	**All variables: mRS**
Super learner	0.55 (0.54–0.56)	0.79 (0.79–0.80)	0.90 (0.90–0.91)
Random forests	0.55 (0.55–0.56)	0.79 (0.79–0.79)	0.91 (0.90–0.91)
Support vector machine	0.53 (0.53–0.54)	0.78 (0.77–0.78)	0.88 (0.88–0.89)
Neural network	0.53 (0.53–0.54)	0.77 (0.76–0.77)	0.88 (0.88–0.89)
LR: AUTOMATED SELECTION**			
Random forests	0.55 (0.55–0.56)	0.78 (0.78–0.78)	0.90 (0.90–0.90)
LASSO	NA¥	0.78 (0.78–0.79)	0.90 (0.89–0.90)
Elastic net	NA¥	0.77 (0.77–0.78)	0.89 (0.88–0.89)
Backward elimination	0.57 (0.57–0.58)	0.78 (0.77–0.78)	0.90 (0.89–0.90)
LR: prior knowledge‡	0.55 (0.55–0.58)	0.78 (0.78–0.79)	0.90 (0.90–0.90)

**Model discrimination is assessed by calculating mean Area Under the Curve (AUC) of the receiver operating characteristic across all outer cross-validation folds.*
***Logistic regression using automated variable selection methods.*
¥ Variable selection not possible, likely due to insufficient signal-to-noise ratio.
‡Logistic regression using variables based on prior knowledge.

Negligible difference of mean AUC (0.01; 95%CI: 0.00–0.01) between best performing machine learning algorithm (Random Forests) and best performing logistic regression model

Prediction of mRS score at 90 days

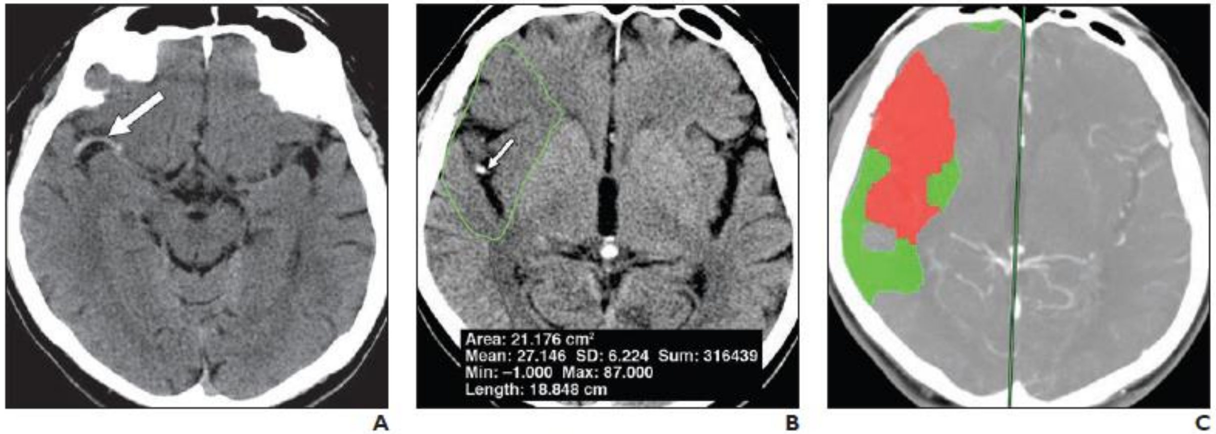
Neuroradiology/Head and Neck Imaging • Original Research

Use of Gradient Boosting Machine Learning to Predict Patient Outcome in Acute Ischemic Stroke on the Basis of Imaging, Demographic, and Clinical Information

A total of 512 patients were enrolled in this retrospective study

TABLE 1: All Features and Selected Features at Two Treatment Time Points for Extreme Gradient Boosting (XGB) and Gradient Boosting Machine (GBM) Models

Features Used in XGB and GBM Models	Time Point: At Admission	Time Point: At 24 Hours
All features	Sex	Sex
	Age	Age
	Total ASPECTS	Total ASPECTS
	Specific ASPECTS[a]	Specific ASPECTS[a]
	Collaterals status	Collaterals status
	HMCAS present	HMCAS present
	Infarct volume on perfusion CT	Infarct volume on perfusion CT
	Penumbra volume on perfusion CT	Penumbra volume on perfusion CT
	Side of stroke	Side of stroke
	Occluded vessels	Occluded vessels
	Baseline TIMI score	Baseline TIMI score
	NASCET degree of stenosis on left side	NASCET degree of stenosis on left side
	NASCET degree of stenosis on right side	NASCET degree of stenosis on right side
	NIHSS score at baseline	NIHSS score at baseline
		NIHSS score at 24 hours[b]
Selected features[c]	NIHSS score at baseline	NIHSS score at 24 hours
	HMCAS present	ASPECTS of M5
	NASCET degree of stenosis on left side	Age
	Age	NASCET degree of stenosis on left side
	ASPECTS of M3 branch of MCA	HMCAS present
	ASPECTS of caudate	
	Baseline TIMI score	



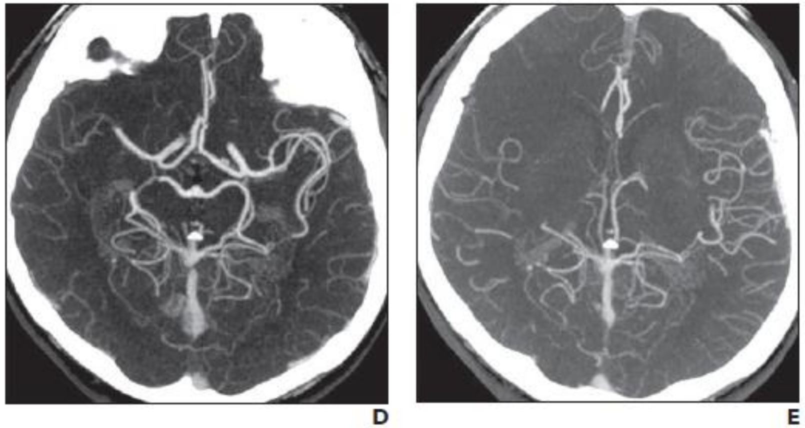


Fig. 1—Feature extraction based on imaging biomarkers from unenhanced CT, perfusion CT, and CT angiography (CTA) of 58-year-old man.
A, On unenhanced CT image, right hyperdense middle cerebral artery sign (*arrow*) can be clearly defined and Alberta stroke program early CT score (ASPECTS) was 6.
B, On unenhanced CT image, ASPECTS is 6. Arrow shows right hyperdense middle cerebral artery sign. Outlined area shows manually depicted area of infarction by experienced radiologist. Min = minimum, Max = maximum.
C, Perfusion CT image shows large-area infarction in middle cerebral artery territory. Red area shows infarct core, and green area shows infarct penumbra. Green line is symmetry axis of brain.
D and **E**, On CTA images, right M1 and M2 are occluded, with thrombolysis in myocardial infarct score 0 (**D**) and collateral score of 1 (**E**).

Prediction of mRS score at 90 days

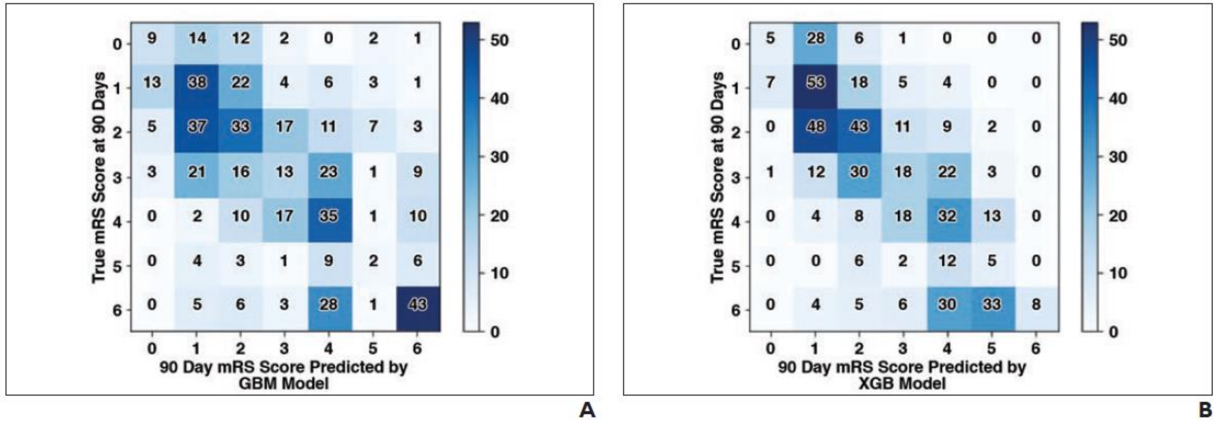
Neuroradiology/Head and Neck Imaging • Original Research

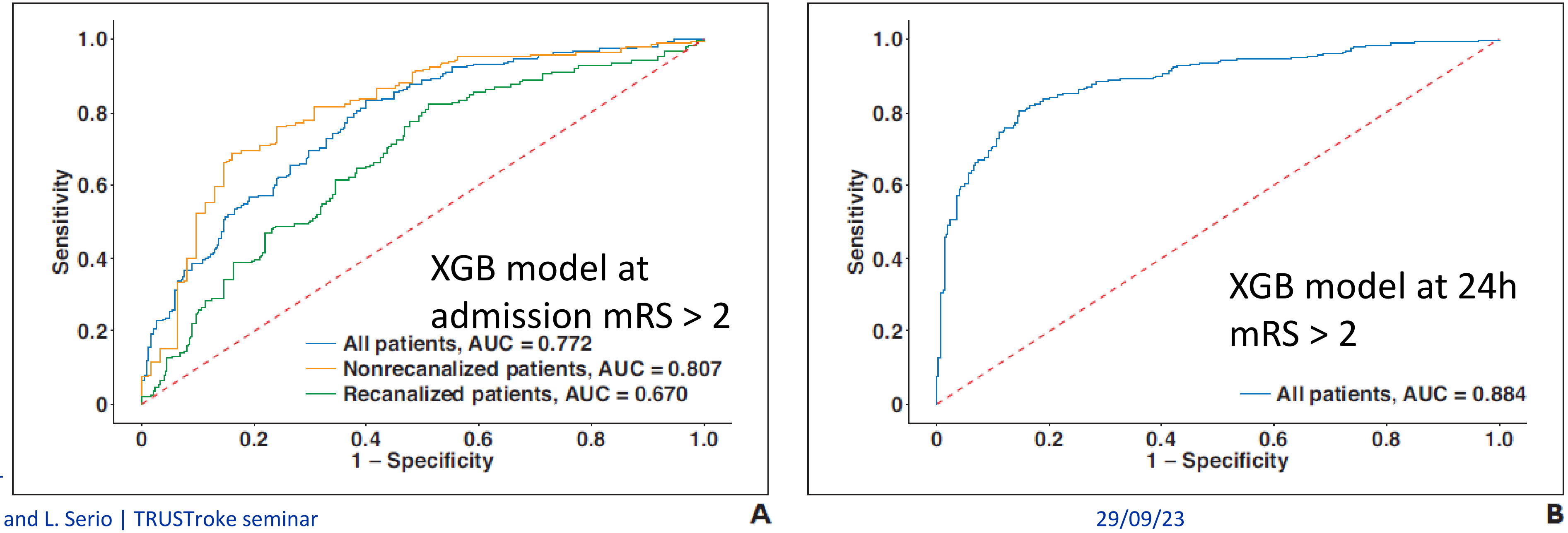
Use of Gradient Boosting Machine Learning to Predict Patient Outcome in Acute Ischemic Stroke on the Basis of Imaging, Demographic, and Clinical Information

TABLE 2: Binomial Prediction Results With All Available Features at 24 Hours and 23 Features at Admission

Features	AUC					
	mRS Score > 0	mRS Score > 1	mRS Score > 2	mRS Score > 3	mRS Score > 4	mRS Score > 5
24 Features						
XGB	0.821	0.812	0.873	0.840	0.840	0.867
GBM	0.798	0.787	0.849	0.825	0.781	0.831
23 Features (without 24-hour NIHSS score)						
XGB	0.778	0.713	0.746	0.770	0.759	0.774
GBM	0.757	0.681	0.748	0.762	0.735	0.758

Note—mRS = modified Rankin scale, XGB = extreme gradient boosting, GBM = gradient boosting machine, NIHSS = National Institutes of Health Stroke Scale.





Prediction of mRS score at 90 days

Computer Methods and Programs in Biomedicine 190 (2020) 105381

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry

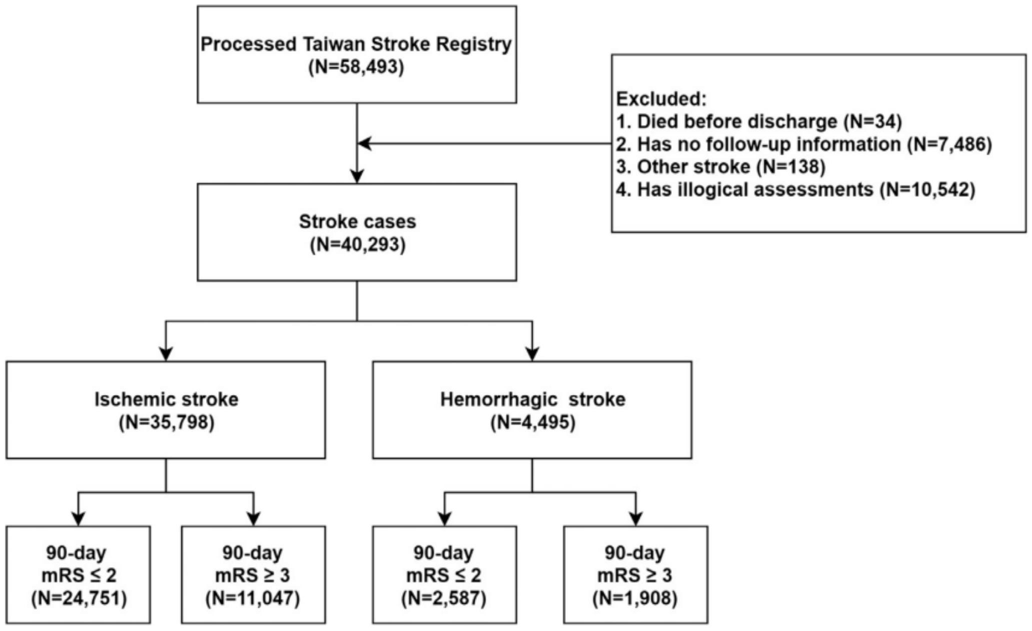
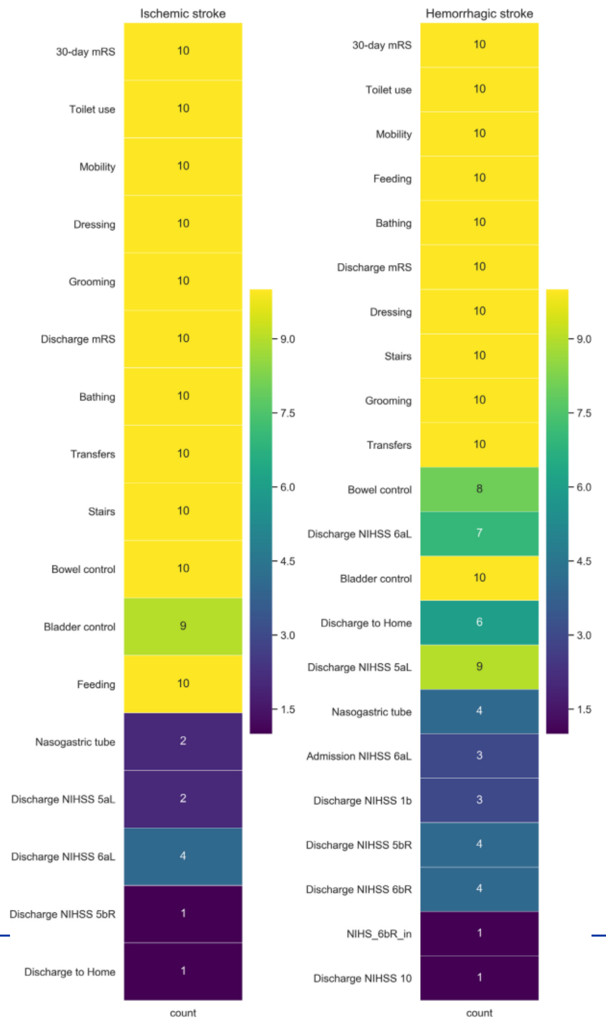


Fig. 1. Flowchart of patient recruitment. mRS indicates modified Rankin Scale.

Preadmission data, and inpatient elements including clinical care during hospitalization in-hospital complications, stroke risk factors, laboratory results of blood tests, electrocardiography, computed tomography (CT) and magnetic resonance imaging (MRI) finding, 30 days follow-up data.



Selected features:

- Discharge NIHSS assessment items
- discharge Barthel index
- the 30-day mRS degree

Prediction of mRS score at 90 days

Computer Methods and Programs in Biomedicine 190 (2020) 105381

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

ELSEVIER

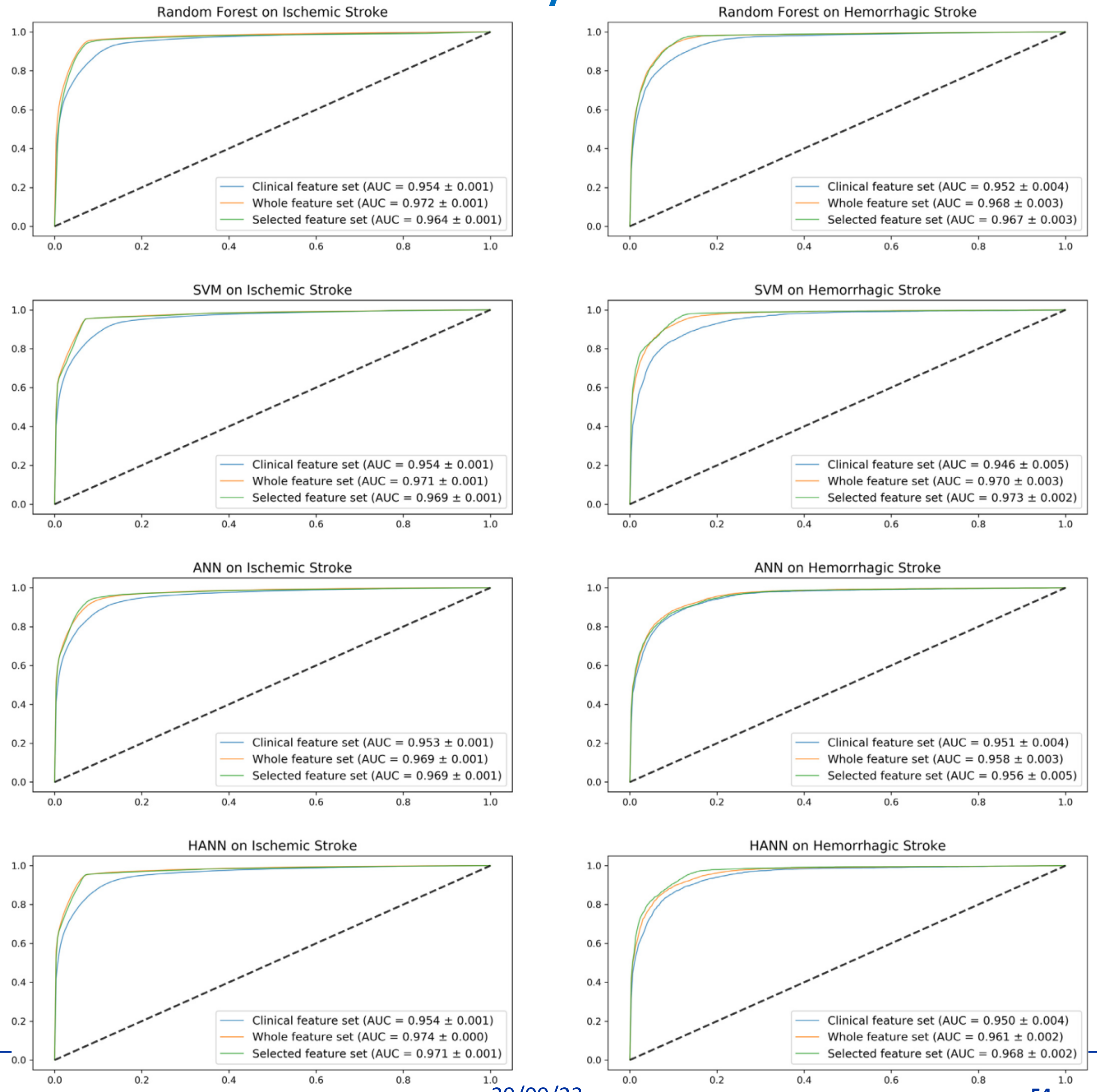
Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry

Table 2
The 10-time holdout testing result of machine learning models on 90-day stroke outcome prediction.

		Ischemic stroke (Mean ± SD)%			Hemorrhagic stroke (Mean ± SD)%		
		Precision	Recall	f1-score	Precision	Recall	f1-score
Clinical feature set (n = 203)	ANN	86.5 ± 0.2	89.2 ± 0.3	87.6 ± 0.6	87.9 ± 0.9	88.3 ± 1.0	88.0 ± 0.9
	RF	86.6 ± 0.2	89.8 ± 0.2	87.8 ± 0.2	87.7 ± 0.7	88.4 ± 0.8	87.9 ± 0.8
	HANN	86.2 ± 0.2	89.4 ± 0.2	87.7 ± 0.2	87.9 ± 0.7	88.2 ± 0.7	88.1 ± 0.7
	SVM	86.7 ± 0.2	89.4 ± 0.2	87.8 ± 0.2	87.0 ± 0.8	87.4 ± 0.8	87.1 ± 0.8
Whole feature set (n = 206)	ANN	89.2 ± 0.6	91.9 ± 0.2	90.3 ± 0.5	88.6 ± 1.0	89.1 ± 0.9	88.8 ± 0.9
	RF	94.1 ± 0.1	93.9 ± 0.1	92.4 ± 0.1	91.2 ± 0.6	92.0 ± 0.6	91.4 ± 0.6
	HANN	91.5 ± 0.5	93.9 ± 0.3	92.6 ± 0.4	89.0 ± 0.7	89.5 ± 0.6	89.2 ± 0.7
	SVM	91.9 ± 0.2	94.2 ± 0.1	92.9 ± 0.1	90.2 ± 0.7	90.9 ± 0.6	90.2 ± 0.8
Selected feature set (n = 17/22)	ANN	90.0 ± 0.7	92.8 ± 0.4	91.1 ± 0.6	88.1 ± 1.2	88.5 ± 1.2	88.2 ± 1.2
	RF	91.3 ± 0.3	93.4 ± 0.3	92.2 ± 0.3	91.6 ± 0.4	92.4 ± 0.5	91.7 ± 0.4
	HANN	91.8 ± 0.3	94.1 ± 0.1	92.8 ± 0.2	90.1 ± 0.4	90.7 ± 0.6	90.1 ± 0.4
	SVM	91.9 ± 0.2	94.2 ± 0.1	92.9 ± 0.1	91.6 ± 0.3	92.4 ± 0.3	91.7 ± 0.3

SVM: support vector machine, RF: random forest, ANN: artificial neural network, HANN: hybrid artificial neural network.
clinical feature set: preadmission and inpatient data, whole feature set: preadmission, inpatient and follow-up data, selected feature set: feature selection on whole feature set.

Excluded:
1. **Died before discharge (N=34)**
2. **Has no follow-up information (N=7,486)**
3. **Other stroke (N=138)**
4. **Has illogical assessments (N=10,542)**



Prediction of mRS score at 90 days

Machine Learning–Based Model for Prediction of Outcomes in Acute Stroke

JoonNyung Heo, MD*; Jihoon G. Yoon, MD*; Hyungjong Park, MD; Young Dae Kim, MD, PhD; Hyo Suk Nam, MD, PhD; Ji Hoe Heo, MD, PhD

(*Stroke*. 2019;50:1263-1265. DOI: 10.1161/STROKEAHA.118.024293.)

This was a retrospective study using a prospective cohort that enrolled patients with acute ischemic stroke.
Favorable outcome was defined as modified Rankin Scale score 0, 1, or 2 at 3 months.

Supplemental Table I. List of collected variables. Numbered variables were used as input for model design.

Patient demographics	26. Atrial fibrillation
1. Age (years)	27. Myocardial infarction
2. Sex	28. Chronic heart failure
3. Smoking status *	29. Cancer
Clinical variables	30. Previous stroke
4. Time from onset to admission (hours)	**Medication history**
5-17. Individual NIHSS scores **	31. Statin ****
18. TOAST classification	32. Antiplatelet *****
19. Systolic blood pressure (mmHg) ***	33. Anticoagulation ******
20. Diastolic blood pressure (mmHg) ***	**Laboratory values**
Previous diseases	34. Hemoglobin (mg/dL)
21. Pre-stroke mRS	35. White blood cell count ($10^3/\mu L$)
22. Hypertension	36. Platelet count ($10^3/\mu L$)
23. Diabetes	37. Prothrombin time (INR)
24. Hypercholesterolemia	38. Glucose level (mg/dL)
25. Metabolic syndrome	**3-month mRS**

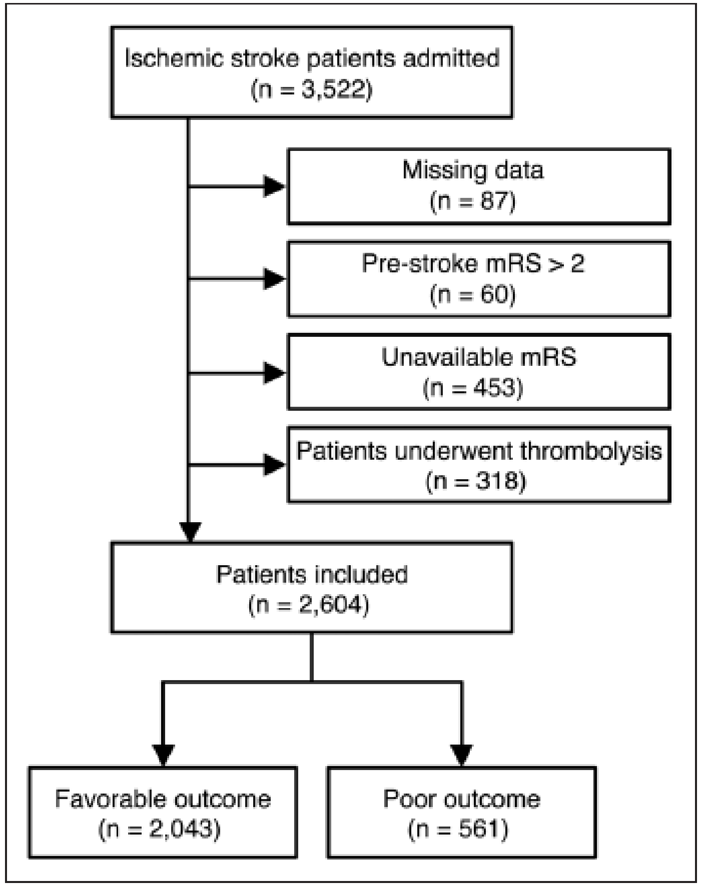
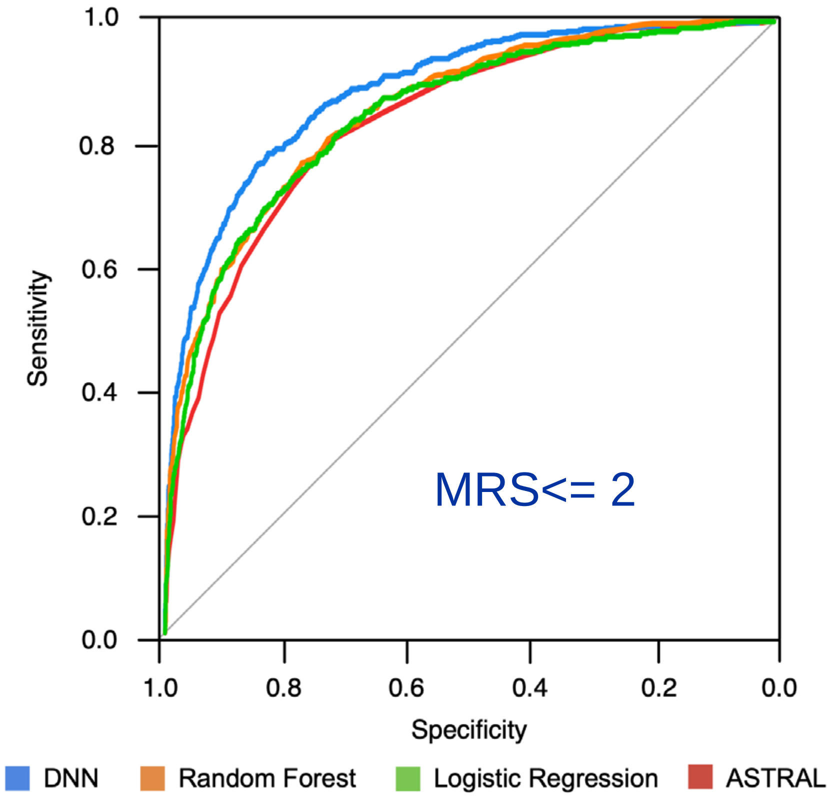


Figure 1. Flow chart illustrating patient selection. mRS indicates modified Rankin Scale.



Most studies have used retrospective data with small sample sizes.

There is a clear need for larger and especially prospective

evaluations building on the successful proof-of-concept reports

Some concerns

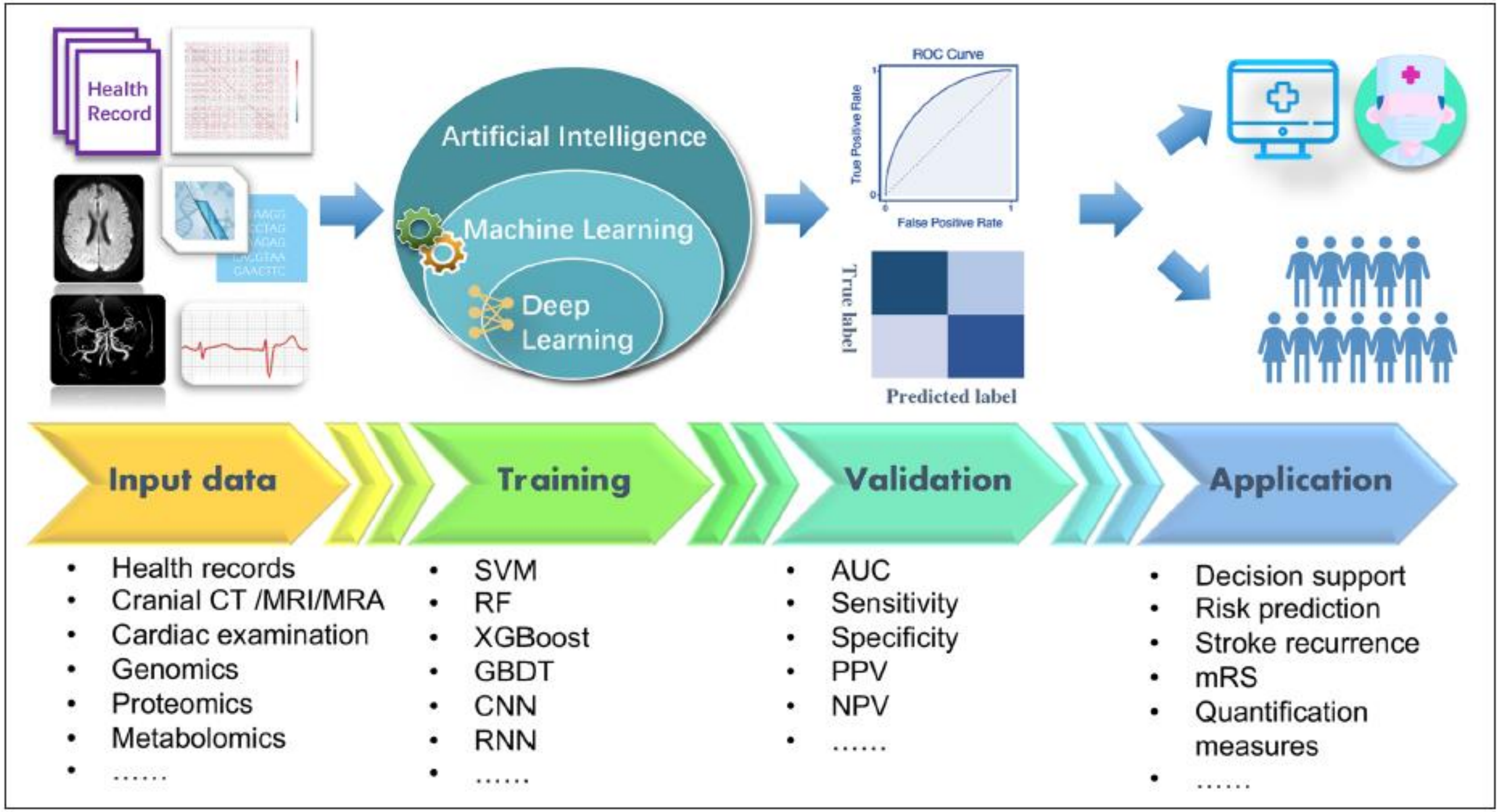
It is not yet known how well AI can perform in the health care setting

The use of AI systems in hospitals may raise new ethical and legal questions

Physicians may face liability issues when she or he follows the AI's recommendations

How integrate AI in International Guidelines?

If integrated into the workflow of physicians, AI can enable physicians to identify medical treatments tailored to each patient



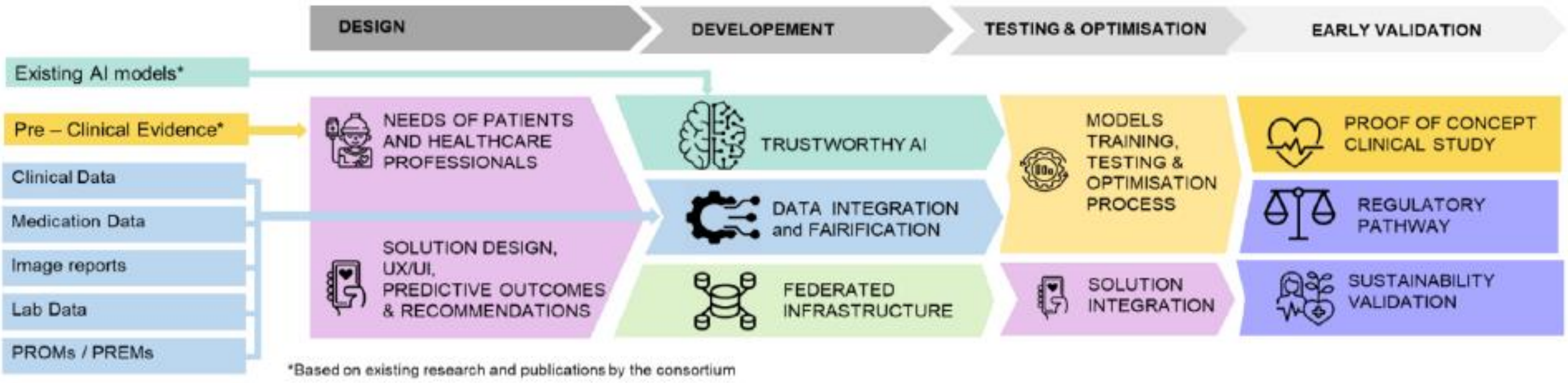
Decision Support System (DSS)

European Commission

Horizon Europe (HORIZON)
Call: HORIZON-HLTH-2022-STAYHLTH-01-two-stage
Project: 101080564 – TRUSTroke

trustroke

TRUSTWORTHY AI FOR IMPROVEMENT OF STROKE OUTCOMES
TRUSTROKE



Crucial Clinical End Points (CEPs)

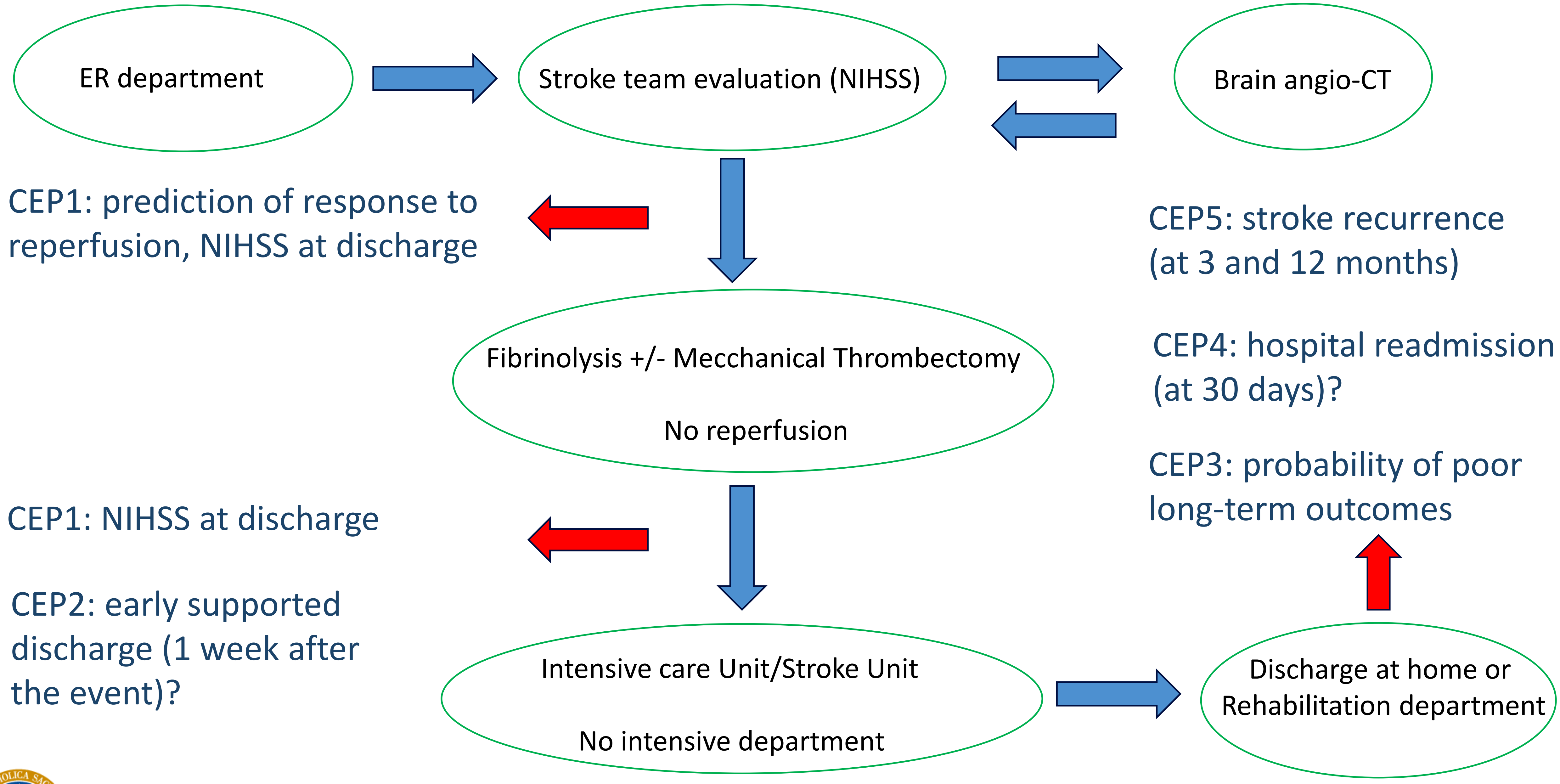
CEP1: What is the individualised prediction of clinical response to acute reperfusion treatment (fibrinolysis / mechanical thrombectomy) and stroke severity at discharge as measured with NIHSS score?

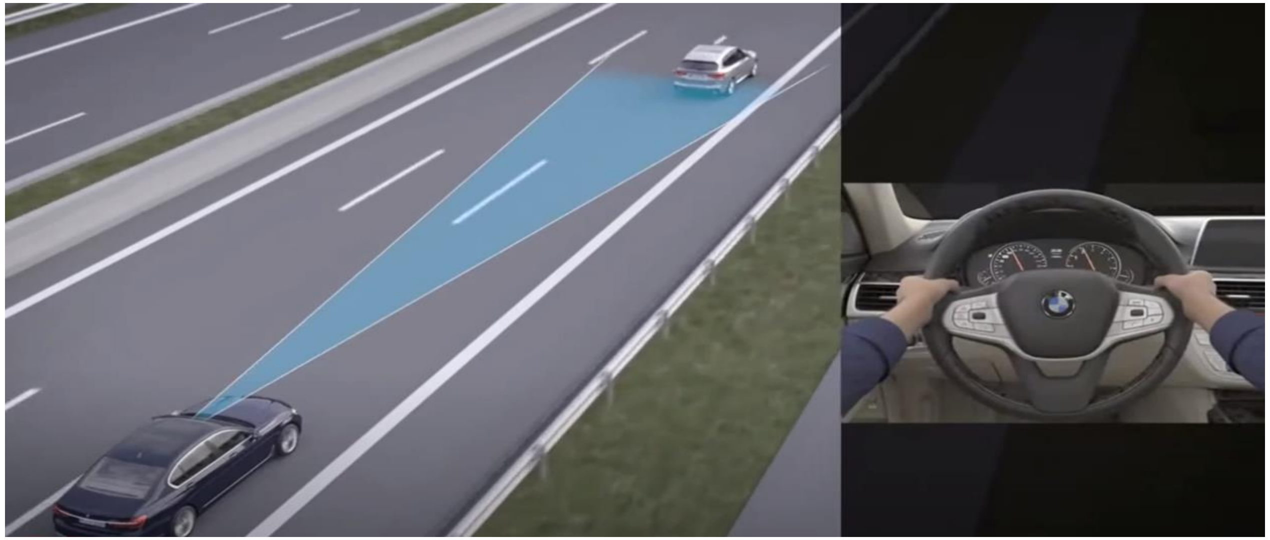
CEP2: What is the probability of early supported discharge (1 week after the event)?

CEP3: What is the probability of poor mobility, incomplete recovery and unfavourable long-term outcomes (in terms of the mRankin Scale, PROMs, PREMs)? When will the patient be able to go back to work?

CEP4: What is the probability of unplanned hospital readmission (at 30 days)?

CEP5: What is the personalised risk of stroke recurrence (at 3 and 12 months? What are the evidence-based actions the patient can do to reduce her/his risk of stroke recurrence?


ER department
Stroke team evaluation (NIHSS)
Brain angio-CT
CEP1: prediction of response to reperfusion, NIHSS at discharge
CEP5: stroke recurrence (at 3 and 12 months)
Fibrinolysis +/- Mecchanical Thrombectomy
No reperfusion
CEP4: hospital readmission (at 30 days)?
CEP3: probability of poor long-term outcomes
CEP1: NIHSS at discharge
CEP2: early supported discharge (1 week after the event)?
Intensive care Unit/Stroke Unit
No intensive department
Discharge at home or Rehabilitation department



https://g.co/kgs/5EMnMF

Thank you

UZ LEUVEN
S·A·F·E
THE STROKE PATIENT VOICE IN EUROPE
eatris
European infrastructure for translational medicine
Vall d'Hebron
Institut de Recerca
Vall d'Hebron
Nora
eurecat
Centre Tecnològic de Catalunya
NACAR
Strategic Design Agency
Gemelli
CERN
Jožef Stefan Institute
univerzitetni klinični center ljubljana
Consiglio Nazionale delle Ricerche
POLITECNICO
MILANO 1863