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
Table of Contents

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)
# Table of Contents

**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)
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, guarantee 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

**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

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

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
South Africa – Thailand - Tunisia – United Arab Emirates – Vietnam

P. Caliandro and L. Serio | TRUSTroke seminar
29/09/23
Four pillars underpin CERN’s mission

- RESEARCH
- COLLABORATION
- TECHNOLOGY & INNOVATION
- EDUCATION & TRAINING
CERN develops technologies in three key areas

ACCELERATORS

DETECTORS

COMPUTING
CERN’s technological innovations have applications in many fields

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

Running jobs: 365644
Active CPU cores: 807139
Transfer rate: 21.54 GiB/sec

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”
Table of Contents

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
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

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

**Augmented Evolutionary and APRIORI Algorithm for the Identification of Functional Dependencies**


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

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

GNNs are versatile and easily scalable

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

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)
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

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

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

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

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)

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

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

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

- 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

P. Caliandro and L. Serio | TRUSTroke seminar
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

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

Estimated number of radiologists per million inhabitants

Artificial intelligence-driven workflow for imaging in patients with cancer

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.

Ref. Lancet Oncology Commission on Medical Imaging and Nuclear Medicine
Perspectives and future work

AI to support healthcare in areas where 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


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 resources.
Perspectives and future work

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

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)
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
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
Leader: CERN
Participants: CNR and POLIMI

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

TRUSTstroke Project’s pathways towards the impact

Year 1
- Retrospective data for 6000 stroke patients
- Locally trained AI models
- FL Platform (R1)
- Trustworthy TRUSTroke platform

Month 18
- Common Data Model & Data management toolbox for Data Harmonisation

Year 2
- FL Platform (R2)
- Global AI models
- Clinical Proof of concept study and adoption by 4 European Hospitals

Month 42
- Trustworthy TRUSTroke platform

6 years
- +40 early adopter hospitals
- +30000 patients assisted
- 210 M€ cost savings
- -35% length hospital stay
- +15% home hospitalization
- -20% 30d hospital readmissions
- +36% treatment adherence
- -22% poor outcomes
- 7k cost savings/patient*

Start

End
4 years

D&C&E

O1 CITIZENS BETTER MANAGE THEIR OWN HEALTH
- Proof of Concept platform for stroke management validated by +500 outpatients and their caregivers

O2 HEALTHCARE PROFESSIONALS BETTER ASSESS RISK & PROGRESSION
- POC platform deployed and validated in 4 stroke units in hospitals at Spain, Italy, Belgium and Slovenia

O3 GUIDELINES FOR USE OF AI-TOOLS TO OPTIMISE HEALTHCARE
- Revised international clinical guidelines for the management of patients using the novel risk scores to determine interventions

O4 QUANTITATIVE INDICATORS FOR BETTER ASSESSMENT OF RISK PROGRESSION
- Robust AI-based scoring systems for 5 crucial CEPs in the management of stroke, with associated interventions

LONGER-TERM IMPACTS
- TRUSTroke reduces 10-20% risk of stroke recurrence, from patient’s better adherence to treatment and healthier habits
- TRUSTroke reduces 5-10% mortality by better management of acute and chronic phases, enabled by realizable risk scores
- Novel quantitative metrics in clinical practise, after regulatory approval and exploitation of TRUSTroke platform

* Lower estimations based on preliminary data gathered at VHIR, where no AI capabilities was included

P. Caliandro and L. Serio | TRUSTroke seminar
Table of Contents

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)
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)
Acute Ischemic stroke

Emergency Department

Clinical Evaluation
Support of vital functions
Blood exams

Brain CT /angio-CT
Perfusion CT

Ischemia

hemorrhage

Fibrinolysis
+/-
Mechanical Thrombectomy

TICI score

---

Table 1  The proposed modified TICI score

<table>
<thead>
<tr>
<th>Score</th>
<th>Revised Thrombolysis in Cerebral Infarction Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No perfusion or anterograde flow beyond site of occlusion</td>
</tr>
<tr>
<td>1</td>
<td>Penetration but not perfusion. Contrast penetration exists past the initial obstruction but with minimal filling of the normal territory</td>
</tr>
<tr>
<td>2</td>
<td>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</td>
</tr>
<tr>
<td>2A</td>
<td>Some perfusion with distal branch filling of &lt;50% of territory visualized</td>
</tr>
<tr>
<td>2B</td>
<td>Substantial perfusion with distal branch filling of ≥ 50% of territory visualized</td>
</tr>
<tr>
<td>2C</td>
<td>Near complete perfusion except for slow flow in a few distal cortical vessels, or presence of small distal cortical emboli</td>
</tr>
<tr>
<td>3</td>
<td>Complete perfusion with normal filling of all distal branches</td>
</tr>
</tbody>
</table>
Acute Ischemic Stroke

Emergency Department
- Support of vital functions
- Blood exams
- Stroke team evaluation

CT cranio/angio-CT
Perfusion CT
- Hemorrhage
- Ischemia

Fibrinolysis
+/-
Mechanical Thrombectomy

TICI score

Stroke Unit

Multiparametric Monitoring

Therapies

Prevent neurological and not neurological complications

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

P. Caliandro and L. Serio | TRUSToke seminar
ER department → Stroke team evaluation (NIHSS) → Brain angio-CT

Fibrinolysis +/- Mechanical Thrombectomy

No reperfusion

Intensive care Unit/Stroke Unit

No intensive department

Discharge at home or Rehabilitation department
Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke

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, IIb or IIIa if no benefit)
- Suggested phrases for writing recommendations:
  - Not recommended
  - 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**
- Moderate-quality evidence from 1 or more RCTs
- Meta-analyses of moderate-quality RCTs

**LEVEL B-NR**
- 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

*OR and LOE are determined independently (any ORR 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 very clear clinical consensuses 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 (CLASS 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.

ORR 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
Le innovazioni digitali all’interno dell’area oncologica:
AI & REAL WORLD DATA

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

Courtesy of A. Dekker – MAASTRO Clinic
Clinical routine

100% of patients

100% of features

80% missing

2000 data points
Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging

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

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

<table>
<thead>
<tr>
<th>Features Used in XGB and GBM Models</th>
<th>Time Point</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>At Admission</td>
</tr>
<tr>
<td>All features</td>
<td>Sex</td>
</tr>
<tr>
<td></td>
<td>Age</td>
</tr>
<tr>
<td>Total ASPECTS</td>
<td>Total ASPECTS</td>
</tr>
<tr>
<td>Specific ASPECTS^</td>
<td>Specific ASPECTS^</td>
</tr>
<tr>
<td>Collaterals status</td>
<td>Collaterals status</td>
</tr>
<tr>
<td>HMCAS present</td>
<td>HMCAS present</td>
</tr>
<tr>
<td>Infarct volume on perfusion CT</td>
<td>Infarct volume on perfusion CT</td>
</tr>
<tr>
<td>Penumbra volume on perfusion CT</td>
<td>Penumbra volume on perfusion CT</td>
</tr>
<tr>
<td>Side of stroke</td>
<td>Side of stroke</td>
</tr>
<tr>
<td>Occluded vessels</td>
<td>Occluded vessels</td>
</tr>
<tr>
<td>Baseline TIMI score</td>
<td>Baseline TIMI score</td>
</tr>
<tr>
<td>NASCET degree of stenosis on left side</td>
<td>NASCET degree of stenosis on left side</td>
</tr>
<tr>
<td>NASCET degree of stenosis on right side</td>
<td>NASCET degree of stenosis on right side</td>
</tr>
<tr>
<td>NIHSS score at baseline</td>
<td>NIHSS score at baseline</td>
</tr>
<tr>
<td></td>
<td>NIHSS score at 24 hours</td>
</tr>
</tbody>
</table>

Selected features^                   | NIHSS score at baseline | NIHSS score at 24 hours |
|                                      | HMCAS present         | ASPECTS of M5          |
|                                      | NASCET degree of stenosis on left side |                  |
|                                      | Age                  |                           |
|                                      | ASPECTS of M3 branch of MCA |            |
|                                      | ASPECTS of caudate    | HMCAS present         |
|                                      | Baseline TIMI score   |                              |

Fig. 1—Feature extraction based on imaging markers from unenhanced CT, perfusion CT, and CT angiography (CTA) of 56-year-old man. A. Unenhanced CT image, right hemisphere middle cerebral artery sign (arrows) can be clearly defined and Alberta stroke program early CT score (ASPECTS) was 8. B. Unenhanced CT image, ASPECTS is 6. Arrow shows right hemisphere middle cerebral artery sign. Outlined area shows manually depicted area of infarction by uninvolved 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. CTA images, right M1 and M2 are occluded, with thrombolyis to intracranial infarct score 0 (B) and collateral score of 1 (E).
Prediction of mRS score at 90 days

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

<table>
<thead>
<tr>
<th>Features</th>
<th>AUC</th>
<th>mRS Score &gt; 0</th>
<th>mRS Score &gt; 1</th>
<th>mRS Score &gt; 2</th>
<th>mRS Score &gt; 3</th>
<th>mRS Score &gt; 4</th>
<th>mRS Score &gt; 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 Features</td>
<td></td>
<td>0.831</td>
<td>0.812</td>
<td>0.873</td>
<td>0.840</td>
<td>0.840</td>
<td>0.867</td>
</tr>
<tr>
<td>XGB</td>
<td></td>
<td>0.798</td>
<td>0.787</td>
<td>0.840</td>
<td>0.825</td>
<td>0.781</td>
<td>0.831</td>
</tr>
<tr>
<td>GBM</td>
<td></td>
<td>0.778</td>
<td>0.713</td>
<td>0.746</td>
<td>0.770</td>
<td>0.759</td>
<td>0.774</td>
</tr>
<tr>
<td>23 Features (without 24-hour NIHSS score)</td>
<td></td>
<td>0.757</td>
<td>0.681</td>
<td>0.748</td>
<td>0.762</td>
<td>0.735</td>
<td>0.758</td>
</tr>
</tbody>
</table>

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

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.

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

Selected features:
- Discharge NIHSS assessment items
- discharge Barthel index
- the 30-day mRS degree
Prediction of mRS score at 90 days

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.

<table>
<thead>
<tr>
<th>Stroke Type</th>
<th>Ischemic stroke (Mean ± SD%)</th>
<th>Hemorrhagic stroke (Mean ± SD%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Sets</td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>Clinical set</td>
<td>n = 203</td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>86.5 ± 0.2</td>
<td>89.2 ± 0.3</td>
</tr>
<tr>
<td>RF</td>
<td>86.6 ± 0.2</td>
<td>89.8 ± 0.2</td>
</tr>
<tr>
<td>HANN</td>
<td>86.2 ± 0.2</td>
<td>89.4 ± 0.2</td>
</tr>
<tr>
<td>SVM</td>
<td>86.7 ± 0.2</td>
<td>89.4 ± 0.2</td>
</tr>
<tr>
<td>Whole set</td>
<td>n = 206</td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>89.2 ± 0.6</td>
<td>91.9 ± 0.2</td>
</tr>
<tr>
<td>RF</td>
<td>94.1 ± 0.1</td>
<td>91.9 ± 0.1</td>
</tr>
<tr>
<td>HANN</td>
<td>91.5 ± 0.5</td>
<td>91.9 ± 0.3</td>
</tr>
<tr>
<td>SVM</td>
<td>91.9 ± 0.2</td>
<td>94.2 ± 0.1</td>
</tr>
<tr>
<td>Selected set</td>
<td>n = 17/22</td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>90.0 ± 0.7</td>
<td>92.8 ± 0.4</td>
</tr>
<tr>
<td>RF</td>
<td>91.3 ± 0.3</td>
<td>93.8 ± 0.3</td>
</tr>
<tr>
<td>HANN</td>
<td>91.8 ± 0.3</td>
<td>94.1 ± 0.1</td>
</tr>
<tr>
<td>SVM</td>
<td>91.9 ± 0.2</td>
<td>94.2 ± 0.1</td>
</tr>
</tbody>
</table>


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)
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.

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

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.

### Prediction of mRS score at 90 days

<table>
<thead>
<tr>
<th>Patient demographics</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>26</td>
</tr>
<tr>
<td>Sex</td>
<td>27</td>
</tr>
<tr>
<td>Chronic heart failure</td>
<td>28</td>
</tr>
<tr>
<td>Smoking status</td>
<td>29</td>
</tr>
</tbody>
</table>

### Clinical variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time from onset to admission</td>
<td>30</td>
</tr>
<tr>
<td>Medication history</td>
<td>31</td>
</tr>
</tbody>
</table>

### Previous diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin (mg/dL)</td>
<td>34</td>
</tr>
<tr>
<td>White blood cell count (10⁹/L)</td>
<td>35</td>
</tr>
</tbody>
</table>

### Laboratory values

<table>
<thead>
<tr>
<th>Value</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prothrombin time (INR)</td>
<td>37</td>
</tr>
<tr>
<td>Glucose level (mg/dL)</td>
<td>38</td>
</tr>
</tbody>
</table>

**Table 1. List of collected variables.**

**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)
TRUSTWORTHY AI FOR IMPROVEMENT OF STROKE OUTCOMES

TRUSTroke

*Based on existing research and publications by the consortium
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?
CEP1: prediction of response to reperfusion, NIHSS at discharge

CEP1: NIHSS at discharge

CEP2: early supported discharge (1 week after the event)?

CEP3: probability of poor long-term outcomes

CEP4: hospital readmission (at 30 days)?

CEP5: stroke recurrence (at 3 and 12 months)
Thank you