PRECISE4Q

PREDICTIVE MODELLING IN STROKE

DELIVERABLE - SUBMISSION

Project Acronym: **Precise4Q**
Grant Agreement number: **777107**
Project Title: **Personalised Medicine by Predictive Modelling in Stroke for better Quality of Life**

D7.5 Launch of European Modelling Platform for Open Stroke Research

Revision 3.0

Authors and Contributors	Dietmar Frey (CUB); Adam Hilbert (CUB); Rainer Thiel (EMP); Malte von Tottleben (EMP)			
Responsible Author	Dietmar Frey		**Email**	dietmar.frey@charite.de
	Beneficiary	CUB	**Phone**	+49 30 450 560398

Revision History, Status, Abstract, Keywords, Statement of Originality

Revision History

Revision	Date	Author	Organisation	Description
0.9	15/03/22	Dietmar Frey	CUB	Initial draft
1.0	28/03/22	Adam Hilbert	CUB	Revision
1.1	03/05/22	Malte von Tottleben	EMP	Revision and editing
1.2	09/06/22	Rainer Thiel	EMP	Revision

Project co-funded by the European Commission within **H2020-SC1-2016-2017/SC1-PM-17-2017**		
Dissemination Level		
PU	Public, fully open	x
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

1.3	05/11/22	Adam Hilbert	CUB	Editing and review
1.4	21/12/22	Malte von Tottleben	EMP	Revision
2.0	30/12/22	Dietmar Frey	CUB	Final review
2.1	25/01/23	Dietmar Frey	CUB	Addressing reviewers' comments
2.2	30/01/23	Adam Hilbert	CUB	Addressing reviewers' comments
2.3	10/02/23	Dietmar Frey	CUB	Adding and revisions
2.4	19/02/23	Adam Hilbert	CUB	Minor modifications and additions
2.5	27/02/23	Malte von Tottleben	EMP	Revision
2.6	28/02/23	Dietmar Frey	CUB	Revision
2.7	01/03/23	Rainer Thiel	EMP	Final Revision
3.0	03/03/23	Dietmar Frey	CUB	Final Version

Date of delivery	Contractual:	30.06.2022	Actual:	0603.2023
Status	final x /draft ☐			

Abstract (for dissemination)	In this deliverable, we exploited the PRECISE4Q data eco-system, integration and application scenarios and used them as a blueprint for EUROPE-Stroke. The aim was to support hospitals, physicians and researchers in stroke modelling and data-guided treatment and rehabilitation planning through various measures. One of the main outputs are successful follow-up projects for the different use cases and phases of the disease which could be initiated both on a European level and national funding levels to ensure sustainability of the results in improving patient care and outcome.
Keywords	Open Science, Platform, Stroke, Research, Follow-up projects, Sustainability

Statement of originality
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of Content

List of figures

Executive Summary

Within Precise4Q multiple AI-based models will be developed to be integrated in solutions guiding prevention, treatment, rehab and reintegration measures. The main medical concept of PRECISE4Q is to target four different stages of stroke in the life trajectory in a novel precision medicine approach. Precision medicine is defined as a concept to tailor prevention, diagnostics and therapeutics individually to any given patient. Thus, we are developing a set of models for each of the four clinical stages of stroke - prevention, stroke therapy, stroke rehabilitation and stroke reintegration - and combine these in a digital stroke patient platform. To provide the digital stroke patient platform and its respective AI-based models to the respective healthcare professionals in the different phases, we have been designing use case scenarios and are subsequently developing user-centered services such as mobile applications (apps) and decision support systems.

This task consisted of planning and designing an Open Innovation Organisation for stroke modelling in prevention and treatment (European Modelling Platform for Open Stroke Research; EUROPE-Stroke). The aim was to fully exploit the PRECISE4Q data eco-system, integration and application scenarios and use them as a blueprint for EUROPE-Stroke. This activity consisted of the organisation, coordination and implementation of the dissemination actions addressed to make available to the target audiences information on, and arising from, the project. We defined the kind of activities to reach the audiences more efficiently, and provide the tools for doing so.

We were successful in ensuring the PRECISE4Q aims and output beyond the project's end by acquiring subsequent funding for the use case scenarios: For prevention and continuous assessment STRATIF-AI (#101080875) with the partners LIN, TUD, UM, CUB, GUT, for Acute Stroke Treatment VALIDATE (#101057263) with the partners CUB, EMP, TUD and for Security of Data in Modeling and AI-based Patient Monitoring CYLCOMED (#101095542).

Outreach to stakeholders will be sustainably ensured by integrating SAFE (Stroke Alliance for Europe https://www.safestroke.eu/) as a patient outreach organization and multiple ongoing research based on PRECISE4Q and documented in 38 peer-reviewed journal articles and 11 papers pending acceptance and 56 conference contributions and 30 additional related papers (see D7.2). Dissemination activities were focused on articles in peer-reviewed journals and conference papers with the addition of workshop publications.

1 Introduction

Within Precise4Q multiple AI-based models will be developed to be integrated in solutions guiding prevention, treatment, rehab and reintegration measures.

The main medical concept of PRECISE4Q is to target four different stages of stroke in the life trajectory in a novel precision medicine approach. Precision medicine is defined as a concept to tailor prevention, diagnostics and therapeutics individually to any given patient. Thus, we have been developing a set of models for each of the four clinical stages of stroke - prevention, stroke therapy, stroke rehabilitation and stroke reintegration - and combine these in a digital stroke patient platform. To provide the digital stroke patient platform and its respective AI-based models to the healthcare professional, we have been designing use case scenarios and are subsequently developing user-centered services such as mobile applications (apps) and decision support systems. The objective of this report is take the results from identification of the most relevant user needs for each patient journey phases, concentrating on the most common and complex questions (D1.2 and D1.3) and translate these into a service blueprint for the different stages enabling the development of user-centric services and product prototypes.

2 Service blueprint as the foundation for the Stroke Modeling Platform

To have the maximum impact on improvement of patient care and patient outcome, it is essential to provide the developed AI-based models in user-friendly and seamlessly integrated services.

We designed and developed the *Stroke Service Blueprint* in an iterative process regarding the four phases of the patient journey:

1. Prevention: One of the most promising approaches to reduce the effects of stroke on individual health and healthcare systems is to prevent stroke. More than 77% of stroke events are first time events.

2. Acute Treatment: There have been advances in the therapy of ischemic stroke in the past decades. Overall therapy success, however, is still poor. For thromboembolic stroke, the most favorable current treatment paradigm is the time-based dissolution of the obstructing blood clot by a drug or its mechanical retrieval. Unfortunately, up to 20% of patients arrive with an unknown time from stroke onset, and most patients present too late in the hospital to receive treatment.

3. Rehabilitation: A multitude of different stroke rehabilitation concepts and methods has been developed to date. However, from an evidence-based perspective only very few general proven recommendations exist: a) Specialized rehabilitation is useful, b) early rehabilitation and mobilization is useful and c) higher intensities of therapy are useful. Beyond this, it is unclear which therapy

options lead to better rehabilitation outcome, i.e. which therapies are best suited for the individual patient.

4. Reintegration: Reintegration is the long-term outcome after stroke. After acute treatment and rehabilitation, reintegration success is measured by the patients' reintegration into their family, communities and workplaces.

As shown in Figure 1, the service blueprint is divided into the four different stages of the patient journey. In the beginning, in the prevention phase, the patient's actions are visualized: The patient is notified about necessary check-ups, the patient visits the doctor, and checks for his/her current risk. This might lead to behavior change and reduction of stroke risk. The touch points are: information by letter, email or insurance app; the medical check-up, the mobile technology in an app and a medical check-up.

For this deliverable, these different stages and the various touchpoints reflect on the distinct models that have been built and that will act independently from each other for the defined use case.

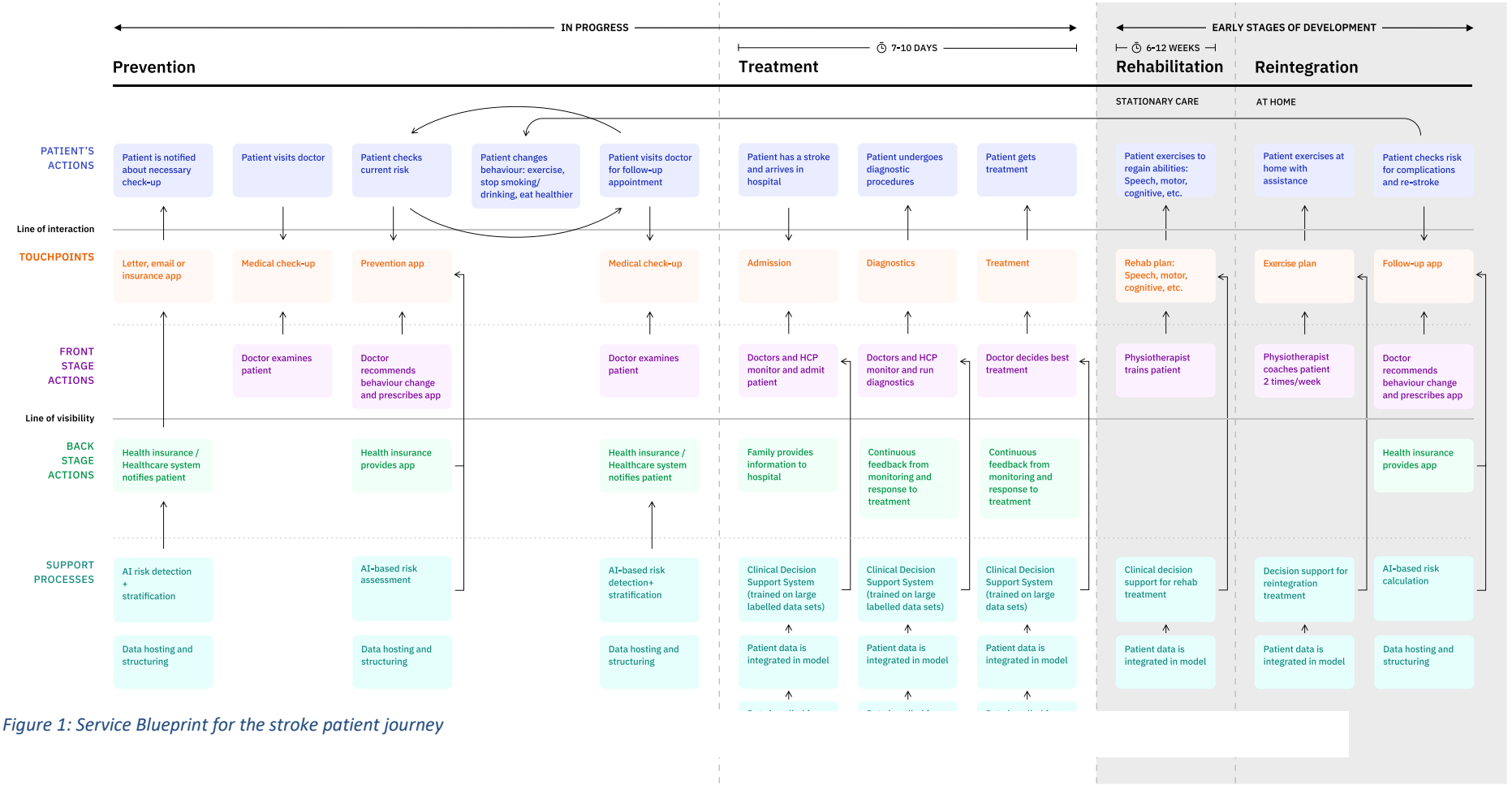

Figure 1: Service Blueprint for the stroke patient journey

Figure 1: The previously developed PRECISE4Q service blueprint serves as a basis for the stroke modeling platform.

3 Application scenarios

As described and based on the work in the project, the application scenarios and the respective technology vehicles are dependent on the different stages of the patient journey and the approaches to be taken: In 1. Prevention: a mobile application (app) for the primary use of the patient at risk for stroke. In 2. Treatment: A clinical decision support system integrated in the Hospital Information System (HIS) supporting the doctors treating acute stroke patients. In 3. Rehabilitation: A decision support system helping healthcare professionals in allocating resources and personalizing rehabilitation procedures. And, finally to close the loop, in 4. Reintegration: a mobile application for the use of patients to prevent complications after stroke and re-stroke (i.e. second stroke).

This is reflected in the technological environment (termed “backstage” in the blueprint). For successful modeling and integration of models various processes and involved participants are grouped visually by placing them in lanes, with one lane for each person, group or relevant sub-process.

3.1 Prevention

Stroke is the end-result of a long chain of etiological events, and each step in this chain can, in principle, be prevented by preventive healthcare. The initial steps towards ischemic stroke, which is the primary version when considering prevention, are metabolic disorders, such as overweight, insulin resistance and dyslipidemia. All of these conditions are important risk factors for ischemic stroke, and all of these conditions can largely be prevented by proper diet and exercise. These conditions increase the risk of development of type 2 diabetes and atherosclerosis. These complications are also important risk factors for stroke, and can be treated and improved by various medications, which often are combined with improved diet and exercise levels. Other early risk factors are high blood pressure, which to some extent is preventable by diet and exercise, but which also is treated with a variety of different medications. Finally, high blood pressure and atherosclerosis increases the risk of ruptures, which together with atrial fibrillation increases the risk of thrombosis - the formation of blood clots - which may travel up to the brain, and cause a stroke. Also, these last steps can be treated by e.g. anticoagulants or electrical cardioversion, which also are important preventive actions, which reduces the risk of stroke. While it is clear that there are a number of different preventive actions that can be made, traditional healthcare is seldom spending enough resources to fully exploit this possibility, and there is a need both to i) identify which preventive actions are the most appropriate for each patient, and ii) increase the understanding and motivation of each patient, to comply and follow those prescribed preventive actions. These needs warrant the development of new technologies, which can provide cost- competitive preventive healthcare, and both help identify treatments and make patients more likely to understand and follow these preventive treatments.

3.2 Acute Stroke

The treatment decision of acute stroke patients is difficult since causes are highly heterogeneous. Current treatment paradigms, however, do not consider individual differences. Stroke care, selecting the optimal treatment option in the acute setting, could be significantly improved by more personalized risk calculation and individualized therapeutic recommendations. Such a precision medicine approach for the treatment of stroke would lead to an informed choice of medical and surgical treatment interventions and a more effective allocation of resources as well as better outcomes in stroke patients.
There have been advances in the treatment of ischemic stroke in the past decades. Overall treatment success, however, is still poor. For thromboembolic stroke, the most favorable current treatment paradigm is the time-based dissolution of the obstructing blood clot by a drug or its mechanical removal. Mainly, there are 2 treatment options: 1. Systemic thrombolysis, i.e. the i.v. administration of a blood-thinning medication or 2. Interventional mechanical thrombectomy, i.e. the neuro-radiologist retrieves the blood clot via a catheter and clears the way.
Both treatment options have risks and typically after 6 hours the risks outweigh the benefits leading to the time-to-treat paradigm that in average is correct but often leads to under-treatment of patients that individually would benefit.
In addition, unfortunately, 20% of patients arrive with an unknown time from stroke onset, and most of these patients present too late in the hospital to receive treatment. Also, only very few patients are eligible for mechanical thrombectomy, whereas the number needed to treat for intravenous thrombolysis drastically increases with time and reaches around 10 in the time window where most patients are treated.
Latest results have shown that the so called DWI-FLAIR-mismatch, a mismatch of stroke related signals in two different MRI-sequences, can identify patients eligible for treatment independent of their onset time. Approaches like these are called “tissue-based” approaches which have a much higher potential for patient selection than the purely time-based approaches of the past. However, they are not widespread and the DWI-FLAIR mismatch – as the only validated one – relies on MR-imaging which is far less often used than CT-imaging in the acute setting. The advent of mechanical thrombectomy and its continuous success has considerably changed the treatment schemas for acute stroke patients making treatment more complex and calling for solutions that encompass the individual features of the patient.
Neurologists are desperate to have more individual fine-grained approaches to treatment. This applies both to the imaging standards (CT scans and MRI) and to the clinical data analytics posing an opportunity that the CDSS will answer. An essential contribution of our prototype will be the adaptation to the current clinical changes (e.g. introduction of mechanical thrombectomy) and to break up the current time-fixated approach to treatment. Also, imaging-based prediction of stroke progression is added here as additional use-case that might give better visual guidance regarding treatment success in contrast to the simple numbers the other models will provide. Thus current processes and guideline-

based treatment are by their nature general and do not account for all the individual characteristics of every patient. That often leads to treatment decisions that deny a patient a potentially life-saving intervention. The deployment and administration of treatment after stroke is highly time-critical and often leads to insufficient outcome. Taken together, the resulting wrong treatment strategies lead to unnecessary disability, death and complications resulting in enormous costs for the healthcare system and society in general. For the scenario of acute stroke the obvious clinical need is a solution that helps the doctor make the right decision based on a solution that is personalized, objective, evidence-based, and easy to use.

For this described scenario, we lay the grounds for providing the following main elements: 1. We support the decision of the doctor for selecting the best individual treatment according to the associated risks. By simulating different scenarios such as surgery, stenting, medication, observation, the treating doctor can compare proactively the different options and can select the treatment with the lowest associated risks. This, of course, can be done before treatment is initiated and will result in the treatment that is best on an individualized level. 2. Doctors can treat according to best individual blood pressure values. Blood pressure navigation is an essential element after acute stroke. A safe blood pressure range can be determined and patients will neither suffer from too high blood pressures resulting possibly in devastating bleeding in the brain nor from low blood pressures that would lead to insufficient blood supply to the brain. 3. Digital monitoring and reporting will enable quality assessments and quantitative follow up on individual outcome of patients feeding back to the models.

3.3 Rehabiliation

Post-stroke cognitive impairment occurs frequently, it ranges from 20% to 80%, varying due to the difference between countries, races, and the diagnostic criteria. Computerized tasks are increasingly being applied over traditional paper and pencil activities. The Guttmann, NeuroPersonalTrainer® platform (https://www.gnpt.es/en) GNPT, which is already integrated into the clinical practice of more than 200 clinical centres in Spain, is the cognitive rehabilitation demonstrator presented in this document for cognitive rehabilitation pilot at Institut Guttmann.

Patients execute computerized cognitive rehabilitation tasks selected by clinicians along a treatment, each task targets a specific cognitive function e.g. Attention, Memory or Executive Function. Each cognitive function has been traditionally subdivided into sub-functions (e.g. Sustained, Selective and Divided Attention). The whole cognitive rehabilitation process in GNPT involves 4 main steps presented in this section.

The neuropsychological scenario is defined as an active process that helps the affected person to optimize the recovery of superior functions, to better understand the nature of

the alterations it presents and to develop strategies to compensate for disorders. The evaluation is the first step for neuropsychological rehabilitation scenario allows to:

- Identify, describe and quantify cognitive, behavioural and emotional alterations as well as the preserved functions.
- Guiding the process in order to rehabilitate the affected functions and modify the maladaptive behaviours.
- Determine the patient's progress more objectively and evaluate the effectiveness of the different interventions
- Provide information and guidance to the family and the members of the rehabilitation team that help to set realistic and functional goals
- Estimate the severity of sequels within the forensic professionals in order to support legal decision-making.
- Contribute, along with other professionals, to the psychosocial orientation that allows the reintegration of the patient to his habitual environment

A standard cognitive rehabilitation treatment takes 2-5 months distributed in 3-5 sessions a week, each session is composed of 5-10 cognitive rehabilitation tasks. Typically, each patient executes a different number of tasks along treatment and in a different order. For each execution, the patient gets a result (ranging from 0 to 100). There is very little research related on the amount and type of practice that occurs during cognitive rehabilitation treatment and its relationship to rehabilitation and not enough on-field experience yet regarding which specific intervention (tasks and performance on them) is more appropriate to help cognitive rehabilitation therapists to design their clinical therapeutic plans. Therefore, we described several use cases addressing representative situations.

3.4 Reintegration

While the majority of stroke survivors return to live in the community, re-integration may be an enormous challenge. The ability to return to an acceptable lifestyle, participating in both social and domestic activities is important for perceived quality of life. Therefore, in this work the Quality of Life Laboratory (QVidLab) is the reintegration demonstrator described in this document for the community reintegration phase. Every stroke patient, after discharge from Institut Guttmann, periodically undergoes follow up evaluations which may also lead to detect early pathology that, due to the characteristics of the specific lesion, could be asymptomatic and/or remain unnoticed until advanced stages. It has a periodicity of 12-24 months, patients can request it by telephone, or in person to the Admissions Service, which, approximately one month before the evaluation, sends a reminder letter of the visit to the patient by mail. The periodic review will be done within the least amount of time as possible, with the objective of interfering as little as possible in the usual activities of the person (it usually takes from 9:00 to 12:30 during one morning). There is also the possibility, for patients from other Autonomous Communities of Spain or abroad, to

perform this procedure within a short admission to the hospital of less than 5 days. Subsequently, within approximately three weeks, the patient receives at home the report with the conclusions of the medical examinations. If problems have been detected that require urgent intervention, patients will be personally contacted or the responsible relative, to give the pertinent information as well as to request additional tests or refer to the adequate service for the follow-up and / or treatment of the eventual complications detected.

In this scenario, periodic evaluations are preventive actions, which aim to reduce the incidence of complications in the population with stroke while allowing the monitoring of the results of long-term treatment, by different clinical professionals, in terms of restriction of participation, as well as the assessment of the family, community and employment insertion.

Periodic integral evaluations are performed within the framework of an interdisciplinary team which will comprehensively assess the patient, the team is integrated by: Medical Doctor (Coordinator of the whole process), Nurse, Urologist, Psychologist, Neuropsychologist, social worker.

4 Set of Models

Within PRECISE4Q models for the different use cases and designing frameworks and user interfaces were developed in parallel and partly independent from each other. For the implementation of these models in the real world, we adhered to iterative development principles integrating user feedback and different deliverables over the course of the project as D1.3 Use cases and their inputs/outputs specifications, D1.4 Set of functional requirements and architecture, D1.5 Empirical study on attitudes towards personalized medicine, D1.6 Ethical framework and oversight mechanisms for big data health research, D1.7 Ethics of personalized medicine and data-driven modeling, D1.8 Release of the deliberative dashboard, D4.3 Data Schema Designs for Each Model, D6.2 Innovation management plan and exploitation report; and will inform D2.7 Reliable interfaces implemented as service connectors, D2.8 Pilot for clinical decision support system, D4.10 An integrated digital stroke patient platform spanning the entire patient life-cycle.

4.1 Prevention

Two different kinds of models were developed and combined in the hybrid model for stroke prevention; a multi-level multi-timescale mechanistic models, and logistic regression models. The mechanistic models simulate the progression of risk factors for stroke given different scenarios, and these long-term simulations are used as input to the logistic regression models. The logistic regression models calculate the evolution of the risk of stroke given the scenarios simulated by the mechanistic models.

The mechanistic model is a combination of several models developed by LIU; a blood pressure model, and a multi-level diabetes progression model. The blood pressure model was developed using two datasets; Framingham data, to capture the increase in blood pressure with age, and data from the VALUE trial, to capture how blood pressure change with medication, specifically angiotensin receptor blocker. The multi-level diabetes progression model is a combination of several diabetes related models developed and used by LIU over many years; a fat cell model, a meal response model, and a weight change model. These models could previously only simulate either the diabetes case or the non-diabetic case, and not the progression of insulin resistance and ultimately diabetes over several years. To include this progression in the combined model, an insulin resistance model was added. The insulin resistance model takes change in fat tissue from the weight model as input, and the resulting insulin resistance has an effect on the meal response model and fat cell model. The entire mechanistic model can then take different scenarios, such as different diets and/or a blood pressure medication, and simulate the evolution of risk factors for stroke given these scenarios. Using these simulations as input to the logistic regression models for risk predictions enables the risk to be calculated continuously over the simulated time span, and the simulations can be used to get an explanation and understanding for the predicted risk scores. Furthermore, since the mechanistic model is multi-level and multi-timescale, the model can be used to simulate things happening on

shorter timescales as well, such as glucose levels during a meal, giving further opportunities for a deeper understanding of the physiological mechanisms underlying the risk predictions.

The logistic regression models predict an individualized risk of stroke with a time horizon of 5 years, and were created using data from a number of longitudinal studies. In creating the models, TU Dublin identified a gap in the literature where the use of age as a risk factor in the model might lead to over or underfitting of certain age groups and that some risk factors might not change proportionality with age. TU Dublin created a framework for determining if stroke risk factors were proportional to age, and then created a set of models that were designed to take the non-proportionality of a number of risk factors by age into account. These age specific models were shown to be better calibrated and produce better predictions compared to models that do not take the non-proportionality of risk factors by age into account. To account for discontinuous jumps in risk scores when moving between age groups, TU Dublin converted the set of age-specific models from a set of independent models to an ensemble, with the predictions from each of the models being integrated to generate a single overall risk for an individual. The models were then combined with the mechanistic models created by LIU to produce a hybrid model that allows for individualized simulations to be run over time, showing the evolution of stroke risk under different treatment conditions such as diet, exercise, and medication.

4.2 Acute Treatment

Hybrid modelling led by CUB included the development of a mechanistic simulation model to estimate the hemodynamic changes for a certain patient in various blood pressure levels. This can reveal vulnerability to hypoperfusion in certain areas of the brain and thus provide clinicians with additional, valuable information about the potential progression of the stroke. The pipeline was first developed with semi-automated steps, together with a software tool to facilitate ground truth creation of the vessel masks and anatomical labelling as well for further automation. The novelty of the approach is that the simulation relies merely on routine TOF-MRA imaging, a widely available imaging method not requiring contrast-based enhancement of the vasculature. The results were that an ensemble model combining a stroke outcome prediction model trained using clinical features with multiple hemodynamic simulation models (each simulation assuming a different blood pressure configuration) achieves slightly higher performance. In addition to pilot studies utilizing quantitative MRI data, efforts led by TU Dublin included convolutional neural networks (CNNs) trained on ADC maps from ischemic stroke patients and healthy stroke-aged controls. Initial predictions of dichotomized chronological age and functional stroke outcomes indicated that the models showed comparably poor results. To ensure that the poor performances were not due to the model architecture, the same architecture was trained on the simpler task of stroke diagnosis, where the results suggested that the selected architecture could capture input-output relationships, suggesting that the bad poor performance of the model on the dichotomized chronological age and function stroke

outcome tasks were not due to model limitations in terms of the model being able to learn a binary task on this dataset. Next, TU Dublin investigated if there were any image or patient-specific patterns in the outcome and age predictions that may be an indication of ``easy'' brain images and found no distinguishable patterns related to data source or gender, and then used class activation mapping to highlight regions in the ADC maps driving the predictions of the age and outcome prediction models. The highlighted regions of the two models vastly overlapped. They were also similar to the age-relevant regions identified in TU Dublin's previous work. To ensure that these regions were not just identical for any poor-performing model independent of the task, the same architecture was trained on a random binary classification task.

Furthermore, to support decision support scenario in its completeness, CUB conducted research on exploitation of Deep Learning for the creation of multimodal models. These models can incorporate all information sources (i.e. clinical parameters, medical history, imaging) available in the acute setting, similarly to the human decision-making processes. Despite the task of multimodal outcome prediction however has proven to be highly challenging, great progress was made in understanding the behavior of these models. We were able to achieve slight improvements in prediction performance compared to the classical clinical models that only consider standard imaging features dependent on manual/semi-automatic extraction. Additionally, CUB has done novel analysis of the interaction of different modalities in this Deep Learning training scenario and found potential reasons for the minor improvements. Findings were presented and greatly welcomed on the World Stroke Congress (WSC) 2022.

Due to data access issues caused by the pandemic, an initial dataset for the development of the quality of life (QoL) models was the Barthel Index (BI) data from GUT. In this deliverable, the methodology and workflow pipeline for predicting a profile of the patient, based on the items of the QoL questionnaire, rather than a single value, was developed and optimized with two different datasets, at the acute stage as well as the reintegration stage. Using BI data at 3 months after stroke, multitask learning (MTL) models were explored and initially constructed using a statistical dependency coefficient Cramer's V. Initial results were unstable due to the small size of the dataset; however, the final BI dataset was assembled into model development, aiding in new methodology of MTL with task affinity grouping (TAG) to be integrated into this work. Both MTL modelling methods were applied to the final BI dataset and to supplementary data extracted from the large clinical trial IST-3. After variables exploration and filtering, two IST-3 data subsets were extracted for optimizing the MTL methodologies, including data of patient with stroke after 6 months as well as 18 months, which encompasses the reintegration stage. The models were run 50 times to enable optimization of hyperparameters, yielding stable results.

4.3 Rehabilitation

The rehabilitation and reintegration taskforce led by UCD completed the stroke rehabilitation modelling which was focused in three clinical use cases: i) the prediction of cognitive deficit reduction (UCD), ii) the prediction of the patient's compliance during therapy (UCD), and iii) the prediction of the motor deficit reduction (GUT). For the first clinical use case, the developed model used as input electronic records from ischemic stroke surviving patients containing demographic information, scores from administered cognitive assessments, and therapy information collected during rehabilitation. A cognitive improvement variable, which integrates scores from different cognitive domains (e.g., attention, memory, language), was selected as target and adapted to fit a binary classification problem. Following a similar approach, the second clinical use case produced a therapy compliance prediction model. The model was trained to predict the proportion of non-executed tasks during therapy. The third clinical use case focused on post-stroke motor recovery using the Fugl-Meyer Assessment Upper Extremity (FMA-UE) as the prediction target. For this use case, the individual prediction capacity of the FMA-UE items (e.g., finger mass flexion) was evaluated. Finally, as introductory research into how stroke rehabilitation data can be leveraged beyond primary tabular models, two use cases that explore the usage of the graph database representation for modelling purposes were presented (UM).

4.4 Reintegration

For the reintegration phase, the rehabilitation and reintegration taskforce led by UCD was focused on the implementation of the reintegration models for the fourth and final stage of the stroke patient journey. Two main models were implemented for this phase: i) the long-term trajectories of community integration (TUD), and ii) the prediction of significant social risk (TUD), followed by the extraction of communities guided by the structure of the harmonised graph database (UM). For the community integration model, growth mixture modelling (GMM) was applied to identify and characterise the long-term trajectories of community integration in adults who suffered stroke episodes. For the social risk prediction model, a gradient boosted modelling (GBM) approach was applied. The model employed functional independence measures (FIM assessment) and social risk scores (EVSF ranking score) to predict social risk upon discharge from the rehabilitation hospital. Finally, a preliminary study applied community detection algorithms over a graph database representation to extract clusters of patients based on their similarities (FIM and EVSF scores). Future research targets using this type of approach to detect relevant patient communities that tabular data cannot identify.

5 Results and sustainability of scientific and technological output

In modification of the original planning we have not established a dedicated genuine platform but have been integrating and will further integrate our findings, models and results in a variety of targeted instruments. This adjustment was necessary after initial research showed that one-size-fits-all outreach would be inefficient and specific audiences require specific approaches.

Outreach to patients will be facilitated and implemented via **SAFE** – the European patient organization that is integral part of one of the Horizon Europe follow-up projects: **VALIDATE** (#101057263: Validation of a Trustworthy AI-based Clinical Decision Support System for Improving Patient Outcome in Acute Stroke Treatment). VALIDATE will serve as a prime example how to achieve continuous agile research within one topic and generate results that can be built upon. Due to the modular design and built the PRECISE4Q acute stroke models will be available for further development, refinement and clinical validation within this follow-up project.

Thanks to the enhanced retrospective study in PRECISE4Q, these models build a solid basis for testing and establishing a robust validation framework, that can bring research findings into actual practice.

In addition, in **STRATIF-AI**, another follow-up project to PRECISE4Q, we will continuously evaluate models that have been developed in PRECISE4Q for the prevention and follow-up of patients with the help of digital twinning. This will serve the purpose of creating a platform for researchers to develop further continuous modeling – in particular for hybrid modeling – and deployment of models.

The scientific outreach and production was significant resulting in 70 peer-reviewed journal articles and 11 papers pending acceptance and 56 conference contributions (posters, workshops, presentations, summer schools, and conference papers), initiating constructive discussions on the subject and enabling collaboration to maximize project output.