PRECISE4Q

PREDICTIVE MODELLING IN STROKE

DELIVERABLE

Project Acronym: **Precise4Q**

Grant Agreement number: **777107**

Project Title: **Personalised Medicine by Predictive Modelling in Stroke for better Quality of Life**

D4.2 – Quality of Life Targets for the Models created in T4.5, T4.6, T4.7, and T4.8

Revision: 1.2

Authors and Contributors	John Kelleher (TU Dublin), Attia Fatima (TU Dublin), Vince Madai (CUB),			
Responsible Author	John Kelleher		**Email**	john.d.kelleher@dit.ie
	Beneficiary	TU Dublin	**Phone**	+353 1 402 4789

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	

Table of contents

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

Revision History

Revision	Date	Author	Organisation	Description
0.1	11.2.19	Attia Fatima	TU Dublin	First draft Prevention Model Section
0.2	12.2.19	John Kelleher	TU Dublin	Reviewed and revised prevention model section
0.3	13.2.19	Attia Fatima	TU Dublin	Added acute model section
0.4	14.2.19	John Kelleher	TU Dublin	Reviewed and revised acute model section
0.5	15.2.19	Attia Fatima	TU Dublin	Added rehabilitation model section
0.6	16.2.19	John Kelleher	TU Dublin	Reviewed and revised rehabilitation model section
0.7	25.2.19	Vince Madai	CUB	comments on stroke prevention and acute stroke treatment
0.8	25.2.19	John Kelleher	TU Dublin	Revised the executive summary
0.9	26.2.19	John Kelleher	TU Dublin	Revised the prevention stroke section
1.0	27.2.19	John Kelleher	TU Dublin	Revised the Acute Stroke section
1.2	28.2.19	John Kelleher	TU Dublin	Updated the Acute Stroke sections, and revised the rehabilitation model section, and conclusions, and revised abstract

Date of delivery	Contractual:	28.2.2019	Actual:	28.2.2019
Status	final **x** /draft ☐			

Abstract (for dissemination)	This document set out the target outputs for the Precise4Q model deliverables D4.5, D4.6, D4. and D4.8. The specification of these model outputs will be used to guide the development of these models.
Keywords	stroke prevention, acute stroke treatment, stroke rehabilitation, hybrid modelling, model targets.

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This deliverable contains original unpublished work except where clearly indicated otherwise.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

Executive Summary

A key goal of the PRECISE4Q project is to create multi-dimensional data-driven predictive computer models enabling personalised stroke treatment, addressing patient needs. PRECISE4Q addresses this challenge by targeting (and developing models for) four major stages in stroke management: stroke prevention, acute stroke treatment, post stroke functional and cognitive rehabilitation, and post stroke social reintegration. Accurate and timely decision making in all these phases is critical to prevent strokes and improve outcomes after stroke.

A fundamental task in developing any predictive model is to define the predictive target of the model. The predictive target of a model is the data structure the model will generate (or predict) in response to a set of inputs. In the context of Precise4Q defining a predictive target entails describing the information a model will generate and present to a clinician in response to a patient profile. This deliverable sets out an initial definition of the predictive targets for the Precise4Q models: D4.5 (prevention), D4.6 (acute 1), D4.7 (acute 2), and D4.8 (rehabilitation). The definition of these targets is informed by the use-case and clinical scenarios definitions set out in D4.1.

For Precise4Q, understanding the predictive target of a model is crucial to understand how the model output fits into the treatment process in a care setting. This understanding is key in ensuring that the final model is relevant to its domain of application. Furthermore, the definition of a model's predictive target informs much of the technical work involved in developing a model, such as: data design, preparation, and annotation; model design; and the definition of clinical studies to validate the model. The fact that so many technical decisions are informed by (and dependent) on the definition of the predictive targets for the models means that this definition must be specified early on in the model develop process. However, an important caveat on the definition of a predictive target for any model is that the definition of the target is likely to evolve as the model development progresses, this is a natural consequence of the fact that data driven development is fundamentally an iterative and experimental process, best addressed through using an agile project methodology that is capable of responding to new opportunities and resources as they arise. Consequently, this deliverable should be understood as describing an informed (but not final) specification of model targets. Indeed, a number of targets for the models have significantly evolved from those envisaged in the proposal.

The PRECISE4Q model development life cycle is a modified version of the CRoss Industry Standard Process for Data Mining (CRISP-DM), see Figure 1. PRECISE4Q adopted the CRISP-DM lifecycle as a fundamental framework because it is non-proprietary, application, industry, and tool neutral, and it considers the data mining/modelling process from both an

technical and application perspective. Indeed, these features have made CRISP-DM lifecycle the most popular project lifecycle for data science and modelling projects. Following the CRISP-DM framework, the PRECISE4Q model development life cycle is split into several stages, namely: Domain Understanding, Data Understanding, Data Preparation and Harmonisation, Modelling, Validation and Platform Integration.

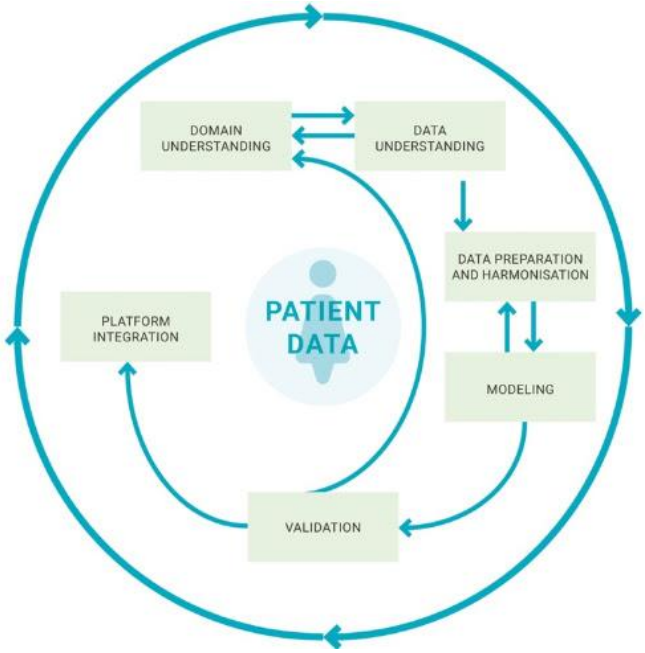

Figure1. PRECISE4Q lifecycle is a modified version of the CRoss Industry Standard Process for Data Mining (CRISP-DM)

A distinctive aspect of the PRECISE4Q approach to model building is that many of the models that will be built by the project are hybrid models: blending phenomenological models with mechanistic models (Nersessian & MacLeod, 2017). The concepts of phenomenological and mechanistic models are drawn on throughout this deliverable so it useful at this point to clarify the meaning of these terms.

Phenomenological models are primarily focused on learning correlations between data points (phenomena). Speaking broadly, the design of these models is generally informed by computational learning theory with a goal of developing a highly accurate model, potentially at the expense of transparency in terms of how the model processes the input information to generate an output. Although these models struggle with questions such as 'why is this the output?', they have the strength that they are capable of handling high-dimensional data, including data for which there is a lack of good domain theory. Deep learning networks are a useful example of phenomenological models. By contrast mechanistic models are informed by domain theory. The internal structure of these models mirror the relationships between components of a theory. The advantage of these models is that the underlying

domain theory provides an explanation for 'why' a model has generated a given output in response to an input. Furthermore, these models often have the ability to be urn as a simulation, enabling the model to be used to examine how a system will evolve into the future. Examples of this type of model includes hand-crafted Bayesian Networks. Figure 2 provides a high level illustration of the distinctions between phenomenological and mechanistic models.

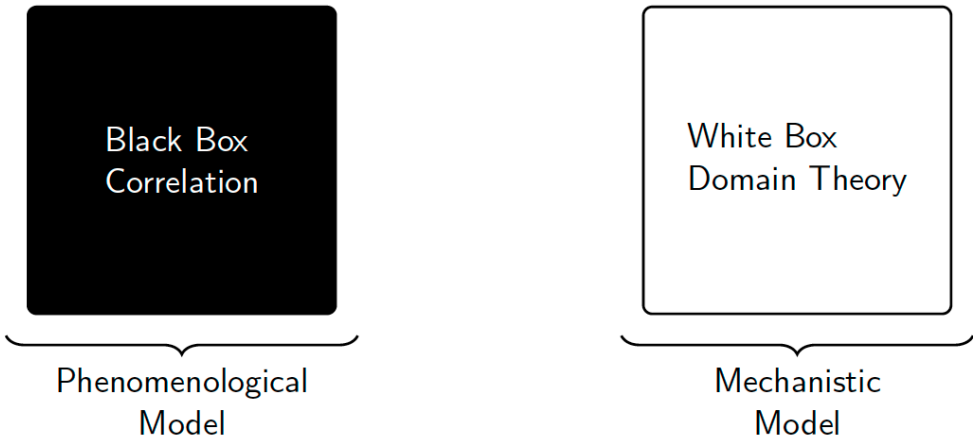

Figure 2. Contrasting phenomenoloigcal versus Mechanistic Models

The development of this deliverable involved a number of interrelated activities. As mentioned above, the definition of any model target should be informed by an understanding of how the model output fits into the treatment process in a care setting. Consequently, defining a model target requires input from a range of stakeholders, including domain specialists, and data and modelling specialists. This deliverable can be understood as arising from the domain understanding stage of the life cycle. An important aspect of the domain understanding stage of the lifecycle is to initiate an interdisciplinary dialog between domain experts (clinicians) and technical experts (data and modelling researchers). This dialog was initiated through the workshops that fed into the creation of this deliverable. The workshop groups were designed to be interdisciplinary in nature, with clinicians, data experts, and modelling experts all collaborating to develop the use-cases and potential model targets. The results of these workshops were also supplemented by reviews of relevant literature, and data exploration activities.

The general structure of the deliverable is that for each of the 4 models (D4.5, D4.6, D4.7, and D4.8) a review of the current instruments used by clinicians to inform their decision making within the clinical setting of the model is presented, this overview is followed by the the specification of the predictive targets for the model, and then a discussion on some future directions that may be investigated to extend the model targets. Note, however, that D4.6 and D4.7 are both designed to be deployed in the acute treatment setting, and so the literature review for this setting is relevant for both of these models.

1. Prediction Targets for D4.5: Hybrid Model for Stroke Prevention

The most direct way to reduce the effects of stroke is to stop it happening. The Precise4Q task T4.5 addresses the challenge of stroke prevention by developing a predictive model that will support clinicians in providing comprehensive and timely evidence-based recommendations to the prevention of stroke. The clinical setting for this this model is patient screening outside of a hospital setting. The goal of this model is to help clinicians provide comprehensive and timely evidence-based recommendations on the prevention of stroke. As set out in Precise4Q deliverable D4.1 Section 1.1 the use case for this model includes both primary prevention (prevention of stroke in patients who have not previously suffered from stroke) and secondary prevention (prevention of recurrence of stroke in an individual). It is likely that building a model for secondary prevention will require a separate model from the model developed for primary prevention; for example, this secondary prevention model will consider a broader set of input features, e.g. neuroimaging, not relevant to primary prevention. However, both the primary and secondary prevention will generate the same output, so for the purposes of this deliverable, which is focused on model outputs, this distinction between primary and secondary prevention will be elided.

Current Practice:

A review of the relevant literature has identified a number of stroke risk scoring schemes that are currently used to screen individuals, these include:

- **CHADS2 score scheme:** The well accepted CHADS2 stroke risk scoring scheme yields a score of 0 to 6, with 1 point each given for congestive heart failure, hypertension, age ≥75 years, and diabetes mellitus and with 2 points given for prior stroke or transient ischemic attack (TIA) (Karthikeyan and Eikelboom, 2010, p. 2). CHADS2 score is based on simple counting system and does not include many well-established risk factors. A variant of this score with additional risk factors included in is the CHA2DS2 -VASc scheme.
- **CHA2DS2 -VASc scheme:** A modification of CHADS2 by adding an age category (1 point for age 65 to 74 years, 2 points for age ≥75 years) and adding 1 point each for diagnosis of vascular disease and for the female sex. CHA2DS2 -VASc gives better stratification of individuals estimated to be at low to moderate risk using CHADS2 (scores of 0 to 1)(Olesen et al., 2012). CHA2DS2 VASc score is better at identifying low risk however both risk scores do not described their outputs in terms of absolute risk of stroke, and, as such, there is no consensus regarding a risk threshold.
- **The Framingham Stroke Risk Score (FSRS)** This scale predicts 10- years probability of stroke. FSRs takes in to account a combination of stroke risk factors including; age, sex, blood pressure, use of antihypertensives, left ventricular hypertrophy, prevalent cardiovascular disease, smoking status, current/previous atrial fibrillation, and

diabetes mellitus(Flueckiger et al., 2018). According to FSRS risk stratification; 20% or greater is considered high risk, 10% to <20% is intermediate risk, and <10% low risk.

- **QStroke:** Another more recent score variant that give absolute value for stroke risk is QStroke. QStroke predict the risk of stroke both in populations with atrial fibrillation and those without atrial fibrillation. For this stroke risk scores the challenge is the Choice of appropriate threshold in complex areas dependent on many variables of clinical outcomes and service provision costs.

Although these scoring schemes are widely adopted many of them neglect a range of potentially important factors that may contribute to stroke risk, (Deliverable D4.1 provides a review of these factors).

Target output for the Hybrid Stroke Prevention Model D4.5

The primary prediction target for model D4.5 is: ***an individualised risk of stroke with a parameterised time horizon of 3 to 5 years.*** This targeted outcome is aligned with Primary prevention Scenario/ Use case 1 and 2 for deliverable 4.1. This will assist clinicians to categorise patients into risk or low or no risk category with the help of the PRECISE4Q models, e.g. as part of a mobile app. It will help clinicians to select from cigarette smoking cessation (risk factor mitigation), diet, and exercise (increase of health factors to reduce stroke risk will help reduction of occurence of stroke).

The concept of risk can be computationally modelled in a number of different ways. For example, in survival analysis the risk of an individual suffering a hazard is often modelled in terms of a risk level relative to a population, sometimes expressed as a hazard ratio. A Cox proportional hazards model is an example of the type of survival model which could generate these types of relative risk scores. Although these models, and relative risk scores, are something that may be considered during the project, the initial prediction target for the prevention model will be framed in terms of an absolute measure of stroke risk with a time horizon of 3 to 5 years. Furthermore, the scale range of risk score will be expressed in terms of percentage (0-99%). From previously developed scores a reasonable threshold of stratification of patients can be; 20% or greater considered high risk, 10% to <20% intermediate risk and <10% low risk.

D4.5 is designed as a hybrid model integrating mechanistic (domain theory) modelling components and phenomenological (machine learning theory) modelling components. This approach is outlined in D4.4. A key (and distinctive) advantage of this hybrid architecture is that the mechanistic component can be run as a simulation of the patients trajectory through time under different intervention/treatment (and patient compliance) conditions. These simulated values can then be used, together with measured data, as input to the hybrid model to calculate the risk prediction (see D4.4 figure 9 for details). The mechanistic

models will integrate (at least), and be capable of simulating the evolution of the following biomarkers for an individual patient over a 3 to 5 year time horizon:

Table 1: Some of the biomarkers simulated by the mechanistic models

Plasma glucose
Plasma insulin
Diabetes
Blood flow
Vessel compliance
Glucose uptake in adipose tissue
Blood oxygenation in brain
Liver function
Body weight
Lean and fat tissue mass
Energy intake

A crucial step in running such a mechanistic simulation is the initialisation of the model. Ideally, the model should be initialised for an individual using biomarkers measurements from the patient. However, it is often the case that only a subset of the biomarkers will be available for a given individual at the initialisation of the model. In these cases, the values for the biomarkers that are not available at initialisation need to be imputed or predicted by secondary models. Consequently, ***the biomarkers listed above are also prediction targets for the prevention stage of stroke management***.

Phenomenological models or imputation can also be used to estimate unmeasured biomarkers relevant for the risk prediction. Below are some such biomarkers listed(**Brännmark**. etal., 2013,Casas et al., 2018)

Table 2: List of biomarkers to estimated in risk prediction.

Biomarker	*Explanation*
Cardiomegaly	Enlargement of heart
Central systolic blood pressure	Blood pressure in the aorta
C-reactive protein	Protein whose levels rise in response to inflammation
(Carotid-femoral) Pulse wave velocity	Velocity at which the blood pressure pulse propagates through the circulatory system. Measure of arterial stiffness in arteries to head and neck
Augmentation index	Measure of systemic arterial stiffness derived from the ascending aortic pressure waveform
Intima–media thickness	Thickness of the two innermost two layers of the wall of an artery. Measure of Atherosclerosis.
Glomerular filtration velocity	Velocity of fluid and waste filtration from blood. Measure of kidney function.

The individualised risk of stroke for a patient will thus be calculated as a function of the state of the simulated biomarkers at a given time point, biomarkers estimated by phenomenological models, and measured biomarkers.The biomarker to risk of stroke function will be defined using a phenomenological prediction model. For example, in the simplest case the phenomenological prediction model will be defined as a logistic regression model that takes the above biomarkers as inputs. Note that the fact that the prediction process involves a sampling of variables from the mechanistic model, and the use of these sampled values as inputs to the phenomenological model opens the possibility that other input features (not included in the mechanistic model) can be included as inputs to the phenomenological model. For example, these extra features might include demographic features, and/or flags indicating events in an individuals history (such as previous strokes). The inclusion of these extra features means that the prediction of risk of stroke is not limited to only the biomarkers modelled by the mechanistic model. Furthermore, it provides a natural way to parameterize the model for both primary and secondary stroke prevention scenario, (we will discuss the secondary stroke prevention scenario in more detail below).

To illustrate the interaction between the mechanistic and phenomenological model, consider the task of generating an individualised risk of stroke with a time horizon of 3 to 5 years for an individual under the condition that they have complied with an intervention programme for six months from today. To generate such a risk assessment the following steps would be followed:

1. Initialise the mechanistic model with the current measured values of biomarkers for the individual and estimated biomarkers.
2. Run the mechanistic model for a time simulation of six months with the condition of the intervention.
3. Sample the evolved values of the biomarkers from the simulation model and use them as inputs to a phenomenological model together with measured and estimated biomarkers to generate the risk of stroke.

In short, the model will be capable of generating a individualised risk of stroke for a patient at different time points by running the simulation to the required time point, sampling from the simulation and using the sampled values as inputs to the prediction model.

Potential Future Directions of Research Arising from D4.5

Identifying Health Factors: The development of a prediction model has a direct benefit in terms of the insight that the model's predictions can provide to a decision maker. However, an secondary benefit of the model is that the structure of the model (for example, in terms of the weighting the model may apply to a biomarker or feature) can be revealing in terms of the importance of a feature in determining an outcome. Here again, the use of a mechanistic model integrating multiple biomarkers and capable of simulating the evolution of these biomarkers through time is very useful. In particular, it provides a basis for analysing the biomarkers covered by the model to identify potential health factors which prevent stroke in contrast to risk factors that increase the likelihood of stroke. Especially - so far maybe unknown - genetic factors could be found this way.

Extended Model Targets: As mentioned in the executive summary the specification of a models prediction target often evolves as the development of the model progresses: new requirements are identified, or new resources become available. An interesting area of potential future work include extending the model to predict an individual's compliance to treatment, or level of understanding or engagement. However, the benefits of such work at present is currently open to debate, Depending on causality, avoidance of a risk factor may or may not lead to lower incidence of the disease (Hollnagl et al., 2000). Also, influencing one parameter - even if causal - does not necessarily lead to decreased incidence, when the genesis of the disease is multifactorial and risk factors influence each other (Hollnagl et al., 2000)

2. Prediction Targets for D4.6 (Hybrid Model for Short Term Stroke Outcome) and D4.7 (Personalised Acute Stroke Quality of Life Prediction Model)

Precise4Q will develop two separate but complementary prediction models that focus on the acute stroke management stage. The first of these models D4.6 Hybrid Model for Short Term Stroke Outcome will be developed through the activities carried out in Task 4.6. This model will use a variety of hospital data (including medical imaging and other forms of data) and integrate mechanistic and phenomenological models to predict short term stroke outcomes (e.g. discharge NIHSS and mRS 3 months post treatment) with a special focus on different treatment options as inputs. The second model D4.7 Personalised Acute Stroke Quality of Life Prediction Model will be developed through the activities carried out in Task 4.7. The basis for this model is to move beyond the estimates of short-term disease outcome generated by D4.6 and instead generate a complex structured quality of life target profile for a patient.

Acute stroke management starts when a neurologist has examined the neuroimage of the patient along with the clinical presentation taken into account the time window as any likely suspicion of ischemia leads to initiation of stroke treatment. As set out in Precise4Q deliverable D4.1 Section 1.2 the use case for this model is a clinician in an acute care hospital who is deciding on the appropriate stroke treatment intervention for an individual, be it time-based dissolution of the obstructing blood-clot by either intravenous thrombolysis or mechanical thrombectomy. The goal of the D4.6 and D4.7 prediction models is to support the clinician in making this decision by providing them with fine-grained estimates of the patient outcomes under different treatment conditions that extends well beyond the functional outcome paradigms that are used in current practice.

Current Practice:

Currently, the most widely used instrument for measuring stroke patient outcomes is the **7-level modified Rankin Scale** (mRS). The mRS outcome scale for acute stroke provides a graded evaluation of global disability, mainly in term of motor functions for a patient in daily living (Rankin, 1957). It is defined between 0 and 6, where an mRS of 0 indicates that no symptoms for disability are present, 5 denotes the most severe disabilities, and 6 records that the patient did not survive (Van Swieten et al., 1988) (Table 1). The strength of mRS is that it covers the entire range of functional outcomes. Also, because mRS is categorized intuitively it is easily understandable by patients as well. However the shortcoming of the mRS scale is the subjective determination between categories and the difficulty in reproducible scoring by clinicians and patients. Besides the 7 level ordinal mRS scale, mRS is

also used as binary or dichotomous scale: good or bad prognosis. This dichotomization of mRS scores into binary is done by binning mRS 0-2 into a favourable outcome group and mRS 3-5 into a severe outcome group. The shortcoming of dichotomous mRS is that it does not include the entire range of outcomes on the other hand ordinal mRS scale is inclusive of both positive and negative outcomes.

Table3: mRS Score Description from D1.3 PRECISE4Q

mRS	Score Description
0	No symptoms at all
1	No significant disability despite symptoms; able to carry out all usual duties and activities
2	Slight disability; unable to carry out all previous activities, but able to look after own affairs without assistance
3	Moderate disability; requiring some help, but able to walk without assistance
4	Moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance
5	Severe disability; bedridden, incontinent and requiring constant nursing care and attention
6	Dead

Recently the STAIR (Stroke Therapy Academic Industry Roundtable) recommended the development of a utility-weighted (UW) version of the mRS. Investigators subsequently calculated utility values for the various levels of the mRS by mapping responses from the EQ-5D (European Quality of Life Scale) (Rabin and Charro, 2001) onto the mRS levels in populations of patients with stroke. In another study, disability weights for mRS levels were derived using the methodology of the WHO GBD (World Health Organization Global Burden of Disease Project) (Hong and Saver, 2009). On the basis of these approaches, a UW-mRS was created. Both Ordinal analyses and the UW-mRS approach capture the entire distribution of outcomes (good and bad) when compared with a dichotomous approach. The advantage of the translation of the UW-mRS into quality-life-

years can communicate the efficacy of a treatment to patients and physicians (Broderick et al., 2017).

The National Institutes of Health Stroke Scale (NIHSS) is an alternative quantification of impairment caused by stroke. It evaluates initial neurologic outcome of stroke and degree of recovery for patients with stroke (Schlegel et al., 2003, 2004). The scale assesses impairment for 15-items, including: level of consciousness, extraocular movements, visual fields, facial muscle function, extremity strength, sensory function,coordination (ataxia), language (aphasia), speech (dysarthria), and hemi-inattention (neglect) (Lyden et al., 1999, 2001).

Although the mRS (and its variants) and NIHSS scales are both widely used in practice, they both suffer from the fact that they reduce the patient outcome to a single number. Such a compression of information, although useful at a aggregate level, leads to a conflation of distinct patient outcome profiles. Precise4Q will address this shortcoming by developing predictive models that provide clinicians with a broader Quality of Life (QoL) profile for patient outcomes under different treatment conditions. This precision medicine approach to treatment of stroke has the potential to lead to a more informed choice of therapeutic interventions and a more rational allocation of resources.

Target Output for the Hybrid Model of Short Term Stroke Outcome D4.6

Model D4.6 is a hybrid model that integrates a mechanistic simulation model of blood perfusion in the brain with a phenomenological model utilizing neuroimaging and clinical data to predict a prospective estimate of a short-term disease outcome.

The mechanistic simulation will take structural vessel imaging (both available in MRI and CT) as input and will generate a 2D image of the Circle of Willis for the patient. Furthermore, leveraging the fact that this mechanistic model is capable of simulating the brain through different time-scales and under different conditions, this model will be used to generate **2D images of the Circle of Willis and major brain arteries at under different blood pressure setting (thereby enabling the exploration of different boundary conditions)**.

The phenomenological model will take a number of different streams of information as input including:

1. 2D Output of Blood Vessels generated by the Mechanistic Simulation Model
2. Clinical Data
3. 3D Neuroimaging

The phenomenological model will then output the short-term stroke outcome estimate. The initial prediction target for this model will be a **binarised mRS scale at 3 months post**

treatment: favourable outcome (mRS 0-2), severe outcome (mRS 3-5). However, we will also explore the possibility of predicting **NIHSS at discharge f**or a patient. (See the above discussion on Acute Current Practice for more information on mRS and NIHSS). Importantly, the clinical data will include information about the chosen treatment option, thus allowing to compare the outcome based on treatment decisions.

Target Output for the Personalised Acute Stroke Quality of Life Prediction Model D4.7

The concept of Quality of Life (QoL) in a health context is somewhat ambiguous and can be given at least two alternative definitions (Fayers and Machin, 2016):

1. the set of outcomes that contribute to a patient's well-being or overall health
2. a summary measure or scale that purports to describe a patient's overall well-being or health

Traditionally most machine learning models focus on predicting a single value. This single value output approach fits naturally with QoL definition 2 ('a summary measure or scale'). The target output for D4.6 can be understood as such a scale. A distinctive aspect of the Precise4Q D4.7 model is that it adopts the broader conceptualization of QoL given in definition 1 above, as a 'set of outcomes'. Importantly this set of outcomes are likely to be interrelated (as opposed to independent). Viewed in this context the task of predicting a QoL for a patient post treatment should be understood as predicting a complex multi-faceted and interrelated structure. Figure 3 provides a graphical illustration of the concept of predicting a complex interrelated structure.

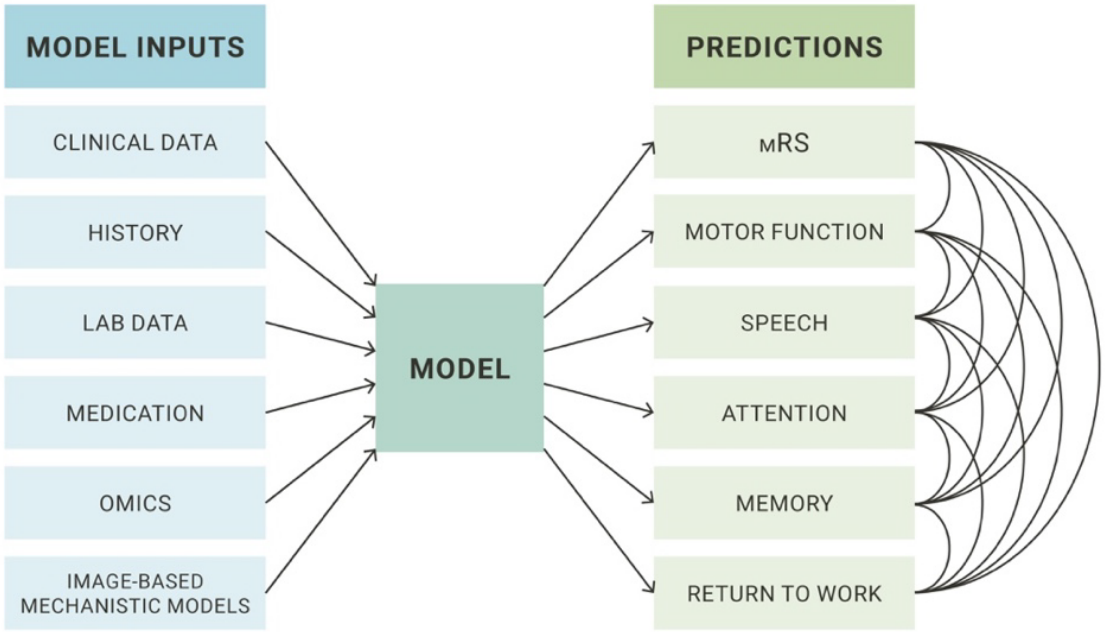

Figure 3: Illustration of a Model Predicting a Complex Interrelation Structure

Given this ambitious goal the task at hand in this document is to define the set of components that should be included within this structure. Within a medical context, QoL is generally measure post-treatment using a patient survey. Many such surveys exists, including:

- **Medical Outcomes Study Short Form 36 (SF-36)**: The Medical Outcomes Study 36-item Short-Form Health Survey is a widely used, generic, patient-report measure created to assess health-related quality of life (HRQOL) in the general population. SF-36 is the most commonly used generic instrument for measuring quality of life (de Haan, 2002). The SF-36 can be used, but is not limited to, persons with stroke. The Stroke Impact Scale (SIS): The Stroke Impact Scale (SIS) is a stroke-specific is a questionnaire for patients to report their health status. It is an assessment of several stroke outcomes, including strength, hand function Activities of Daily Living / Instrumental Activities of Daily Living (ADL/IADL), mobility, communication, emotion, memory and thinking, and participation.
- **EuroQoL:** EuroQoL is based on a standardized health state description, consisting of five domains including; mobility, self-care, usual activities, pain/discomfort and anxiety/depression, each of which has three levels of severity (Brooks and Group, 1996).
- **Health Utilities Index (HUI):** HUI is a family of generic health profiles and preference-based systems for the purposes of measuring health status, reporting health-related quality of life, and producing utility scores(Mathias et al., 1997).

- **Stroke and Aphasia Quality Of Life Scale (SAQOL-39):** The stroke-specific quality of life (SS-QOL) ((Williams et al., 1999))questionnaire to measure the QoL of people with aphasia following the stroke (Stroke-Aphasia Quality of Life – SAQOL scale). These authors evaluated the psychometric properties of the initial SAQOL version (containing 53 items) and refined it to a shorter version consisting of 39 items (SAQOL-39).
- **Stroke-Specific quality Of Life Scale:** The Stroke-Specific Quality of Life Scale (SS-QOL) (Williams et al., 1999) is a survey questionnaire with 49 items to access multiple domains of QOL based on the past week information. This scale takes into account the physical, psychological and social aspects of people's lives, these 12 quality domains are: energy, family roles (defined by patients as relationships and work within the family), language, mobility, mood, personality, self-care, social roles, thinking, upper extremity function, vision, and work/productivity. Each of these domains are individually scored and their global sum is also calculated. The scale is formatted to record three response sets in 5-point Likert format based on amount of help required (no help - total help), amount of trouble experienced when attempting tasks (unable to do it-no trouble), degree of agreement with statements regarding function (strongly agree-strongly disagree. The scale is in agreement with the International Classification of Functioning, Disability, and Health (Lin et al., 2010). SS-QOL has moderate ability to detect change in patients between 1 and 3 months post-stroke. the SS-QOL, has a small to moderate ability to detect change in patients between 3 and 12 month post-stroke. the minimal clinically detectable difference for the mobility, self-care and upper extremity function subscales was defined as a mean change in score of at least 1.5, 1.2 and 1.2 respectively.

There are a number of data sources available to the Precise4Q consortium. One such source is that Swedish Stroke Registry (Riksstroke). Rikstroke carry out a 3-month post-stroke treatment survey of stroke sufferers. This survey covers a range of topics, including question relating to stroke services. However, many of the questions relate to quality of life and span topics from pain, tiredness, dependence on support to carry out activities of daily living (such as getting dressed and undressed, visiting the toilet, etc.). Based on a review of these surveys it has been decided that they would provide a useful basis to train models to predict a quality of life profile. Note that the majority of questions on these forms have ordinal responses. Using the 2018 Riksstroke 3 Month Follow-up survey as reference, model D4.7 will predict the patient's response to the following questions at 3 months post treatment: Q3, Q4, Q5, Q6, Q7, Q14, Q15, Q19, Q20, Q21, Q22, Q23, Q24. Additionally, the developed QoL-targets will be included in the acute clinical study protocol in WP5.

3. Prediction Targets for D4.8 (Personalised Rehabilitation Model)

Stroke rehabilitation is managed by an interdisciplinary team working cohesively and closely to provide a comprehensive program for each patient. Rehabilitation has been shown to be most beneficial when started early, although recovery of stroke-related impairments is still possible even years later. Stroke recovery is influenced by a variety of intrinsic and extrinsic factors that influence the likelihood and degree of neurological reorganization.

During rehabilitation multidisciplinary teams establish or regularly revise rehabilitation goals and plans, assess patient progress, identify barriers or complications, and develop plans for discharge or transfer to another type of rehabilitation program. These programs may vary in the types of therapies offered as well as their intensity, frequency, and duration.

Stroke rehabilitation focuses on both cognitive and functional impairments. Here we discuss the domains of cognitive and functional impairments associated with stroke and the scales used for their assessment at various stages of rehabilitation.

Current Practice: Post-stroke Cognitive Rehabilitation

20% to 80% of ischemic stroke result in cognitive impairment (Sun et al., 2014). These impairments often effect the main cognitive function involved in performing activities of daily living (ADLs); such as attention, memory, and executive functioning. Consequently, stroke patients are given specific cognitive training tasks to perform as part of the rehabilitation. Meaningful task-specific training (MTST) is important for improvement in function and focus.

At the initiation of a cognitive rehabilitation (CR) program patients are assessed for various cognition domains using an Neuropsychological Assessment Battery (NAB) and also before and after each therapy. Differences between before and after therapy NAB test scores helps access patient improvement in the cognitive field. For example, recording deficit reduction is a helpful tool in cognitive skill improvement therapy whereby after each session of treatment, the level of reduction in a particular deficit per specific function is recorded and also the overall deficit reduction at global level using an NAB.

There are various treatment configurations for cognitive rehabilitation based on the variety of tasks done and the number of repetitions of these tasks. The performance of patients is monitored for each training task. There are several factors that are considered in the design and updating of a CR program, including:

- **Neuro-rehabilitation range** (NRR) (García-Rudolph and Gibert, 2014) is another component of cognitive rehabilitation programs. NRR metric is two dimensional: the number of executions of a task during a CR treatment and the performance in each execution of the task.This range assessment is useful as it helps to choose tasks

suited the performance ability of the patient. NRR is used in repetitive task training in order to estimate the repetition of achievable tasks.

- The **effectiveness of selected training tasks** measured a ratio between the patient potential (in terms of skills of the targeted patient) and the level of difficulty associated with the fulfillment of the task (Carey et al., 2007).
- **Patient compliance** directly impacts the end goal of the task. Therefore level of compliance for patients at training task are recorded for both global cognition and at at functional cognitive level. Patient compliance assessment involves comparison of the extent of activity actually performed to the levels of performance recommended (For example a recommendation can be at least 30 minutes of non-stop activity three times a week). Three months after the end of the programme, the patient's statement of activity performed and actual activity performance recommended (in terms of frequency and duration) are compared. Patient Compliance can be given as a percentage of assigned activity actually performed by the patient.

Current Practice: Post-stroke Functional Rehabilitation

Rehabilitation outcomes in terms of functional independence are assessed on the basis of the ability to carry out different activities of daily living (ADLs). ADLs include activities such as: feeding, bathing, grooming, dressing, bowel control, bladder control, toileting, chair transfer, ambulation and stair climbing.The level of functional independence metrics reflect the amount of support needed by stroke survivor to carry out these activities. Broadly speaking these levels of functional independence include: No Helper required (Complete Independence and Modified Independence where no help is needed); Supervision needed; Minimal Assistance; Moderate Assistance; Maximal Assistance, and Complete Assistance required. There are several metrics are available for evaluation for functional independence outcomes, we discuss some commonly used here.

- **Barthel Index** (BI) is a classic functional independence metric. BI measures the extent to which a stroke survivor can function independently and has mobility in their ADLs. The index also gives indication on the need for assistance in care (Collin et al., 1988). A baseline Barthel index score is recorded at the time of discharge from stroke unit Baseline BI. Later during the rehab the update in BI score is recorded and compared with Baseline BI to access improvement in functionality.
- **Functional Independence Measure (FIM)** is a global measurement for disability, comprised of a checklist of 18-item of physical, psychological and social functionality (Linacre et al., 1994). This FIM scale setting is based on the International Classification of Impairment. FIM Scoring can be carried out at patient admission to a stroke unit and is termed an admission score and a discharge score just before discharge from a stroke unit. This short version of FIM scoring is called AlphaFIM. It is also recommended to set a goal FIM score based on the intervention chosen.

- **The Frenchay Activities Index (FAI)** is a measure of instrumental activities of daily living. One benefit of using this metric is that it provides a broader measurement of actual activities patients have undertaken in the recent past (Wade, Legh-Smith, & Langton, 1985).
- **Stroke Impact Scale (SIS)** is a stroke specific self-report metric designed to assess multidimensional stroke outcomes. It covers ADL/IADL, mobility, communication, emotion, memory and thinking, and participation.

The functional independence of a stroke rehabilitation patients is assessed on the first day of rehabilitation and is the basis for the initial design of the functional rehabilitation programme for a patient. However, to maximize the benefit of a rehabilitation programme for a patient the programme should be updated on a regular basis in response to the patient progress (for example, as measured by task completion and repetition.

Target output for model 4.8

The updating and maintenance of rehabilitation programmes is an complex process and can have a significant effect of the success of the rehabilitation. Model D4.8 is designed to support the updating of a patient's rehabilitation programme. The model will be run daily and will **generate a personalised rehabilitation activity schedule for the day**. Each activity in the schedule will be annotated with a target performance level. The model will also **forecast the patient's cognitive and functional status on the final day of discharge**.

Workshops with rehabilitation specialists at Guttman also identified a novel model target which will quantify the '**Fragility**' of a patient post rehabilitation. Within the Guttman Rehabilitation program the concept of patient Fragility captures the likelihood (or not) of a patient to relapse, or retreat from life, once they are discharged from rehabilitation. Unfortunately, many of the functional and cognitive gains that a patient achieves during rehabilitation (with the intensive support of the rehabilitation team) can be lost post rehabilitation if the person becomes inactive post discharge. There are a broad range of factors that can contribute to this unwanted outcome. For example factors extrinsic to an individual such as lack of family or financial support, or living remotely. Rehabilitation specialists intuitively model these factors through the concept of 'Fragility' (that idea of fragility is mostly applied to most severe cases in terms of risk of psicosocial exclusion but in this work we could refer to it as an indicator of shot, mild and long term trajectories or evolution) and patients who are deemed to be Fragile at discharge are provided with extra supports as they transition back to independent living, outside of the rehabilitation hospital. Deciding which patients are Fragile (and thereby receive these extra supports) is a difficult and multi-factored task. However, based on discussion with Guttman staff, it should be possible to analysis Guttman records to identify patients that were deemed Fragile or non-

Fragile at discharge. On the basis that this data preparation task is completed successfully **a binary Fragile flag** will be added to the model outputs and will be used by rehabilitation professionals both to identify Fragile patients, and also to plan and resource the supports required by Fragile patients.

4. Conclusions

In this report we identify target outputs pairs for the Precise4Q predictive models D4.5, D4.6, D4.7, and D4.8. The definition of these targets was informed by the definition of relevant use-cases (see Deliverable D4.1), the data available to the consortium, and workshops that initiated and facilitated multi-disciplinary dialog on requirements. As noted at the start of this document, the primary purpose of this document is to provide clear (agreed) targets to inform model development. However, these targets may be updated as the research and model development progresses. Indeed, it is likely that these target models outputs will be updated and verified with time during model building and validation process, taking in to account cross disciplinary feedback on their usage in decision support system for stroke management.

5. References

Brännmark, C., Nyman, E., Fagerholm, S., Bergenholm, L., Ekstrand, E.-M., Cedersund, G., Strålfors, P., 2013. Insulin Signaling in Type 2 Diabetes. J. Biol. Chem. 288, 9867–9880.

Brooks, R. and Group, E., 1996. EuroQol: the current state of play. Health policy, 37(1), pp.53-72.

Broderick, J.P., Adeoye, O. and Elm, J., 2017. Evolution of the modified Rankin scale and its use in future stroke trials. Stroke, 48(7), pp.2007-2012.

Casas, B., Viola, F., Cedersund, G., Bolger, A.F., Karlsson, M., Carlhäll, C.-J., Ebbers, T., 2018. Non-invasive Assessment of Systolic and Diastolic Cardiac Function During Rest and Stress Conditions Using an Integrated Image-Modeling Approach. Front. Physiol. 9.

Carey, J.R., Durfee, W.K., Bhatt, E., Nagpal, A., Weinstein, S.A., Anderson, K.M., Lewis, S.M., 2007. Comparison of finger tracking versus simple movement training via telerehabilitation to alter hand function and cortical reorganization after stroke. Neurorehabil. Neural Repair 21, 216–232.

Collin, C., Wade, D.T., Davies, S., Horne, V., 1988. The Barthel ADL Index: a reliability study. Int. Disabil. Stud. 10, 61–63.

Flueckiger, P., Longstreth, W., Herrington, D. and Yeboah, J., 2018. Revised Framingham Stroke Risk Score, Nontraditional Risk Markers, and Incident Stroke in a Multiethnic Cohort. Stroke, 49(2), pp.363-369.

García-Rudolph, A., Gibert, K., 2014. A data mining approach to identify cognitive neurorehabilitation range in traumatic brain injury patients. Expert Syst. Appl. 41, 5238–5251.

Hollnagel, H. and Malterud, K., 2000. From risk factors to health resources in medical practice. Medicine, Health Care and Philosophy, 3(3), pp.255-262.

Lin, K.C., Fu, T., Wu, C.Y., Hsieh, Y.W., Chen, C.L. and Lee, P.C., 2010. Psychometric comparisons of the stroke impact scale 3.0 and stroke-specific quality of life scale. Quality of Life Research, 19(3), pp.435-443.

Linacre, J.M., Heinemann, A.W., Wright, B.D., Granger, C.V., Hamilton, B.B., 1994. The structure and stability of the Functional Independence Measure. Arch. Phys. Med. Rehabil. 75, 127–132.

Lyden, P.D., Lu, M., Levine, S.R., Brott, T.G. and Broderick, J., 2001. A modified national institutes of health stroke scale for use in stroke clinical trials. Stroke, 32(6), pp.1310-1317.

Lyden, P., Lu, M., Jackson, C., Marler, J., Kothari, R., Brott, T. and Zivin, J., 1999. Underlying structure of the National Institutes of Health Stroke Scale: results of a factor analysis. Stroke, 30(11), pp.2347-2354.

Mathias, S.D., Bates, M.M., Pasta, D.J., Cisternas, M.G., Feeny, D. and Patrick, D.L., 1997. Use of the Health Utilities Index with stroke patients and their caregivers. Stroke, 28(10), pp.1888-1894.

Olesen, J.B., Torp-Pedersen, C., Hansen, M.L., Lip, G.Y., 2012. The value of the CHA2DS2-VASc score for refining stroke risk stratification in patients with atrial fibrillation with a CHADS2 score 0–1: a nationwide cohort study. Thromb. Haemost. 108, 1172–1179.

Rankin, J., 1957. Cerebral vascular accidents in patients over the age of 60: II. Prognosis. Scott. Med. J. 2, 200–215.

Schlegel, D., Kolb, S.J., Luciano, J.M., Tovar, J.M., Cucchiara, B.L., Liebeskind, D.S. and **Kasner**, S.E., 2003. Utility of the NIH Stroke Scale as a predictor of hospital disposition. Stroke, 34(1), pp.134-137.

Sun, J.-H., Tan, L., Yu, J.-T., 2014. Post-stroke cognitive impairment: epidemiology, mechanisms and management. Ann. Transl. Med. 2.

Van Swieten, J.C., Koudstaal, P.J., Visser, M.C., Schouten, H.J.A., Van Gijn, J., 1988. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 19, 604–607.

Wade, D.T., Legh-Smith, J. and Hewer, R.L., 1985. Social activities after stroke: measurement and natural history using the Frenchay Activities Index. International rehabilitation medicine, 7(4), pp.176-181.

Williams, L.S., Weinberger, M., Harris, L.E., Clark, D.O. and Biller, J., 1999. Development of a stroke-specific quality of life scale. Stroke, 30(7), pp.1362-1369.