Digital twins and hybrid modelling for simulation of physiological variables and stroke risk

Tilda Herrgårdh[1]*, Elizabeth Hunter[2]*, Kajsa Tunedal[1,6], Håkan Örman[1], Julia Amann[3], Francisco Abad Navarro[4], Catalina Martinez-Costa[4], John D. Kelleher[2,5], Gunnar Cedersund[1,6]

1. Department of Biomedical Engineering, Linköping University, Linköping, Sweden
2. Information, Communications, and Entertainment (ICE) Institute, Technical University Dublin, Dublin, Ireland
3. Health Ethics and Policy Lab, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland
4. Department of Informatics and Systems, University of Murcia, Murcia, Spain
5. ADAPT Research Centre, Technical University Dublin, Dublin, Ireland
6. Center for Medical Image Science and Visualization, Linköping University, Linköping, Sweden

* Contributed equally

ABSTRACT

One of the more interesting ideas for achieving personalized, preventive, and participatory medicine is the concept of a digital twin. A digital twin is a personalized computer model of a patient. So far, digital twins have been constructed using either mechanistic models, which can simulate the trajectory of physiological and biochemical processes in a person, or using machine learning models, which for example can be used to estimate the risk of having a stroke given a cross-section profile at a given timepoint. These two modelling approaches have complementary strengths which can be combined into a hybrid model. However, even though hybrid modelling combining mechanistic modelling and machine learning have been proposed, there are few, if any, real examples of hybrid digital twins available. We now present such a hybrid model for the simulation of ischemic stroke. On the mechanistic side, we develop a new model for blood pressure and integrate this with an existing multi-level and multi-timescale model for the development of type 2 diabetes. This mechanistic model can simulate the evolution of known physiological risk factors (such as weight, diabetes development, and blood pressure) through time, under different intervention scenarios, involving a change in diet, exercise, and certain medications. These forecast trajectories of the physiological risk factors are then used by a machine learning model to calculate the 5-year risk of stroke, which thus also can be calculated for each timepoint in the simulated scenarios. We discuss and illustrate practical issues with clinical implementation, such as data gathering and harmonization. By improving patients' understanding of their body and health, the digital twin can serve as a valuable tool for patient education and as a conversation aid during the clinical encounter. As such, it can facilitate shared decision-making, promote behavior change towards a healthy lifestyle, and improve adherence to prescribed medications.

Introduction

Healthcare is currently moving towards P4 medicine - personalized, predictive, preventive, and participatory medicine (1). This development away from reactive medicine, towards P4 medicine, is driven and facilitated by the increase in chronic diseases (2), by a general increase in mechanistic understanding behind disease development, and by technological advancements (3,4).

One already existing clinical practice, which has the potential to include all P4 principles, is the preventive health dialogue (Figure 1A). During such a health conversation, an individual and a specialized nurse discuss different health aspects regarding the individual's current lifestyle and health status. The discussions cover the current and projected risk of getting different lifestyle related diseases, such as a stroke. The discussions also cover what an individual can do to decrease that risk (Figure 1A and B). With this as a basis, the individual decides which lifestyle changes to do, and is also responsible for maintaining those lifestyle changes (Figure 1C).

While the existing health dialogue already features some aspects of the P4 principles, these principles could be further strengthened in various ways. Firstly, today, the health dialogue does not make use of advanced predictions using personalized computer models. In Figure 1B, such predictions show the effect of two different lifestyles on the risk of a stroke, by simulating physiological variables such as blood pressure and BMI. Secondly, another possibility with computer models is to zoom in on things happening on shorter timescales and other biological levels, using multi-level and multi-timescale models (Figure 1C). By simulating short-term changes, one can compare those predicted changes with measured outcomes during different time points. If those predictions turn out to be correct, it has the potential to increase faith in the long-term predictions, give continuous feedback on progress, and thereby increase motivation. In Figure 1C, this potential is exemplified with a simulated long-term decrease in BMI, which is associated with changes in short-term meal response. One technology that could contribute to all these possibilities is digital twins.

A digital twin of an individual is a personalized computer model of that individual, and it has been suggested that such twins can be used for prevention of cardiometabolic diseases such as diabetes and stroke (5–7). Digital twins can be described by black box phenomenological models or by mechanistic physiology-based models. Phenomenological black box models are suited for making predictions and statistical assessments, such as the risk of a stroke. In contrast, mechanistic models describe physiology and biochemistry using mechanistic data, typically time series. These mechanistic models are first developed using varying amounts of *in vitro*, *in vivo*, and clinical population data, to be able to describe the physiological and biochemical processes (8). Then, these models are individualized using data from the particular individual, to be able to make personalized predictions. Finally, a twin can be either online, if it is continuously updated, or offline, if it is occasionally updated (9)
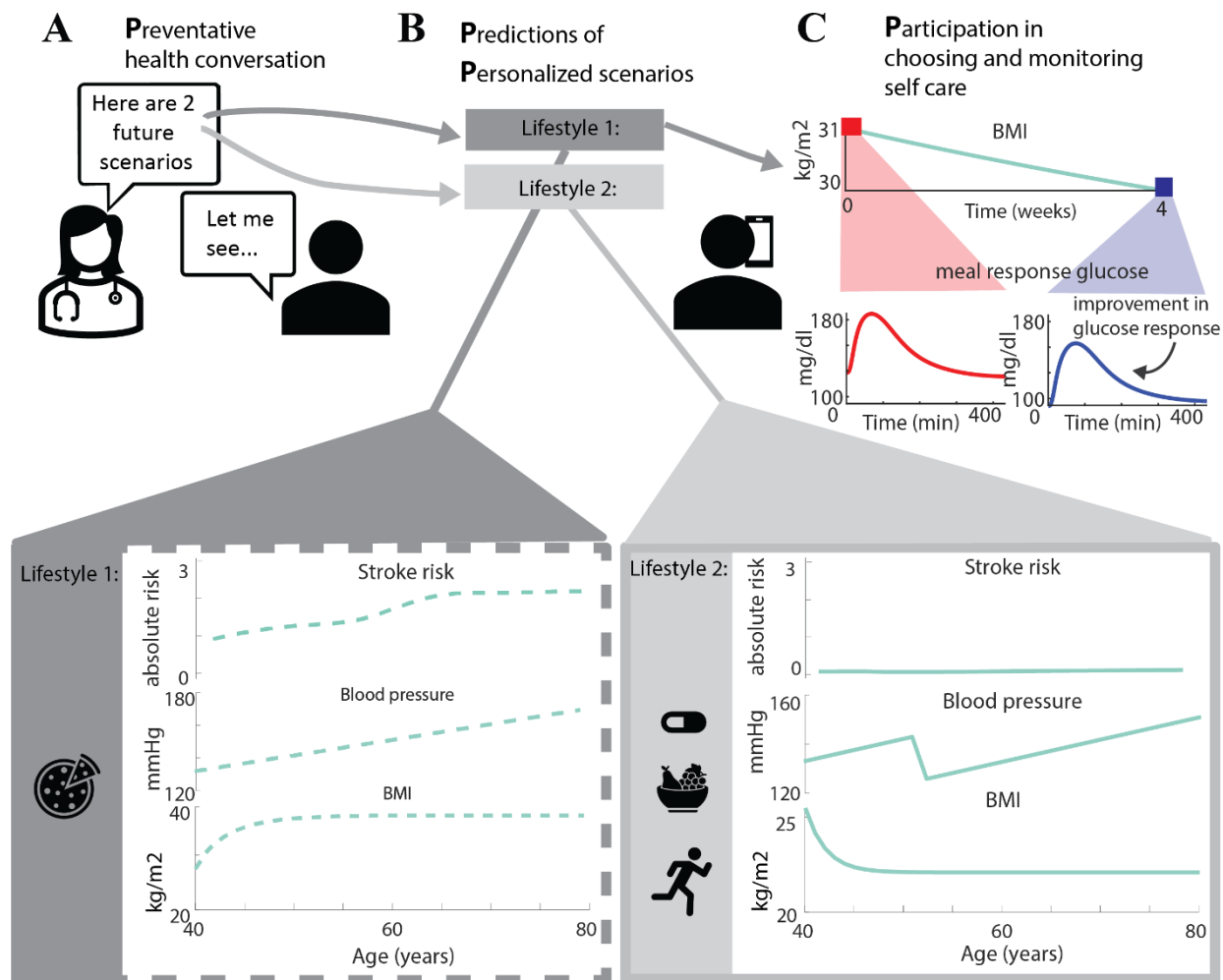
Figure 1: Vision for P4 medicine - preventive, predictive, personalized, and participatory medicine. A) during preventive health dialogues, different future scenarios are presented and discussed. Here, two examples of such scenarios are shown - B) Lifestyle 1: The person does not take any blood pressure medication. The person continues the unhealthy lifestyle, thereby increasing BMI, and eventually developing diabetes. The risk increases during the simulated 40 years. Lifestyle 2: The person starts to lose weight at 40, consequently decreasing BMI, and starts to take blood pressure medication when turning 50. The risk of stroke does not increase much during the 40 years. C) The person participates in his/her own care by choosing what lifestyle changes to make and by continuously monitoring the progress and results of the lifestyle changes. Here, this monitoring is exemplified by looking at how weight loss intervention results in improved meal glucose response. By continuously comparing such predictions with data, the person can continuously deepen their understanding of their own body, and hopefully better maintain motivation.

One area where digital twins are especially suited is for preventative care of cardiovascular diseases such as stroke. These diseases are at least in part preventable with life-style changes and early medical interventions, and have a complex pathophysiology. The complexity of cardiovascular diseases comprises several medical and environmental factors, and involves multiple organs, timescales, and control mechanisms (10). This complexity makes it difficult to make individual predictions without help from advanced analytical tools such as digital twins. Capturing the complexity with a digital twin could

potentially help physicians and patients choose patient-specific preventive measures, by first testing the different measures on the digital twin and thus identifying the one with the most promising effects for that individual. Despite this potential, digital twins have only rarely been used in healthcare, partly because the models are not yet ready for usage.

One of the reasons that the models are not ready for usage is that the development of these models has relied upon either of two different modelling approaches: mechanistic or black box machine learning models. Mechanistic models describe the underlying physiological processes, leading to changes over time (Figure 2A). Specifically, they can take lifestyle variables (e.g. energy expenditure) and personal parameters (e.g. age) as input, and can then be used to simulate the evolution of biomarkers and their covariation with other biomarkers. These kinds of models can then be used to simulate different scenarios (e.g. different lifestyles) and compare them with each other, on both the long term-level and on short-term effects. Since this approach directly uses prior physiological understanding to develop the models, the resulting models can give comprehensive physiological explanations for their predictions. Mechanistic models are primarily used for improved mechanistic understanding or for predicting the dynamics of physiological variables. Machine learning models, on the other hand, can map between inputs and outputs, such as risk factors and risk scores, without needing any information about the physiological mechanisms in that mapping. These models usually do not learn anything new about the mechanisms of the physiological system that produced the predicted output. Machine learning has within health care been successful in e.g. classifying images (11) and calculating risk scores (12–15). Both modelling approaches have different needs regarding data: machine learning models need a lot of data (1000s of subjects), preferably from different sources/demographics to ensure generalizability, and mechanistic models need time-series data for the response to some perturbation of a system, e.g. a meal or a medication.
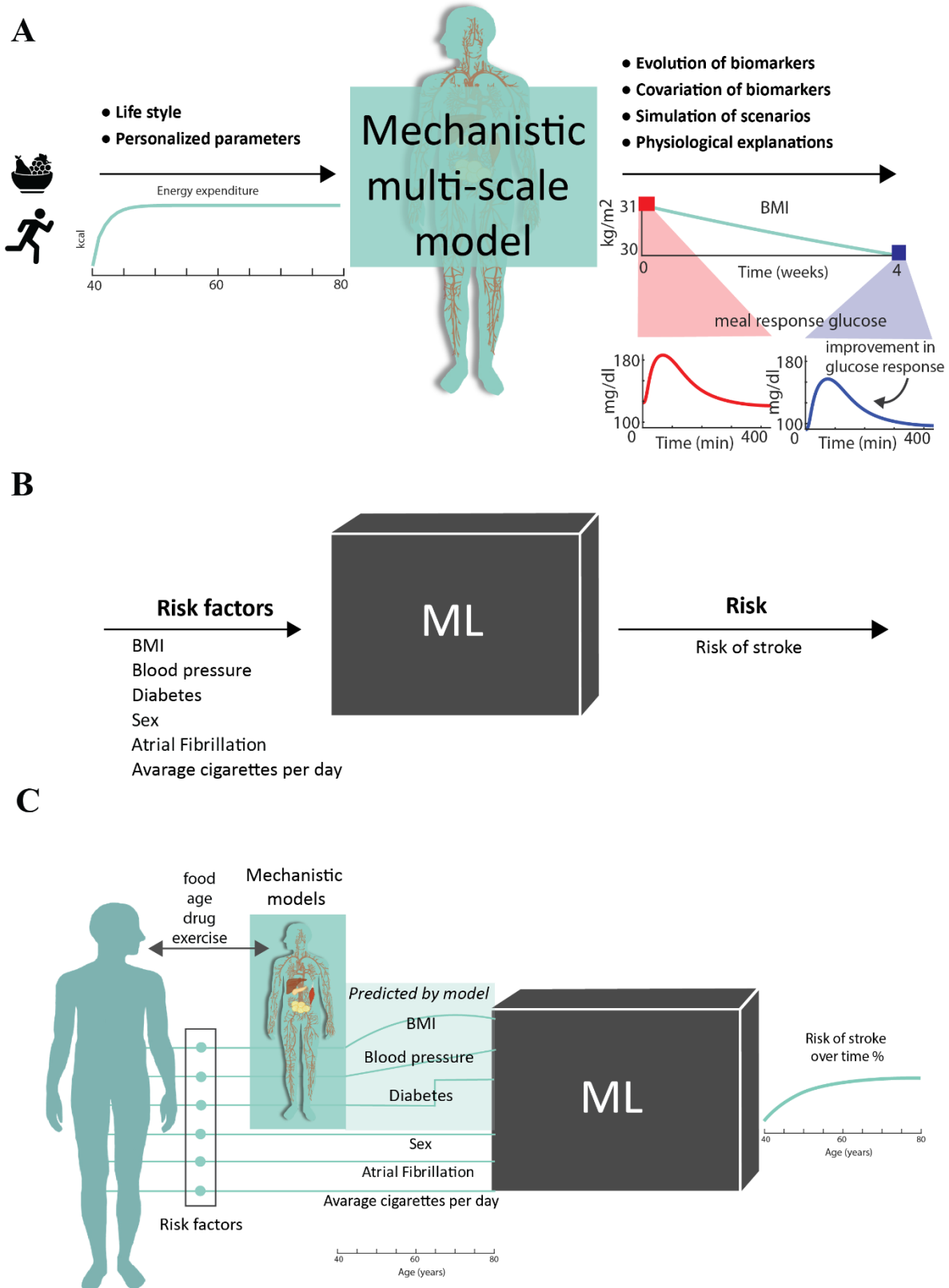
A

- Life style
- Personalized parameters

Energy expenditure

kcal

40   60   80

Mechanistic multi-scale model

- Evolution of biomarkers
- Covariation of biomarkers
- Simulation of scenarios
- Physiological explanations

BMI

kg/m2

31
30

0   Time (weeks)   4

meal response glucose

mg/dl

180

100
0   Time (min)   400

mg/dl

180

100
0   Time (min)   400

improvement in glucose response

B

Risk factors

BMI
Blood pressure
Diabetes
Sex
Atrial Fibrillation
Avarage cigarettes per day

ML

Risk

Risk of stroke

C

food
age
drug
exercise

Mechanistic models

Predicted by model

BMI

Blood pressure

Diabetes

Sex

Atrial Fibrillation

Avarage cigarettes per day

Risk factors

40   60   80
Age (years)

ML

Risk of stroke over time %

40   60   80
Age (years)

Figure 2: Two different modelling approaches, mechanistic and machine learning can be used on their own and in a combined hybrid scheme. A) Mechanistic models can make use of personal parameters and lifestyle data to simulate the evolution of different biomarkers (e.g. BMI and glucose meal response). As such, mechanistic models can be used to simulate different scenarios (e.g. different lifestyles or drugs) and give physiological explanations of these predictions. In this figure, an increase

in energy expenditure is used as input, and the model is used to simulate BMI over 80 years, as well as meal glucose levels at different time points during these 80 years. B) Machine learning models can use different risk factors as input and give an estimated risk score as output. These models are so called black box models, i.e. one cannot look at their model structure and discern any physiological understanding. C) The complementary benefits of the two different modelling approaches can be utilized in a hybrid scheme. Here we suggest and present such a scheme. First, data from an individual are used to personalize the mechanistic model. Second, lifestyle inputs, such as food, age, drug, exercise, are used to make simulations of relevant biomarkers over time. Finally, these simulations, together with other risk factors not simulated by the mechanistic models, are then used to simulate the time-varying risk of stroke.

Since the two modelling approaches have complementary abilities, a combination of the two has sometimes been proposed (10,16–19), even though few examples exist. Such combinations have for instance been used to improve MRI classification (20–22), for predicting tumor growth and density (23), and for predicting phenotypes based on genotypes (24). Hybrid schemas, where ML is used to formulate the mechanistic models, or where mechanistic models are used to constrain ML models, have also been proposed and developed (17,25). For example, ML techniques have been used to train and develop a mechanistic network model that were then used to predict cell responses to combinations of perturbations (e.g. drugs) (26). Within the field of cardiovascular diseases specifically, a hybrid model that combines a discrete model with an agent-based model has been used to simulate treatment procedures of heart failure and comparing both health outcomes and financial aspects (27). However, despite these many suggestions and first examples, and despite the importance of developing digital twins for stroke, no hybrid model underlying digital twins for stroke has yet been developed.

Herein, we present a digital twin for stroke prediction, combining multi-level and multi-timescale mechanistic models with a machine learning model that calculates the 5-year risk of stroke, and show how it can be used in health care. The mechanistic models combine a model for the progression of type 2 diabetes with a new model of how blood pressure changes with age and medication (Figure 3A). We show how this mechanistic model can simulate how different lifestyle choices lead to different scenarios for the evolution of both the overall risk of stroke, and for known physiological risk factors, such as weight, diabetes development, and blood pressure (Figures 9-11). Finally, we discuss outline solutions to practical issues regarding implementation, such as data harmonization, and ethics.

A

Multi-level multi-time scale model of diabetes progression
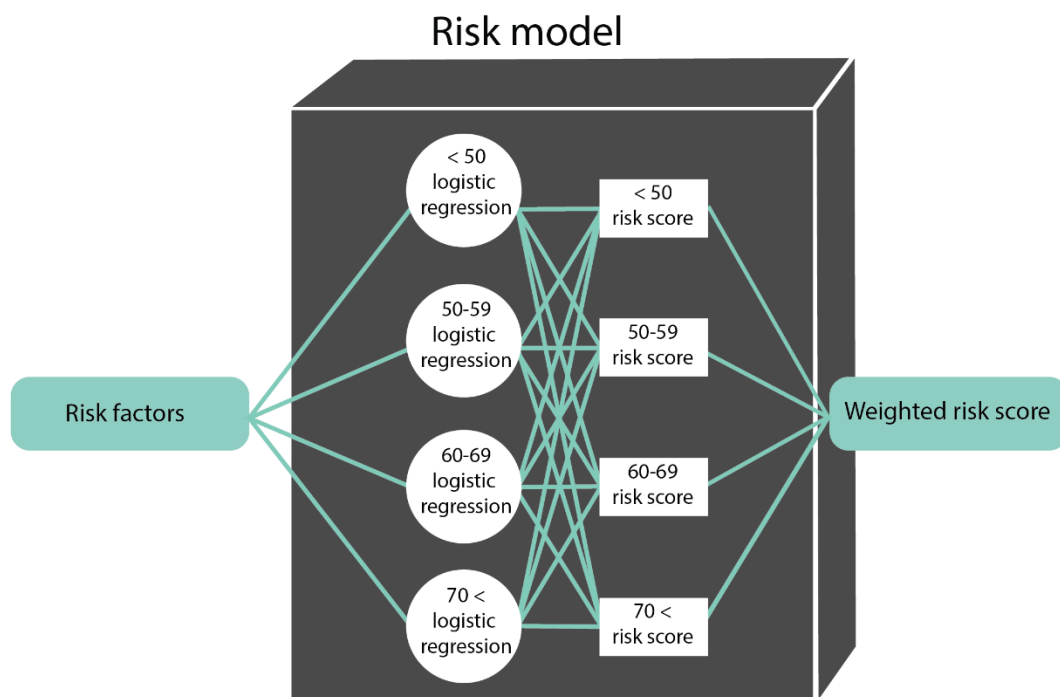

B

Risk model
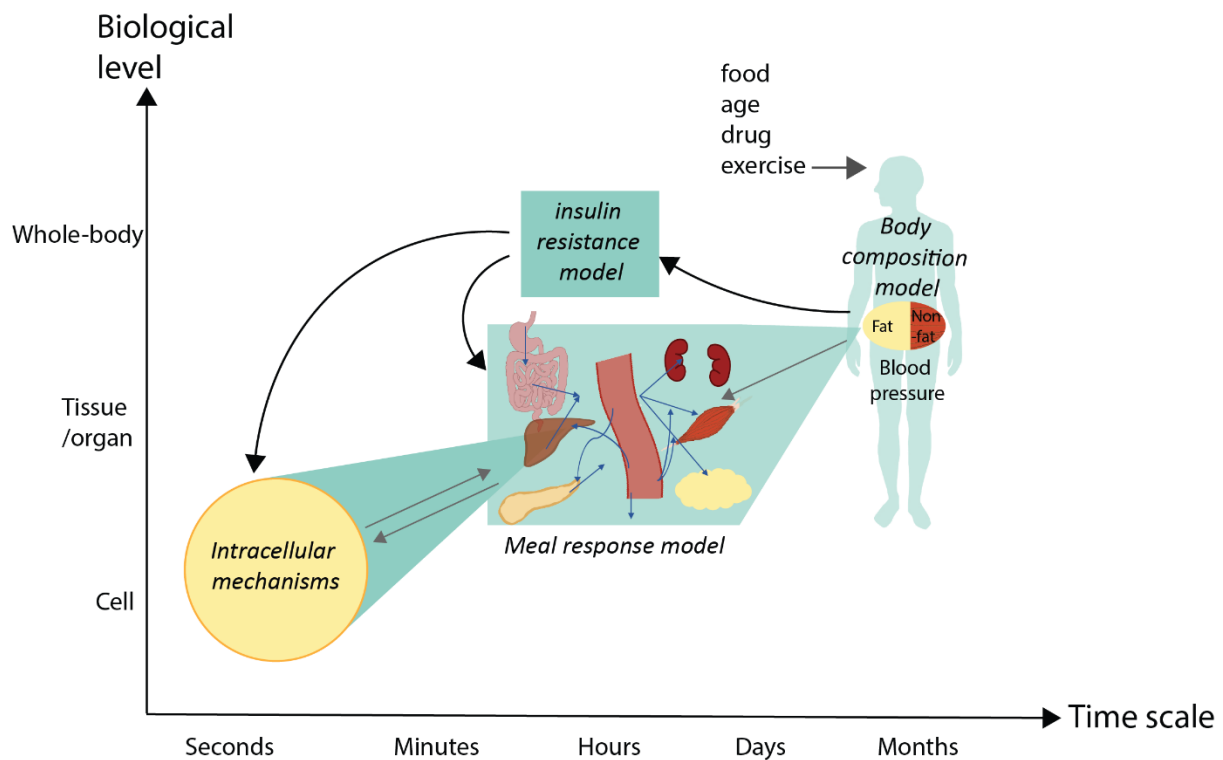
Figure 3: The models included in the herein presented hybrid model. A) The mechanistic part consists of three different biological levels: whole-body, organ/tissue, and cell level. The models can also simulate dynamics on different time scales: on seconds and minutes up to days and months, even years.

B) The risk model consists of an ensemble of 4 different logistic regression models for 4 different age groups, taking in the same risk factors as input. The final risk score is a weighted sum of the risk scores from each of the models.

2 Method

2.1 Mechanistic models

The mechanistic models are built up by ordinary differential equations (ODEs) in the standard form:

$$\frac{dX}{dt} = f(X, t, q, u) \quad (1)$$

$$X(t_0) = X_0(q) \quad (2)$$

$$\dot{y} = g(X, t, q, u) \quad (3)$$

where X is a vector of state variables, f and g are non-linear smooth functions, g a vector of model parameters, u an input signal, $X(t_0)$ initial conditions, and $\dot{y}$ is the simulated model output.

The complete interconnected model consists of four different sub-models describing metabolic control at three different levels: cell, organ/tissue, and whole-body (Figure 3A). All of the equations are given in the Supplementary material (which is based on (28,29)) both as equations and as simulation files. The initial values were obtained through steady state simulation, set to match the scenarios, or kept the same as in original articles. All initial values used in the simulations can be found in the Supplementary material. The fit of the 4 insulin resistance parameters were adjusted by hand, to get a diabetic behavior of relevant variables, corresponding to that of the diabetes parameters in (30). The simulation of the mechanistic model, and numerical optimization of model parameters, were carried out using MATLAB 2018b and the SBtoolbox2 package.

2.2. Parameter estimation in the blood pressure model

The parameter estimation, done only for the new blood pressure model, was done using the traditional weighted least square cost function:

$$v(\theta) = \sum_t \left(\frac{y_t - \hat{y}_t}{SEM_t} \right)^2 \quad (4)$$

where y_t is the measured data at timepoint t, $\hat{y}_t$ the simulation value at time point t, and SEM_t is the standard error of mean at time point t. This cost function is then minimized over the free parameters. Only the parameters of the new blood pressure model were fit to data, and there is no cross-talk between the blood pressure model and the metabolic models. Two datasets were used: Framingham and VALUE. For the Framingham data, the blood pressure parameters were optimized to minimize the cost using two MATLAB functions, first particleswarm and then simulannealbnd. For the VALUE data, the model parameters were manually adjusted to fit to data. Since no standard deviation was given in the Framingham data, the standard deviation was approximated to 3 in the SBP and 2 in the DBP when using a χ^2-test for the training, and standard deviations of 4 for SBP and 2 for DBP in the cross-validation.

2.3 χ2-test

The parameter estimation was evaluated using a χ2-test of the size of the residuals. The null hypothesis for this test is that the experimental data have been generated by the model with additive and normally distributed noise. The test statistic value that is compared with the cost function is given by the inverse cumulative density function,

$$T_{\chi2}^{0} = \mathcal{F}_{\chi2}^{cdf-inv}(1 - \alpha, v) \quad (5)$$

where $T_{\chi2}^{0}$ is the test statistic, $\mathcal{F}_{\chi2}^{cdf-inv}$ is the inverse cumulative density function, α is the significance level (α = 0.05 was used), and v is the degrees of freedom, which was equal to the number of data points in the training dataset. If the model cost is larger than the χ2-threshold, the model is rejected.

2.4 Blood pressure model

The new blood pressure model consists of 2 states:

$$\frac{d}{dt}(SBP) = \left((k1_{SBP} + k2_{SBP} \cdot age) \cdot SBP_{vsIC}\right) - drugS \quad (6)$$

$$\frac{d}{dt}(DBP) = \left((k1_{DBP} - k2_{DBP} \cdot age) \cdot DBP_{vsIC}\right) - drugD \quad (7)$$

$$SBP_{vsIC} = \frac{IC_{SBP} - bSBP}{initSBP - bSBP} \quad (8)$$

$$DBP_{vsIC} = \frac{IC_{DBP} - bDBP}{initDBP - bDBP} \quad (9)$$

Where SBP and DBP are blood pressures, $drugS$ and $drugD$ are the drug effects, $k1_{SBP}$, $k2_{SBP}$, $k1_{DBP}$, $k2_{DBP}$, $k1_{DBP}$, and $k2_{DBP}$ are parameters, SBP_{vsIC} and DBP_{vsIC} are patient-specific constants that depend on the initial blood pressure values at age 30, parameters $bDBP$ and $bSBP$ are parameters that determine the difference in the age-related increase in blood pressure between different patients with different initial blood pressure values, and $initSBP$ and $initDBP$ are constants set to the initial values in the non-hypertensive group 2 in the Framingham data.
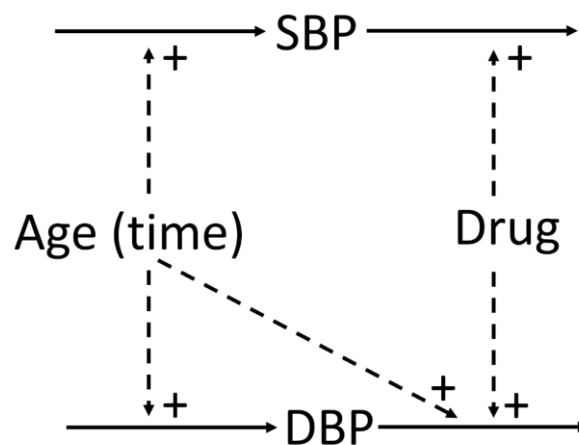

,

Figure 4: Overview of the blood pressure model. SBP is systolic blood pressure, DBP is diastolic blood pressure, and both are affected by both age and by the drug.

2.5 ML risk model

The used machine learning risk model is based on a set of four independent logistic regression models that are designed to predict the 5-year risk of initial ischemic stroke for an individual, belonging to four different age groups, as already presented in (31). There is a logistic regression model for each of the following four age groups: under 50, 50-59, 60-69 and over 70. Each of the four age-specific models has the same risk factors as predictors. The risk factor factors included in the model are: sex, systolic blood pressure (SBP), diastolic blood pressure (DBP), BMI, average cigarettes smoked per day, atrial fibrillation and diabetes. To determine an individual's risk score, their risk factors are used as input into the appropriate model for the individual's age. To use the age models in the hybrid modelling approach, we have adapted the independent set of logistic regression models into an ensemble model as described below.

Age-specific models have been shown to capture the non-proportionality of risk factors by age and are better calibrated to the younger and older populations compared to a model that is trained on all ages with age included as a risk factor. Thus, age-specific models are particularly suited to this hybrid modelling approach that aims to calculate the risk based on simulated scenarios across time, as they will identify and use the risk factors that are most important to the individual's stroke risk at a given time point (31).

However, there are some disadvantages associated with the use of independent age-specific models for the hybrid model. In particular, when the age switches from one age-group to another, there may be a discontinuous jump in the risk, since one switches from one model to another. To avoid this problem, instead of assigning an individual to a given age group and basing their stroke risk solely on the prediction of the age specific model corresponding to their age group, we calculate their risk in each age group in parallel and then take a weighted average of these risks. In other words, we convert 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. To weigh the stroke risk score for an individual returned by each age-group specific model, we use the distance from the individual's age to the corresponding age group of the model that returned the risk score. The function used to calculate the distance d is given in Equation 10 which returns the square root of the sum of the difference between an individual's Age and the minimum value in an $AgeGroup$ squared with the difference between an individual's age and the maximum value in an age group squared.

$$d = \sqrt{(Age - \min AgeGroup)^2 + (Age - \max AgeGroup)^2} \quad (10)$$

Where Age is the age, and $AgeGroup$ is the age group. Within an age group, the smallest value for the distance will be for the age in the middle of the age group and the largest distance will be at the maximum or minimum age. For example, in the 50 to 59 age group, an individual aged 54 or 55 would have the smallest distance to the group while an individual aged 50 or 59 would have the largest. For those not in the age group, the ages closest to the group will have the smallest distance. Someone aged 60 would have a shorter distance to the 50 to 59 age group than someone aged 70. For an individual, the distance to each age group is calculated and then these distances are used to weigh risk scores from corresponding age group specific models and the weighted sum of the risk scores is the individuals total risk score. If we take the example of a 57-year-old person, they will have four distance

parameters, d_{50}, d_{5059}, d_{6069} and d_{70} , and four weights w_{50}, w_{5059}, w_{6069}, w_{70}. These distances would be calculated as follows:

$$d_{50} = \sqrt{(57 - 25)^2 + (57 - 49)^2} = 32.98 \quad (11)$$

$$d_{5059} = \sqrt{(57 - 50)^2 + (57 - 59)^2} = 7.28 \quad (12)$$

$$d_{6069} = \sqrt{(57 - 60)^2 + (57 - 69)^2} = 12.37 \quad (13)$$

$$d_{70} = \sqrt{(57 - 70)^2 + (57 - 96)^2} = 41.11 \quad (14)$$

The weights for each risk model are created using inverse distance weights and are defined below:

$$w_{50} = \frac{\frac{1}{d_{50}}}{\frac{1}{d_{50}} + \frac{1}{d_{5059}} + \frac{1}{d_{6069}} + \frac{1}{d_{70}}} \quad (15)$$

$$w_{5059} = \frac{\frac{1}{d_{5059}}}{\frac{1}{d_{50}} + \frac{1}{d_{5059}} + \frac{1}{d_{6069}} + \frac{1}{d_{70}}} \quad (16)$$

$$w_{6069} = \frac{\frac{1}{d_{6069}}}{\frac{1}{d_{50}} + \frac{1}{d_{5059}} + \frac{1}{d_{6069}} + \frac{1}{d_{70}}} \quad (17)$$

$$w_{70} = \frac{\frac{1}{d_{70}}}{\frac{1}{d_{50}} + \frac{1}{d_{5059}} + \frac{1}{d_{6069}} + \frac{1}{d_{70}}} \quad (18)$$

For the 57 year old, if we plug the distances found in Equations 11 - 14 we get the following weights: w_{50} = 0.11, w_{5059} = 0.50, w_{6069} = 0.30, and w_{70} = 0.09.

To calculate the final risk score we then use the following formula where the variables R_{50}, R_{5059}, R_{6069}, R_{70} are the risk scores calculated by each of the four independent age group specific models:

$$RiskScore = w_{50} * R_{50} + w_{5059} * R_{5059} + w_{6069} * R_{6069} + w_{70} * R_{70} \quad (19)$$

For the 57 year old, we see that the risk score from the 50-59 age model is weighted the highest followed by the risk score from the 60-69 age model; then the under 50 model and finally the over 70 model.

2.6 Code availability

All code will become available, when this manuscript has been accepted in a journal.

3. Results

3.1. Development and validation of the new mechanistic blood pressure model

The blood pressure model describes the effect of aging on systolic and diastolic brachial blood pressure. The model of aging was trained on data from the Framingham Heart Study (32), including four groups with different SBP at the end of the study (Figure 5). The model passes a χ2-test for all 4 groups ($V(\theta)$=32.5 < 111= $\chi^2_{cum,inv}$(88,0.05)). A model validation was performed on the same data by

leaving out one group during model training and then use the left-out group as validation data (Figure 6). The cross-validation in the Framingham data also passes a χ^2-test (V(θ)= 3.03, 0.97, 17,2, 26.5<33.9 = $\chi^2_{cum,inv}$(22,0.05)).

A drug mechanism was added to the model to describe the decline in blood pressure due to anti-hypertensive drugs (Figure 7). The combined long-term model of blood pressure was then trained on data from an angiotensin receptor blocker (ARB)-based treatment from the VALUE trial, where treatment started at the age of 67.2+-8.2 (33). The fit of the trained model to the drug-treatment data passes a χ^2-test (V(θ)=0.8 < 26.3 = $\chi^2_{cum,inv}$(16,0.05)).

The VALUE trial data were also used to choose the timepoint for start of drug treatment used in the hybrid simulation of blood pressure (Figure 10). This timepoint was chosen as the timepoint where the group from the Framingham-data with the highest SBP had the closest SBP to the mean baseline value from the VALUE trial and was found at t=62 years. This choice of timepoint was based on the fact that the person in the simulated scenario is within the age span of the subjects from the VALUE trial (67.2+-8.2). The reduction in blood pressure was assumed to be the same for all values of the blood pressure at the start of treatment. The simulation of normal aging used is the same as was fitted to all four groups from the Framingham study.
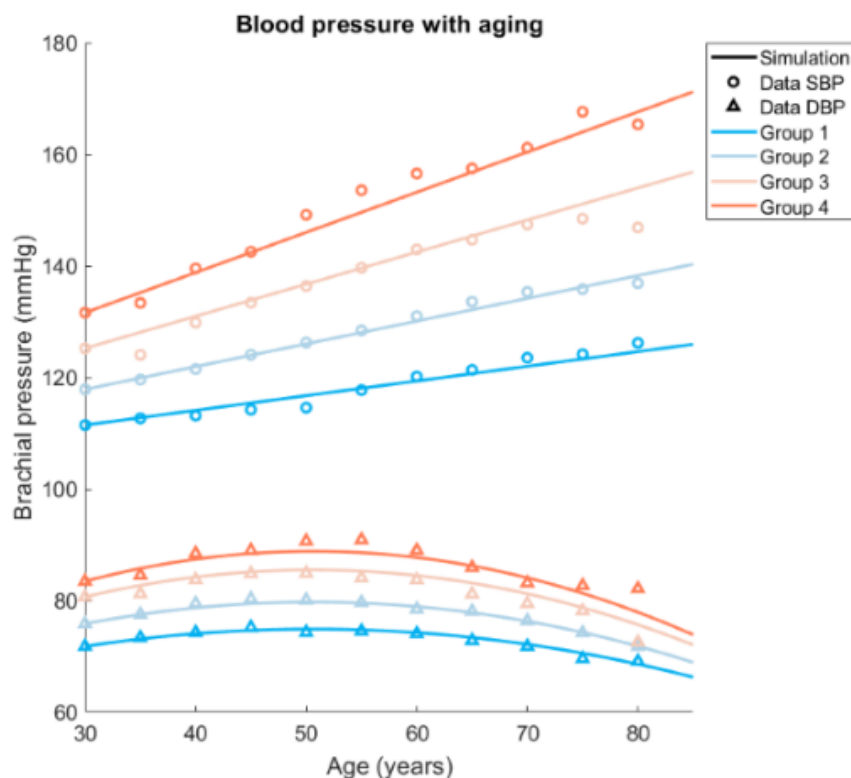

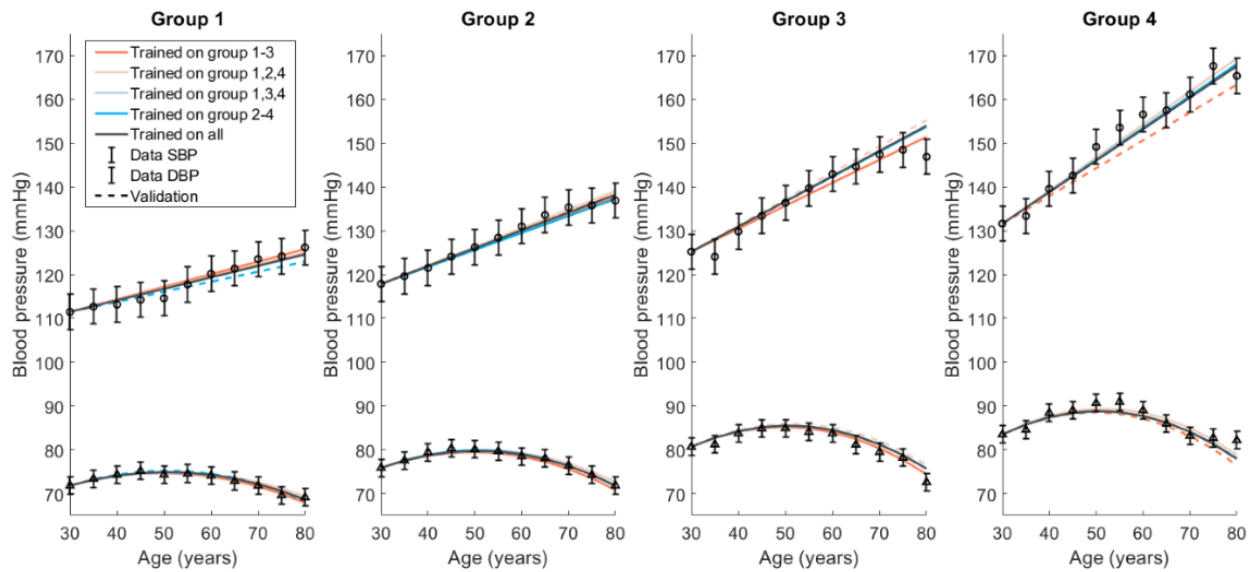*Figure 5: Fit to data of aging from Framingham Heart Study.*

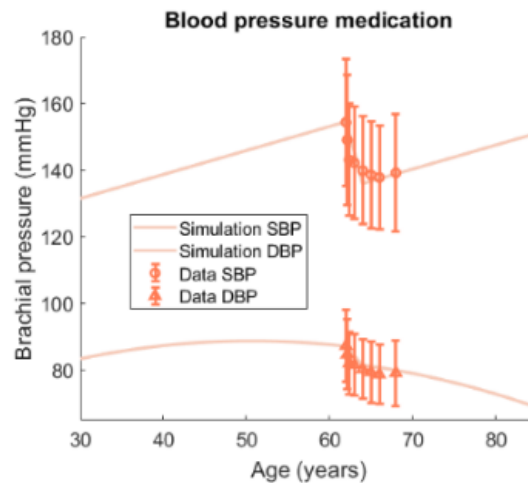*Figure 6: Cross-validation on data from the Framingham Heart Study.*

Figure 7: Fit to data of Valsartan-based treatment, where the final trained model was used to simulate an ARB-based treatment starting at the age of 45 compared to no treatment.

### 3.2. Validation of the ensemble-based risk model

The independent age-specific models, without the ensemble approach, were evaluated in (31). However, since we do a new ensemble-based combination of these age-specific models, it is necessary to look at the evaluation metrics of the ensemble model. We look at discrimination and calibration for the model as a whole and for each of the age groups using an 80:20 train:test split, where we use 80% of each age group data set to train the corresponding age group specific model and then combine the left over 20% of data from each age group to test the weighted model. Looking at both discrimination and calibration is important to fully understand the performance of the model. While the

discrimination metrics provide a measure of the predictive capabilities of the model, the calibration measures the difference between the estimated and true risk, which can help to identify if the model does not predict well for certain groups (34). For discrimination we look at the F1 measure, the AUC, and the accuracy of the model. The F1 measure is the harmonic mean of how often a model makes a positive prediction that is true (precision) and the true positive rate (recall). The AUC is the area under the receiver operating characteristic (ROC) curve which plots the true positive rate against the false positive rate as the prediction threshold of the model moves between 0 and 1. Accuracy is the portion of predictions that are correct (35). To look at calibration, we use the Spiegelhalter's Z-test, which separates the calibration aspects out of the Brier Score (the mean squared error between the outcome and the estimated probabilities), and the Hosmer-Lemeshow test, which examines goodness of fit. If the p-value is significant in either of these two tests, the model is assumed to be not well calibrated.

Using both the discrimination and calibration metrics described above, we compare the ensemble of age-specific models with two models from (31): a set of independent age-specific models (i.e., where the predicted risk for an individual is based solely on the prediction of the age-group specific model that covers their age), and a single model that includes age as a risk factor and that is trained on data from all the age groups. Note that the set of independent age-specific models are, in fact, the component models that are used to construct our ensemble model. In other words, the comparison between these two models directly assesses the contribution of our function for weighting different risks. Figure 8 provides a visualization of the three stroke risk prediction models.
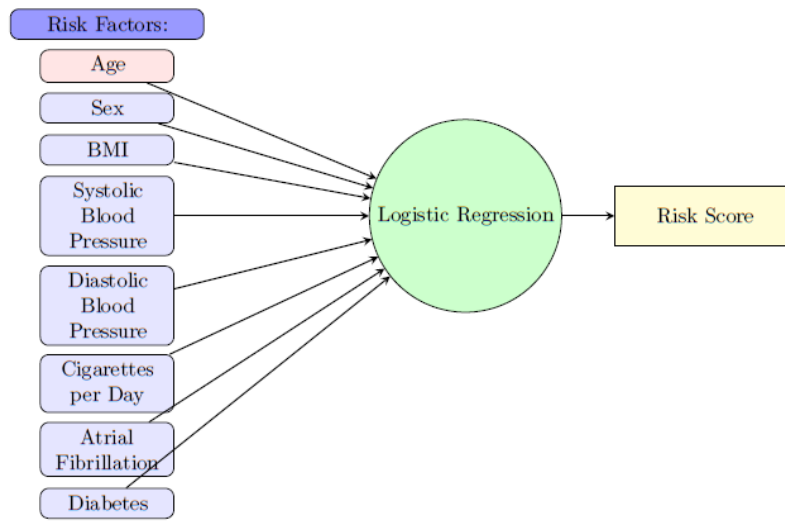
(a) The single model that includes age as a risk factor
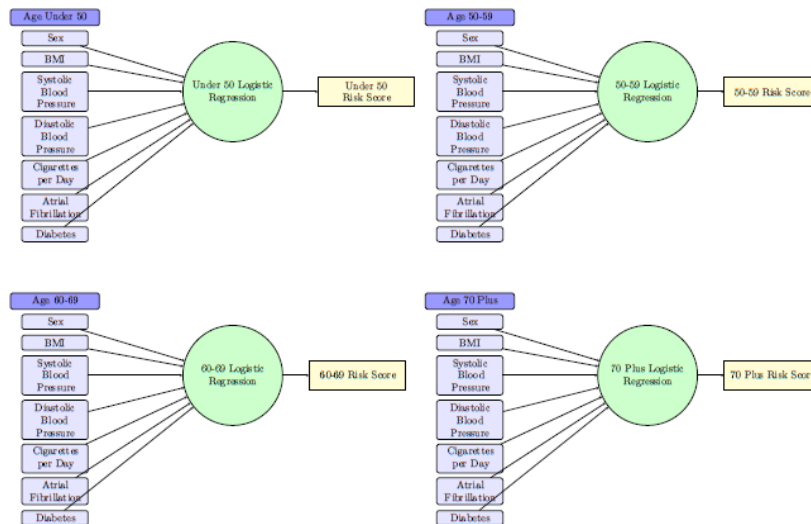
(b) The set of independent age-specific models
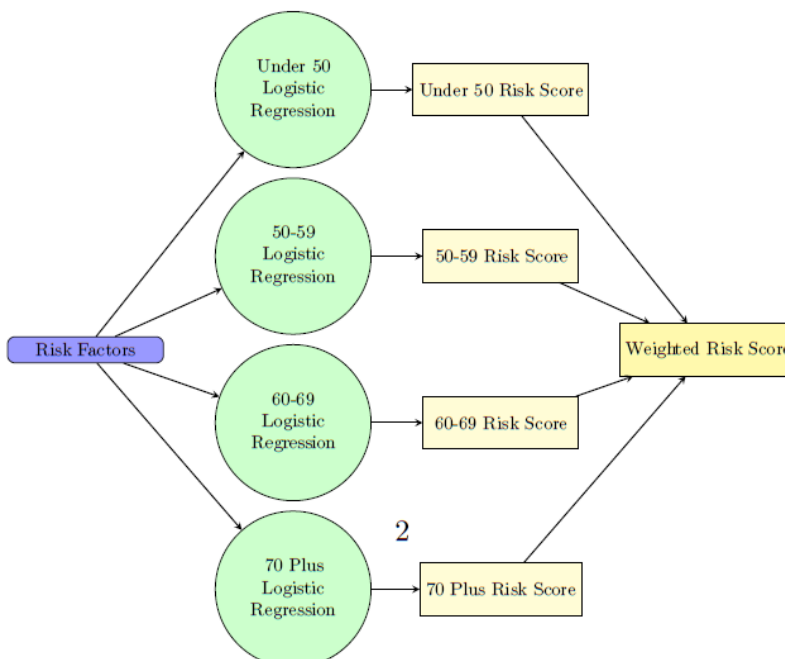
(c) The ensemble model. The risk factors in the ensemble model are the same as those used in the independent age-specific models.

Figure 8: The ensemble model.

3.2.1 Discrimination analysis of the stroke risk model

Table 1 shows the values for AUC, F1, and accuracy for the ensemble of age-specific models tested on the individual age groups, as well as the same measures for the set of independent age-specific models and the single model with age as an input. Comparing the results in Table 1, we see that the set of independent age-group specific models has the highest AUC in all categories and the single model with age as an explicit input (which is representative of a standard approach to stroke risk modeling) has the lowest AUC in all age categories. With respect to F1 the ensemble of the age-specific models has the best performance in all age categories except 60-69, where the independent age specific model has best performance. Finally, looking at accuracy, the independent age-specific models have the best accuracy in the younger age-groups (Under 50, and 50-59). However, in the older age groups the single model with age as an input has slightly better performance and the performance of the ensemble of age specific models either matches or exceeds the performance of the independent age-specific models. Overall, however, we conclude that our new ensemble-approach used to smooth out transitions between the age group model still produces a roughly equal discriminative ability.

Table 1: Discrimination Metrics

	AUC	F1	Accuracy
Under 50			
Ensemble of Age-Group Models	0.57	0.37	0.66
Independent Age-Group Models	0.67	0.22	0.70
Single Model	0.52	0.31	0.64
50-59			
Ensemble of Age-Group Models	0.66	0.45	0.68
Independent Age-Group Models	0.70	0.41	0.71
Single Model	0.64	0.43	0.66
60-69			
Ensemble of Age-Group Models	0.71	0.41	0.68
Independent Age-Group Models	0.72	0.47	0.68
Single Model	0.69	0.37	0.69
70 Plus			
Ensemble of Age-Group Models	0.69	0.49	0.72
Independent Age-Group Models	0.70	0.44	0.71
Single Model	0.69	0.49	0.73

3.2.2 Calibration analysis of the new risk model

Table 2 shows the p-values for the two calibration tests for the ensemble of age-group-specific models, the independent age-group models, and the single model. From the table, we can see that in both the ensemble and independent age models, all age groups have p-values for the goodness of fit tests greater than 0.05, suggesting that these models are well-calibrated across all age groups. This is an improvement from the single-age risk model and suggests that the new ensemble approach does not alter the acceptable calibration property of the previous model (31).

Table 2: Calibration Metrics

	Hosmer-Lemeshow	Spiegelhalter's
Under 50		
Ensemble of Age-Group Models	0.19	0.14
Independent Age-Group Models	0.23	0.34
Single Model	3.45e-5	0.002
50-59		
Ensemble of Age-Group Models	0.17	0.40
Independent Age-Group Models	0.34	0.52
Single Model	0.03	0.09
60-69		
Ensemble of Age-Group Models	0.32	0.54
Independent Age-Group Models	0.32	0.44
Single Model	0.44	0.24
70 Plus		
Ensemble of Age-Group Models	0.38	0.65
Independent Age-Group Models	0.35	0.58
Single Model	0.12	0.38

3.3. Simulation of scenarios with the hybrid model

Five different scenarios are simulated by the hybrid model, and each time two related scenarios are compared. The Five scenarios are summarized in Table 3. In all these scenarios, the person whose variables are simulated is, at the start, a 40-year-old man, with a start weight of 90 kilos, 36% fat, and a height of 185 cm.

Table 3: Risk factors for the 5 different scenarios.

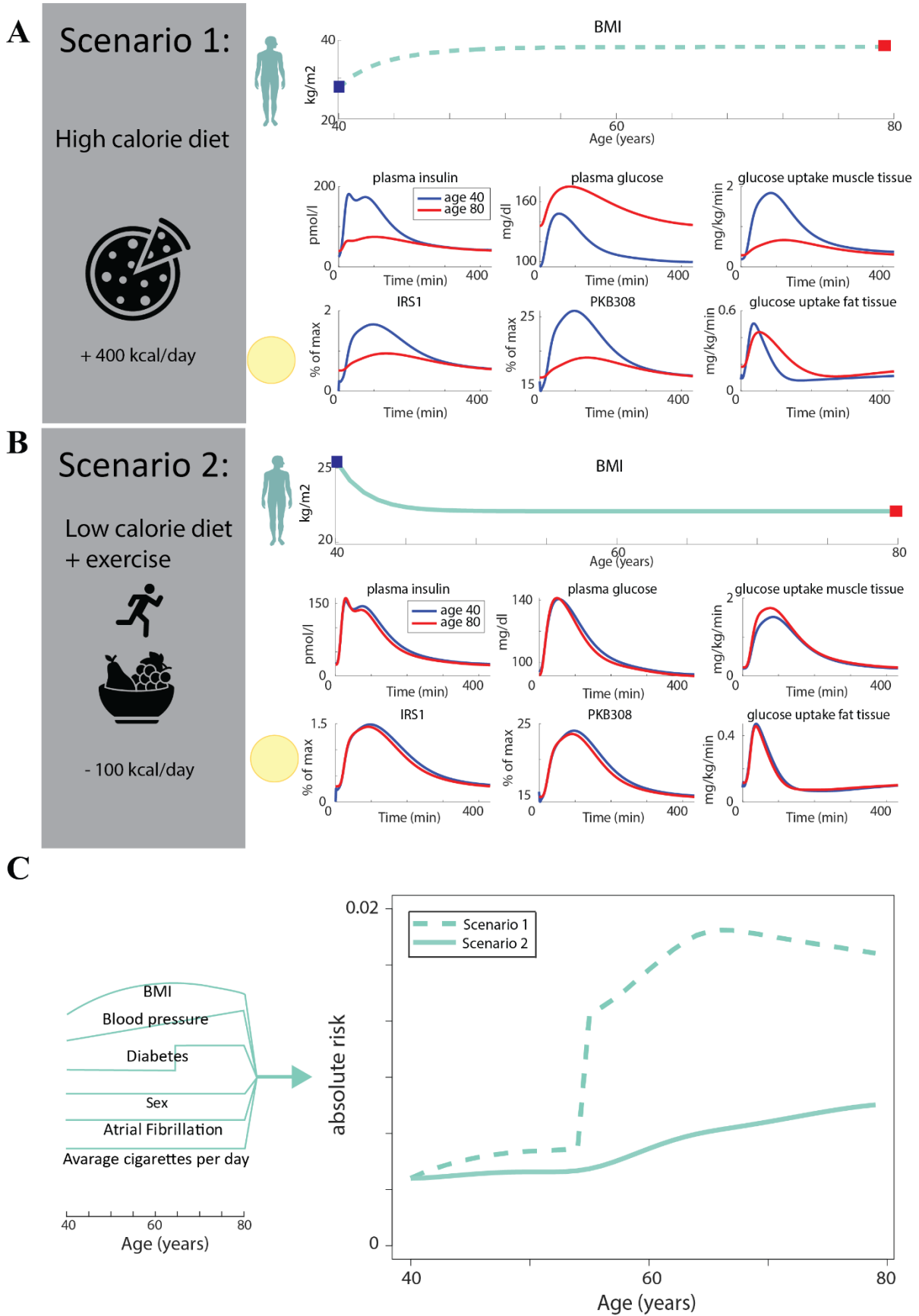|	Weight loss	Weight gain	Blood pressure medication	Diabetes	Smoking	Atrial fibrillation	Gender
Scenario 1		X		X			Male
Scenario 2	X						Male
Scenario 3		X	X				Male
Scenario 4		X		X	X	X	Male
Scenario 5	X		X				Male

3.3.1 Scenario 1 and 2

Figure 9: Simulations and comparisons of scenario 1 and 2. A) Scenario 1: A high-calorie diet, with a daily increase of 400 kcal, and a lack of exercise, leads to an increase of BMI over 40 years. The short-term dynamics at age 80 are here different from the ones at age 40 – plasma insulin has decreased, plasma glucose has increased to diabetic levels, glucose uptake in muscle tissue has decreased, glucose uptake in fat tissue is at the same total amount, but done over a period of time, and insulin-induced signaling IRS1 and PKB308 in isolated fat cells taken from the patient at these two time-points have decreased. B) Scenario 2: A low-calorie diet, decreasing the amount of food per day by 100 kcal, while also increasing the amount of exercise, which leads to a decrease in BMI over the 40 years. The short-term dynamics at age 40, before the weight decrease started, and at age 80, are more or less the same at both the organ/tissue and at the cell level. C) The risk of having a stroke as a function of age for the two scenarios. The risk is constantly higher for scenario 2.

In scenario 1 (Figure 9A), a 40-year-old person that eats approximately 400 extra kcal/day for 40 years, goes from a BMI of around 28, indicating a slight overweight, to a BMI of almost 40, indicating obesity. In contrast, in scenario 2 (Figure 9B), the same person instead decreases their energy intake with 100 kcal/day and therefore has a BMI of 22 after 40 years. When zooming in and comparing the meal response for scenario 1 at age 40 and 80 (Figure 9A), one can see that for example the plasma glucose levels are higher both at start and at its peak, at age 80 compared with age 40. Insulin, on the other hand, is lower at age 80, due to the beta cell collapse, meaning that the person has gone from pre-diabetes to fully developed type 2 diabetes. Another difference that can be seen is that the glucose uptake in muscle tissue has decreased at age 80, but not in the expanded adipose tissue, which is another hallmark of diabetes. When looking at the risks for these two scenarios (Figure 9C), the risk increases slightly more for scenario 1, until age 50, when the diabetes kicks in, and the risk increases prominently. Note that the discrete nature of this jump probably is over-exaggerated, and due to the fact that the risk model is based on discrete diagnosis variable, and not on the more continuously changing physiological variables.
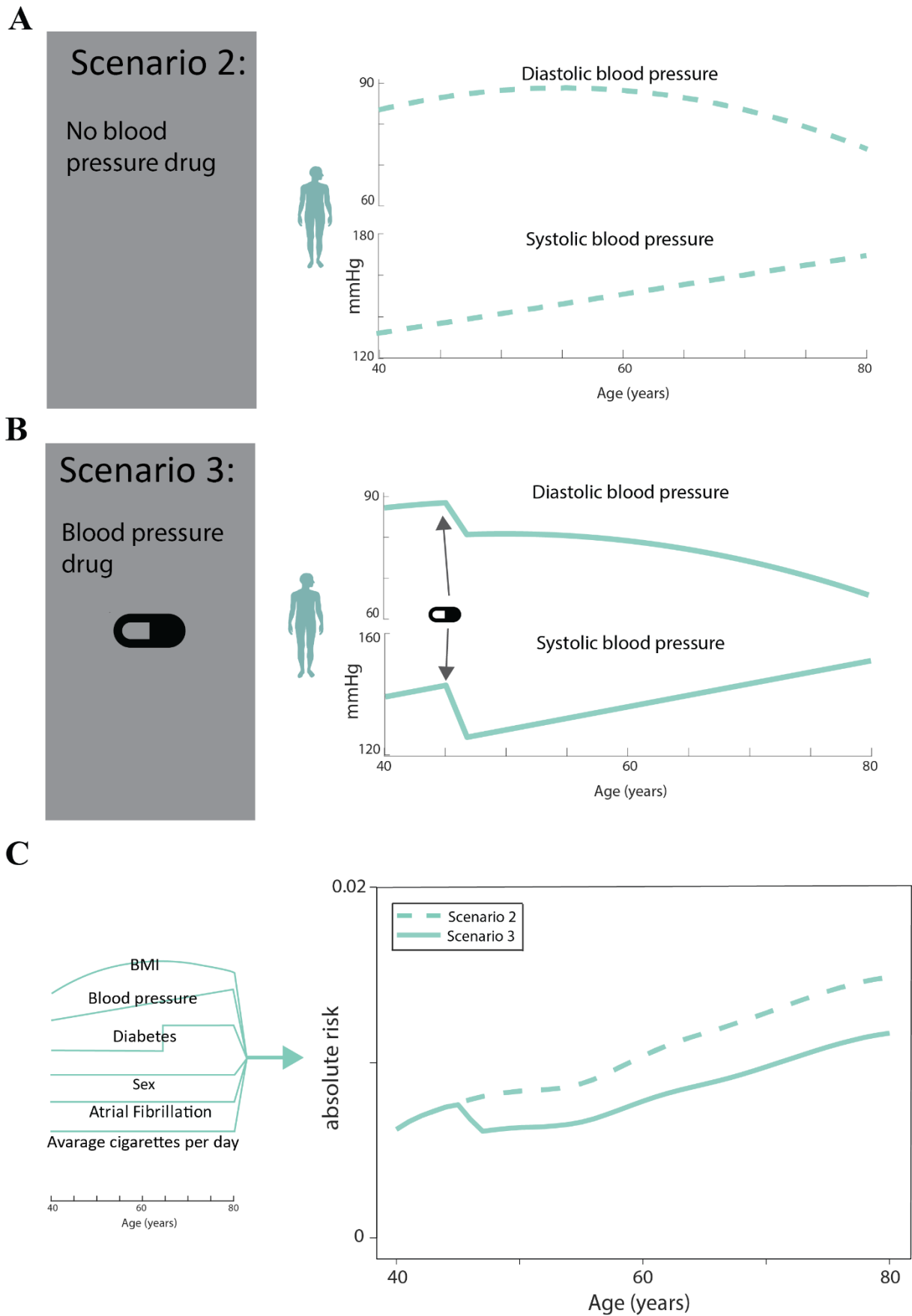
3.3.2 Scenario 2 and 3

Figure 10: Simulations of scenario 2 and 3 compared. A) Scenario 2: No blood pressure drug is taken. B) Scenario 3: Blood pressure increases with age, and at age 50, the person starts taking a blood

pressure drug. C) The risk of having a stroke as a function of age for the two scenarios. The risk is constantly higher for scenario 2, because of the blood pressure medication has taken in scenario 3.

In scenario 3, the person starts taking a blood pressure medication at age 45, consequently decreasing both the SBP and DBP (Figure 10B). This scenario is compared with scenario 2 (Figure 10A), where no medication is taken, and the SBP thus consistently increases with time. Note, however, that the DBP increases at first, and then starts to decrease around age 50. All other risk factors are kept the same for these two scenarios. When comparing the risks for the two scenarios (Figure 10C), one can see a drop in the stroke risk for scenario 3 at the same time that the blood pressure medication is introduced. Note that this reduced risk is maintained throughout the remaining 30 years of simulated time.

3.3.3 Scenario 4 and 5

Let us, finally, compare two more extreme scenarios, to see how much of a difference the lifestyle really can make (Fig 11). In the best-case scenario, the 40-year old reduces weight (as in scenario 1), stops smoking, and also takes a blood pressure medication. Just as in scenario 1, the risk is consistently low, throughout the 40 years of simulated time. In contrast, in scenario 5, the patient gains weight and does not take blood pressure medications, just as in scenario 2. However, here these outcomes are further compounded by the advent of diabetes and atrial fibrillation and continued smoking. As can be seen, the difference in absolute risk is more than 10 times higher in the worst-case scenario, compared to the best case. This shows how big of a difference lifestyle changes can make, and could be the basis of a new digital twin technology, useful in preventive healthcare.
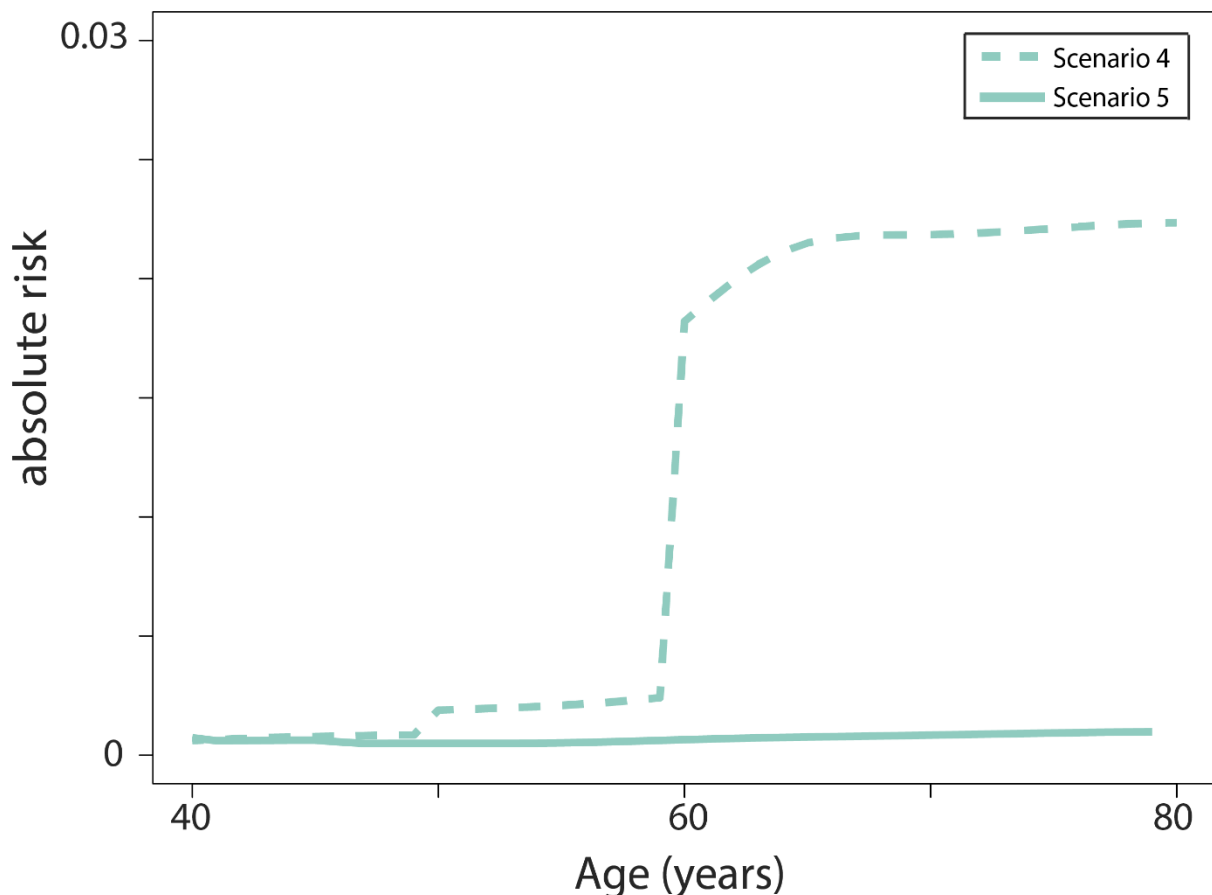

Figure 11: Risk scores for scenario 4 and 5, the worst- and best-case scenarios. In Scenario 4, a weight gain is compounded by continued smoking, diabetes, atrial fibrillation, and the omission of blood

pressure medications. In Scenario 5, weight loss and quitting smoking is combined with blood pressure medication.

3.4 Health informatics harmonization, to help bring the hybrid model into the clinic

If the digital twin models should become a ready-to-use eHealth technology, additional work needs to be done, for example concerning health informatics and end-usage design. There is some, but not much, previous work that focus on data management concerning digital twins (36). Some challenges mentioned include data heterogeneity, the potentially large size of data, the fact that the data is dynamic, i.e. changes over time, and the related data mining problems that these challenges give rise to (36).

Here we focus on the data heterogeneity and the need for semantic harmonization and integration. In the healthcare domain, future digital twins will need to access large amounts of complex heterogeneous data (e.g. omics data, EHR data, wearables data, etc.) which are heterogeneously structured and represented using a wide variety of formats.

There are multiple standards to enable data sharing among systems such as the EHR data exchange standards ISO 13606 (37), HL7 FHIR (38), and openEHR (39) as well as the OHDSI OMOP (40) and i2b2 (41) approaches, which are more oriented to clinical research. Thus, one of the main challenges is to deliver methods that enable the meaningful transformation of data across the existing heterogeneous representation formats. Among others, such methods will need to support the building of federated architectures for digital twins (42,43).

Within the Precise4Q project, where these digital twins for stroke have been developed, we have implemented a semantic-driven framework for semantically harmonizing and integrating data across systems, which facilitates data analysis and exploitation. A data transformation pipeline semantically harmonizes and standardizes the data input to the digital twins. The core of the framework is an ontology-based data model, which uses a top-level ontology to standardize data modelling and ensure interoperability among different ontologies, and the SNOMED CT terminology is used as a reference ontology to represent the clinical concepts.
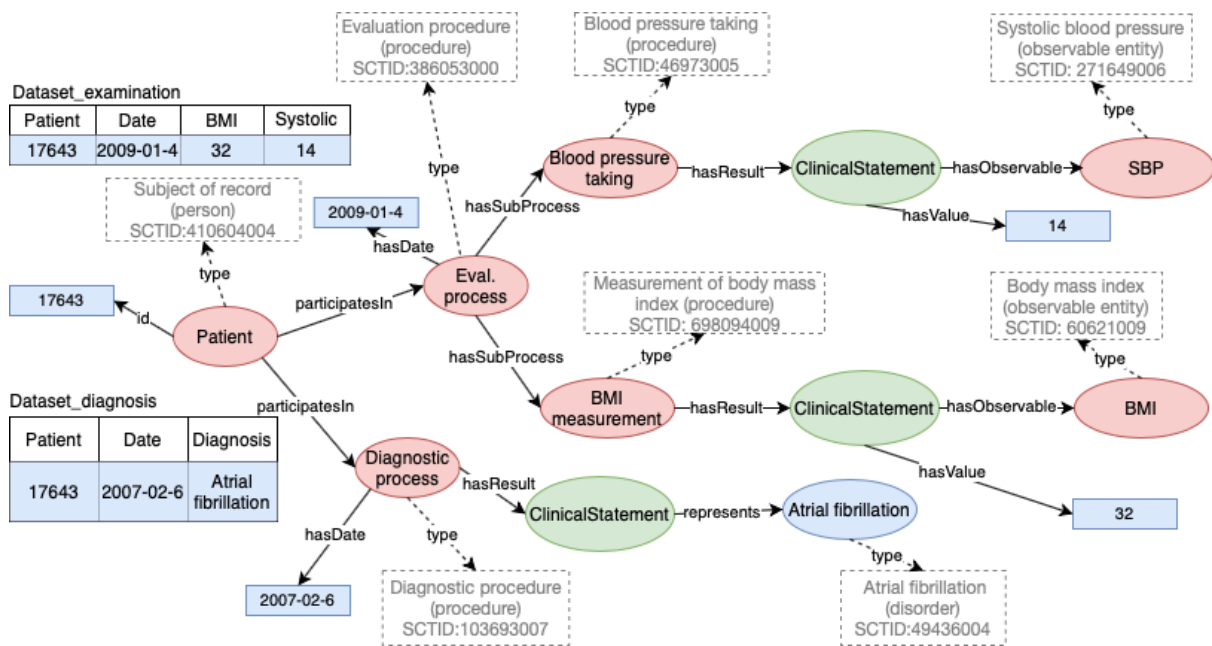
Figure 12: Simplified representation of the result of transforming csv data from two datasets into RDF triples according to the ontology model. Data is shown in blue. SCT concepts are represented as dashed grey squares

Data from the different sources are usually provided as comma-separated values (CSV). We have therefore implemented a data transformation pipeline that takes CSV data as input, transforms it into RDF triples (44) according to the ontology data model, and stores it in a graph-database. Finally, data can be accessed by ML models by using a REST API. Figure 12 depicts the graphical representation of harmonized data from two datasets with patient examination and diagnostic data used in the models herein. Dashed grey squares are SNOMED CT concepts that univocally identify the data elements (e.g. BMI, SPB, Atrial fibrillation, etc.). These variables are related according to the mentioned ontology-based data model. Each data element is always represented together with its contextual information given by the specific process used to acquire the data (e.g. blood pressure taking process at some specific time point). This graph-based structure allows for further exploitation of biomedical knowledge and relations within the heterogenous data, for example for predicting new relations between terms such as genes, diseases, drugs, or symptoms (45,46).

4. Discussion

4.1. Summary

Herein, we have presented the basis for a new digital twin technology, which uses a hybrid combination of a machine learning model with a multi-level, multi-timescale, mechanistic model. This hybrid model can show both the evolution of stroke risk and simulations of physiological variables, such as the progression of diabetes, weight, and blood pressure. The new hybrid model is based on a combination of existing and new models. The multi-level and multi-timescale model describes weight changes and glucose homeostasis (28,29). One of the new models (Figure 4) describes blood pressure as a function of age (Figure 5-7) and medications (Figure 7), and can both agree with estimation data (Figure 5,7), and correctly predict validation data (Figure 6). The second new model is a modification

of the stroke risk model from (31), using a new ensemble approach (Figure 8). This ensemble approach was introduced to smooth out discontinuities, when moving from one age group model to another, and we show that the performance of the new model is as good as the previous model (31), both regarding discrimination (Table 1) and calibration (Table 2). We illustrate how the hybrid model can be used, by simulating and comparing different lifestyle scenarios: for example comparing the health trajectory for an individual who gains weight and develops diabetes (Figure 9, scenario 1) with that of an individual who loses weight and thereby improves their glucose meal response (Figure 9-10, scenario 2); other scenarios considered include an individual who gains weight but take blood pressure medication (Figure 10, scenario 3), an individual who smokes and develops atrial fibrillation (Figure 11, scenario 4), and an individual who loses weight and takes blood pressure medication (Figure 11, scenario 5). Finally, we also discuss and exemplify how health informatics solutions could be used to facilitate the usage of our hybrid model, both in the clinic and at home.

4.2. Strengths of our new models

New blood pressure model

The blood pressure model is new because it can describe both how blood pressure increases with age and how medication lowers the blood pressure. In the literature, there exist several models of the regulation of blood pressure, but only a few of them focus on long-term regulation. Guyton's models created the groundwork for modeling long-term regulation of blood pressure based on simulations and animal data, including detailed descriptions of the salt-fluid balance (47,48). After Guyton's model, many similar models have been developed (49), along with short-term Windkessel- (50) and time-varying elastance (51)-based models with more focus on the arterial system and the heart (52,53). For instance, Maksuti et al used a Windkessel-based model to describe blood pressure changes in aging using some of the Framingham data (54). However, the Maksuti model only describes blood pressure changes during normal aging, thus leaving out the three other groups in the Framingham data (Figure 5-6), and did not include effects of blood pressure-lowering drugs, as in the model presented herein (Figure 7). Apart from these models, only a few models describe the dynamics of long-term changes in systolic as well as diastolic blood pressure, such as aging and effects of anti-hypertensive drugs (49,54). Additionally, many of the previous models are not, or are only sparsely, based on human data, do not consider parameter estimation or validation with independent data, and/or do not combine several effects, such as both drugs and normal and hypertensive aging (49,54).

New ensemble-based risk model

The ensemble-based machine learning model deals with many of the most important risk factors, does so in an age-dependent way, while avoiding discontinuities, when moving from one age group to another. The model enables different risk factors to contribute to the risk scores in different amounts at different stages of life by having different models for different age groups. This is important, because the same risk factors will have different contributions in different stages of life, meaning that it is not sufficient to have age as an independent risk factor, along with the other ones (31). Moreover, the model generates smooths changes in risk score when entering a new age group by using a combined risk score for all age groups. Without this combination, an individual that ages and moves between age groups may display an unrealistic jump, between two risks. For example, a 49-year-old might see a decrease in risk score when they turn 50, even though none of their risk factors have changed. Another reason why this new ensemble-based approach is valuable is that there can be a difference between

a person's chronological age, which would be used to select the appropriate model to predict a person's age, and their biological age. A number of studies have shown that stroke patients tend to have higher biological ages compared to non-stroke patients with the same chronological age (55). For this reason, it may be beneficial to down-play the importance of chronological age, and focus more on the other risk factors, which are a reflection of the biological age. This down-playing of chronological age appears since we use a weighted combination age-specific sub-models, thereby smoothly integrating information from different age cohorts. In these sub-models, age is less important, compared to in all-age models, where age is a more dominant risk factor.

The new hybrid model

The combined hybrid model is the first of its kind for stroke care, since it both can predict the risk of stroke, as well as predict the evolution of risk factors and other physiological variables, ranging from intracellular insulin signaling to organ fluxes, and whole-body variables. This combined ability is possible from our sequential hybrid modelling approach, where the mechanistic model is simulated first, and where that serves as an input to the stroke risk model. This combined ability is important in a clinical context, since one then needs to both see how the risk is expected to change as a function of lifestyle, and understand why these changes are happening, i.e. see the underlying physiological changes (Figure 1). Furthermore, it is also important to be able to simulate on a both long and short timescale, since it is only via short-term simulations (days-months) that one can compare predictions with observed changes and thus gain more faith in the models, and since it is only via long-term simulations (years-decades) that risks are substantially altered. Finally, as mentioned in the introduction, there are some, but not many, hybrid models in biomedicine already available (20–27), but no other model exists for multi-scale simulations of stroke risk.

4.3. Limitations of our models

Multi-level multi-timescale model

The multi-level, multi-timescale mechanistic model used herein constitutes a first step towards a model of stroke progression, but it has as several crucial limitations. Firstly, the weight-driven insulin resistance is constructed in a top-down fashion. The top-level is the weight-change, which is altered by changes to the balance between energy intake and consumption. These top-level weight-changes then affect tissue and cellular processes. However, the short-term dynamics on the lower levels does not propagate up to the top-level. This is a limitation, since long-term changes are nothing else than the cumulative impact of short-term changes, and this means that short-term changes seen during a few days do not impact the long-term predictions. Second, some of the organs and tissues represented are described by highly simplified sub-models. Therefore, in such organs, e.g. muscle, liver and brain, only a few variables can be simulated, such as glucose uptake. There are more detailed models for these organs (56–58), but these models have not yet been integrated into our multi-level model. A third limitation is the lack of other processes related to stroke represented in the mode. Now, the model focuses on the development of insulin resistance and diabetes, which are important in the early etiology of the metabolic syndrome. Future versions of our digital twin models should include mechanisms such as the development of arterial fibrillation, atherosclerosis, and thrombosis.

New blood pressure model

While our new blood pressure model is the first to both describe changes with age and medications, and while it is able to both describe estimation data and correctly predict independent validation data, it is still a simple, isolated, and phenomenological model. First, the new blood pressure model is not based on mechanistic Windkessel models, and can thus not simulate short-term dynamics, such as the blood pressure propagation during a single heartbeat. Second, for the same reason, our model is phenomenological, and predicts the long-term changes in systolic and diastolic blood pressure, without predicting the underlying changes in the mechanisms that are involved in regulating the blood pressure. For instance, the model does not include changes to arterial properties, such as stiffness and elastance, or changes to kidneys, etc. The reason for this limitation is that the required data for such a model - e.g. 4D-flow MRI data collected over long-term longitudinal studies - does not yet exist. Finally, the blood pressure model is not connected to the multi-level, multi-timescale model for weight and glucose homeostasis, even though blood flow has a mutual interplay with e.g. glucose uptake and diabetes (28).

New ensemble-based risk model

The risk model presented herein includes 7 of the most important risk factors for stroke (Figure 4), but there are several other risk factors that are not included. These risk factors include other heart diseases, birth control pills, previous strokes or TIAs, high red blood count, high blood cholesterol and lipids, alcohol and drug use, cardiac structural abnormalities, and social and economic factors. All these risk factors are manageable to some extent at least, and not including them in a risk model used in the clinic could mean that important possible changes to reduce the risk are overlooked. To include more risk factors would, however, also mean that missing values need to be addressed in some way, both during training and during usage of the risk model. The risk factors currently included are all easy to know or measure, and are measured in all known databases, and this is the case also for several other risk factors not included. Such variables would be straightforward to add to our model, and missing values could be handled by using some imputation method. In contrast, some other risk factors are measured only rarely, and are thus not included in all large databases. For instance, cardiac structural abnormalities are diagnosed using echocardiography, which is done in the clinic only if there is a strong suspicion of disease, and not a part the Framingham study, used for our risk models. Addition of such variables would thus require the usage of other databases, combine with usage of a transfer learning approach to merge the insights from analysis of such different databases (59).

The new hybrid model

Our hybrid risk model for stroke is the first of its kind, and is to be considered a first step, with many future improvements possible. The limitations of the different sub-models of course carry over to the combined hybrid model as well. Apart from this, there are several hybrid approaches that are not utilized in our model. The approach that we do use, sequential hybrid modelling, is the simplest one, since you can simulate one model first, and then use that as an input to the second model. A more intricate form of hybrid modelling is called blended modelling, where the different models interact with each other, and thus must be analyzed together. One way to do such blended modelling, would be to use nonlinear mixed-effects modelling (60,61), which is a way of bringing in statistical properties, regarding e.g. the distribution of parameters across a population, into the mechanistic simulations. Another way to do such blended modelling would be to have a sub-model to the big multi-level model that is described using a phenomenological machine-learning model. Finally, another important sequential hybrid combination not pursued herein, would be to use a machine-learning or

bioinformatics network model as an input to the ultimate risk model. Such a model could e.g. be a polygenic risk score model, which takes all the genes and calculate a smaller set of risk factors, which then goes into the ultimate risk model. These different future possibilities are reviewed for the context of stroke in (10).

4.4. Future steps towards potential clinical implementation

The hybrid model presented herein opens up for a future, fully realized, implementation of P4 medicine in health care, i.e. Predictive, Preventive, Personalized and Participatory healthcare (Figure 1). The Predictive part comes through the predictive simulations and risk scores (Figure 9-11, the scenarios); the Preventive part comes from using the predictions in e.g. the preventive health dialogue; Personalization comes through the fact that the models are personalized; and finally, the Participatory part comes from the fact that the simulation empowers patients to make informed decisions about their health behaviors, lifestyle choices, and interventions, not just in connection to the Health Conversation, but continuously throughout life. This last potential was also mentioned in the introduction (Figure 1). Short-term simulations can be compared with data and can thus be used to examine and verify the short-term effects (e.g. glucose meal response) of the long-term changes (e.g. weight changes). This can be done whenever the patient so decides, i.e. after a year, half a year, or even a day. By improving patients' understanding of how different health and lifestyle factors and changes will impact their stroke risk, the digital twin can serve as a valuable tool for patient education and as a conversation aid during the clinical encounter. Being able to simulate how different types of interventions or lack thereof may reduce or increase the risk of suffering a stroke, enables patients to make informed decision about their health at eye level with healthcare professionals. Moreover, if predictions are consistent with the observed changes, this can increase patients' trust in the model, including for the long-term predictions, and as a result increase their motivation and compliance to lifestyle changes.

Despite the great potential of predictive health information to empower individuals, it is important to consider some of the pitfalls and potential challenges patients may face in making effective use of this information. Even though the digital twin can foster awareness of much needed lifestyle changes, it does not help patients to overcome the environmental or social factors that may prevent them from adopting a healthier lifestyle. It is also important to note that not all patients are willing and capable to take on such an active role in managing their own health and may feel burdened by the responsibility of having to do so. This is where healthcare professionals are called upon to support patients in identifying not only the right type of lifestyle intervention but also how it can be achieved.

Finally, when bringing the herein presented hybrid model all the way to an implemented eHealth technology, there are several additional steps needed. For instance, data from different sources need to be integrated and harmonized in a secure way. User interfaces need to be co-designed with end-users, including both patients, clinicians, and other persons that may be involved, such as personal trainers, coaches, relatives, and friends.

5. References

1. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. Personalized Medicine. 2013 Aug;10(6):565–76.

2. United Nations. United Nations General Assembly, Sixty-Sixth Session. [Internet]. High-Level Meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases.; 2011. Available from: https://www.who.int/nmh/events/un_ncd_summit2011/3rd_plenary_meeting.pdf

3. Prepare for the Digital Health Revolution [Internet]. Fortune. [cited 2022 Mar 21]. Available from: https://fortune.com/2017/04/20/digital-health-revolution/

4. Hood L, Balling R, Auffray C. Revolutionizing medicine in the 21st century through systems approaches. Biotechnology Journal. 2012 Aug;7(8):992–1001.

5. Golse N, Joly F, Combari P, Lewin M, Nicolas Q, Audebert C, et al. Predicting the risk of post-hepatectomy portal hypertension using a digital twin: A clinical proof of concept. Journal of Hepatology. 2021 Mar;74(3):661–9.

6. Corral-Acero J, Margara F, Marciniak M, Rodero C, Loncaric F, Feng Y, et al. The 'Digital Twin' to enable the vision of precision cardiology. European Heart Journal. 2020 Dec 21;41(48):4556–64.

7. Shamanna P, Saboo B, Damodharan S, Mohammed J, Mohamed M, Poon T, et al. Reducing HbA1c in Type 2 Diabetes Using Digital Twin Technology-Enabled Precision Nutrition: A Retrospective Analysis. Diabetes Ther. 2020 Nov;11(11):2703–14.

8. Cedersund G, Roll J. Systems biology: model based evaluation and comparison of potential explanations for given biological data: Model based evaluation in systems biology. FEBS Journal. 2009 Feb;276(4):903–22.

9. Shao G, Helu M. Framework for a digital twin in manufacturing: Scope and requirements. Manufacturing Letters. 2020 Apr;24:105–7.

10. Herrgårdh T, Madai VI, Kelleher JD, Magnusson R, Gustafsson M, Milani L, et al. Hybrid modelling for stroke care: Review and suggestions of new approaches for risk assessment and simulation of scenarios. NeuroImage: Clinical. 2021 Jan 1;31:102694.

11. Chan H-P, Samala RK, Hadjiiski LM, Zhou C. Deep Learning in Medical Image Analysis. In: Lee G, Fujita H, editors. Deep Learning in Medical Image Analysis [Internet]. Cham: Springer International Publishing; 2020 [cited 2022 Mar 21]. p. 3–21. (Advances in Experimental Medicine and Biology; vol. 1213). Available from: http://link.springer.com/10.1007/978-3-030-33128-3_1

12. Segar MW, Vaduganathan M, Patel KV, McGuire DK, Butler J, Fonarow GC, et al. Machine Learning to Predict the Risk of Incident Heart Failure Hospitalization Among Patients With Diabetes: The WATCH-DM Risk Score. Diabetes Care. 2019 Sep 13;42(12):2298–306.

13. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLOS ONE. 2017 Apr 4;12(4):e0174944.

14. Alaa AM, Bolton T, Angelantonio ED, Rudd JHF, Schaar M van der. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLOS ONE. 2019 May 15;14(5):e0213653.

15. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak. 2019 Dec 21;19:281.

16. Doyle OM, Tsaneva-Atansaova K, Harte J, Tiffin PA, Tino P, Díaz-Zuccarini V. Bridging Paradigms: Hybrid Mechanistic-Discriminative Predictive Models. IEEE Transactions on Biomedical Engineering. 2013 Mar;60(3):735–42.

17. Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, et al. Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. npj Digit Med. 2019 Nov 25;2(1):1–11.

18. Pathak J, Wikner A, Fussell R, Chandra S, Hunt BR, Girvan M, et al. Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model. Chaos. 2018 Apr 1;28(4):041101.

19. Baker RE, Peña J-M, Jayamohan J, Jérusalem A. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biology Letters. 2018 May 31;14(5):20170660.

20. Bicego M, Ulaş A, Castellani U, Perina A, Murino V, Martins AFT, et al. Combining information theoretic kernels with generative embeddings for classification. Neurocomputing. 2013 Feb 4;101:161–9.

21. Davatzikos C, Shen D, Gur RC, Wu X, Liu D, Fan Y, et al. Whole-Brain Morphometric Study of Schizophrenia Revealing a Spatially Complex Set of Focal Abnormalities. Archives of General Psychiatry. 2005 Nov 1;62(11):1218–27.

22. Batmanghelich NK, Taskar B, Davatzikos C. Generative-Discriminative Basis Learning for Medical Imaging. IEEE Trans Med Imaging. 2012 Jan;31(1):51–69.

23. Gaw N, Hawkins-Daarud A, Hu LS, Yoon H, Wang L, Xu Y, et al. Integration of machine learning and mechanistic models accurately predicts variation in cell density of glioblastoma using multiparametric MRI. Sci Rep. 2019 Jul 11;9:10063.

24. Zhang J, Petersen SD, Radivojevic T, Ramirez A, Pérez-Manríquez A, Abeliuk E, et al. Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. Nat Commun. 2020 Sep 25;11(1):4880.

25. Hughes JH, Keizer RJ. A hybrid machine learning/pharmacokinetic approach outperforms maximum a posteriori Bayesian estimation by selectively flattening model priors. CPT: Pharmacometrics & Systems Pharmacology. 2021;10(10):1150–60.

26. Yuan B, Shen C, Luna A, Korkut A, Marks DS, Ingraham J, et al. CellBox: Interpretable Machine Learning for Perturbation Biology with Application to the Design of Cancer Combination Therapy. Cell Systems. 2021 Feb 17;12(2):128-140.e4.

27. Lassnig A, Rienmueller T, Kramer D, Leodolter W, Baumgartner C, Schroettner J. A novel hybrid modeling approach for the evaluation of integrated care and economic outcome in heart failure treatment. BMC Medical Informatics and Decision Making. 2019 Nov 21;19(1):229.

28. Herrgårdh T, Li H, Nyman E, Cedersund G. An Updated Organ-Based Multi-Level Model for Glucose Homeostasis: Organ Distributions, Timing, and Impact of Blood Flow. Front Physiol. 2021;12:619254.

29. Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, et al. Quantification of the effect of energy imbalance on bodyweight. The Lancet. 2011 Aug;378(9793):826–37.

30. Man CD, Rizza RA, Cobelli C. Meal Simulation Model of the Glucose-Insulin System. IEEE Transactions on Biomedical Engineering. 2007 Oct;54(10):1740–9.

31. Hunter E, Kelleher JD. Age Specific Models to Capture the Change in Risk Factor Contribution by Age to Short Term Primary Ischemic Stroke Risk. Frontiers in Neurology [Internet]. 2022 [cited 2022 Mar 2];13. Available from: https://www.frontiersin.org/article/10.3389/fneur.2022.803749

32. Franklin SS, Gustin W, Wong ND, Larson MG, Weber MA, Kannel WB, et al. Hemodynamic Patterns of Age-Related Changes in Blood Pressure: The Framingham Heart Study. Circulation. 1997 Jul;96(1):308–15.

33. Julius S, Kjeldsen SE, Weber M, Brunner HR, Ekman S, Hansson L, et al. Outcomes in hypertensive patients at high cardiovascular risk treated with regimens based on valsartan or amlodipine: the VALUE randomised trial. The Lancet. 2004 Jun;363(9426):2022–31.

34. Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. Journal of the American Medical Informatics Association. 2020 Apr 1;27(4):621–33.

35. Kelleher JD, Mac Namee B, D'arcy A. Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. MIT press; 2015.

36. Singh S, Shehab E, Higgins N, Fowler K, Reynolds D, Erkoyuncu JA, et al. Data management for developing digital twin ontology model. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture. 2021 Dec;235(14):2323–37.

37. 14:00-17:00. ISO 13606-1:2019 [Internet]. ISO. [cited 2022 Mar 23]. Available from: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/78/67868.html

38. HL7 FHIR standard. FHIR. [Internet]. [cited 2022 Mar 23]. Available from: https://www.hl7.org/fhir/

39. Kalra D, Beale T, Heard S. The openEHR Foundation. Stud Health Technol Inform. 2005;115:153–73.

40. C Blacketer. The common data model [Internet]. The book of OHDSI: observational health data sciences and informatics. [cited 2022 Mar 23]. Available from: https://www.ohdsi.org/data-standardization/the-common-data-model/

41. i2b2: Informatics for Integrating Biology & the Bedside [Internet]. [cited 2022 Mar 23]. Available from: https://www.i2b2.org/

42. Ricci A, Croatti A, Mariani S, Montagna S, Picone M. Web of Digital Twins. ACM Trans Internet Technol. 2022 Feb 3;3507909.

43. D'Amico RD, Addepalli P, Erkoyuncu JA. Is a Top Level Ontology Based Digital Twin the Solution to Human-Machine Interoperability? SSRN Journal [Internet]. 2021 [cited 2022 Mar 21]; Available from: https://www.ssrn.com/abstract=3945058

44. RDF - Semantic Web Standards [Internet]. [cited 2022 Mar 24]. Available from: https://www.w3.org/RDF/

45. Santos A, Colaço AR, Nielsen AB, Niu L, Strauss M, Geyer PE, et al. A knowledge graph to interpret clinical proteomics data. Nat Biotechnol. 2022 Jan 31;1–11.

46. Himmelstein DS, Baranzini SE. Heterogeneous Network Edge Prediction: A Data Integration Approach to Prioritize Disease-Associated Genes. PLOS Computational Biology. 2015 Jul 9;11(7):e1004259.

47. Guyton AC. Long-term arterial pressure control: an analysis from animal experiments and computer and graphic models. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 1990 Nov 1;259(5):R865–77.

48. Guyton AC, Coleman TG, Granger HJ. Circulation: Overall Regulation. Annual Review of Physiology. 1972;34(1):13–44.

49. Hallow KM, Lo A, Beh J, Rodrigo M, Ermakov S, Friedman S, et al. A model-based approach to investigating the pathophysiological mechanisms of hypertension and response to antihypertensive therapies: extending the Guyton model. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2014 May;306(9):R647–62.

50. Sagawa K. Translation of Otto frank's paper "Die Grundform des arteriellen Pulses" zeitschrift für biologie 37: 483–526 (1899). Journal of Molecular and Cellular Cardiology. 1990 Mar;22(3):253–4.

51. Suga H, Sagawa K, Shoukas AA. Load Independence of the Instantaneous Pressure-Volume Ratio of the Canine Left Ventricle and Effects of Epinephrine and Heart Rate on the Ratio. Circulation Research. 1973 Mar;32(3):314–22.

52. Segers P, Stergiopulos N, Westerhof N. Quantification of the Contribution of Cardiac and Arterial Remodeling to Hypertension. Hypertension. 2000 Nov;36(5):760–5.

53. Ursino M. A mathematical model of the carotid baroregulation in pulsating conditions. IEEE Trans Biomed Eng. 1999 Apr;46(4):382–92.

54. Maksuti E, Westerhof N, Westerhof BE, Broomé M, Stergiopulos N. Contribution of the Arterial System and the Heart to Blood Pressure during Normal Aging – A Simulation Study. Minamino T, editor. PLoS ONE. 2016 Jun 24;11(6):e0157493.

55. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo RM, Ois A, Rodríguez-Campello A, et al. Ischemic stroke patients are biologically older than their chronological age. Aging (Albany NY). 2016 Aug 25;8(11):2655–65.

56. Silfvergren O, Simonsson C, Ekstedt M, Lundberg P, Gennemark P, Cedersund G. Digital twin predicting diet response before and after long-term fasting [Internet]. 2021 Nov [cited 2021 Nov 23] p. 2021.11.04.467307. Available from: https://www.biorxiv.org/content/10.1101/2021.11.04.467307v1

57. Sundqvist N, Sten S, Engström M, Cedersund G. Mechanistic model for human brain metabolism and the neurovascular coupling [Internet]. Systems Biology; 2022 Feb [cited 2022 Mar 23]. Available from: http://biorxiv.org/lookup/doi/10.1101/2022.02.15.480629

58. Roy A, Parker RS. Dynamic Modeling of Exercise Effects on Plasma Glucose and Insulin Levels. J Diabetes Sci Technol. 2007 May;1(3):338–47.

59. The Personal Health Train Network | The Personal Health Train [Internet]. [cited 2022 Mar 23]. Available from: https://pht.health-ri.nl/personal-health-train-network

60. Mixed-Effects Models in S and S-PLUS [Internet]. New York: Springer-Verlag; 2000 [cited 2022 Mar 23]. (Statistics and Computing). Available from: http://link.springer.com/10.1007/b98882

61. Karlsson M, Janzén DLI, Durrieu L, Colman-Lerner A, Kjellsson MC, Cedersund G. Nonlinear mixed-effects modelling for single cell estimation: when, why, and how to use it. BMC Systems Biology. 2015 Sep 4;9(1):52.