MISSION-T2D

Multiscale Immune System SImulator for the Onset of Type 2 Diabetes integrating genetic, metabolic and nutritional data

Work Package 4

Deliverable 4.4

Validation and refinement of the models in the overall workflow

Document Information

Grant Agreement	Nº	600803	Acronym	MISSION-T2D
Full Title	Multiscale Immune System SImulator for the Onset of Type 2 Diabetes integrating genetic, metabolic and nutritional data			
Project URL	http://www.mission-t2d.eu			
EU Project Officer	Name	Dr. Adina Ratoi		

Deliverable	No	4.4	Title	Report on the validation and refinement of the models in the overall workflow
Work package	No	4	Title	Metabolic data provision and modelling of aggregated metabolic and inflammatory processes

Date of delivery	Contractual	29.02.2016	Actual	22.03.2016
Status	Version 1.4		Final	X

Nature	Prototype		Report	X	Dissemination		Other	

Dissemination level	Consortium+EU	X
	Public	

Target Group (If Public)		Society (in general)	
Specialized research communities		Health care enterprises	
Health care professionals		Citizens and Public Authorities	

Responsible Author	Name	Albert de Graaf	Partner	TNO
	Email	albert.degraaf@tno.nl		

Version Log

Issue Date	Version	Author (Name)	Partner
29.02.2016	1.0	Albert de Graaf	TNO
07.03.2016	1.1	Filippo Castiglione	CNR
17.03.2016	1.2	Albert de Graaf	TNO
18.03.2016	1.3	Filippo Castiglione	CNR
22.03.2016	1.4	Albert de Graaf	TNO

Executive Summary	This document describes the work done in Task 4.5 Validation and refinement of the models in the overall workflow. In close cooperation of partners TNO and CNR, the predictions of the lower aggregation metabolic model embedded in the overall workflow were assessed by comparing with datasets collected throughout the project in WP7. As a consequence, the low level metabolic model was refined especially to correctly model the action of insulin on metabolic pathway activities across different tissues so as to yield realistic simulated time profiles of extracellular metabolites during feeding. The MT2D-Marvel model (low level of detail, month - 6 year time scale) was simplified and refined based on the available Whitehall II cohort data. Initial model subdomain analyses gave satisfactory results but challenges for overall validation remain. Concluding, the metabolic model embedded in the MISSION-T2D simulator (high level of detail, 1 minute - 1 year time scale) was refined and partially validated and as a result can successfully simulate multi-organ metabolism dynamics for various lifestyle scenarios differing in food intake and exercise frequency and -intensity.
Keywords	Model refinement, model validation, metabolism, inflammation

Contents

1 Deliverable description

Deliverable 4.4 is the result of Task 4.5: Validation and refinement of the models in the overall workflow.

We performed simulations of different lifestyles with the Dynamic-E-MF (low aggregation level) model embedded in the overall workflow and the MF-HOMA (high aggregation level) model to validate the model predictions and to discover the relative importance of factors governing the intervention response on a personalized (i.e., sub-group-specific) basis. A subset of immunological and metabolic measurement variables was selected for incorporation in the e-coaching applications (WP8). Section 2.1. re-addresses the issue of the different timescales of the MF-HOMA (high aggregation) model and the Dynamic-E-MF (low aggregation) model. In section 2.2, the results with regard to the integrated Dynamic E-MF model (detailed metabolic-inflammatory mechanisms, 1 day - 12 month time scale) are described. In section 2.3, the results with regard to the integrated MF-HOMA model (high aggregation, multi-year time scale) are described. In section 2.4, we briefly mention a newly funded H2020 project that will provide for a follow-up for the work performed in the MISSION-T2D project and the models developed for the purpose.

2 Deliverable Results

2.1 Validation and refinement of the MISSION-T2D models at different timescales in the overall workflow

2.1.1 Re-iteration of the issue of the different time scales of the models

As foreseen in the Description of Work for WP4, and described in D4.1 and 4.2, it was anticipated to develop an integrated multi-scale model by developing and integrating models at two different aggregation levels and time scales. The short time scale, low aggregation level model (Dynamic E-MF model) corresponds to the challenge test time scale, i.e., minutes-to-days, and contains detail on metabolic pathways, inflammatory processes and their interactions, whereas the long time scale, high aggregation level (MF-HOMA) model would cover the week-to-months timescale and contain little mechanistic details. In practice, the time scales and the aggregation levels of the models have developed into somewhat different ranges. The integrated model ultimately constructed in WP6 by partner CNR in collaboration with most of the partners of the present project, integrates the immune system simulator originally from CNR, the metabolic model contributed by TNO (see Deliverable 4.1), and the models of physical activity, beta cell function and microbiota function contributed by partners UniCAM and UniBO respectively. This model is henceforth called the MISSION-T2D simulator. To all practical

purpose the MISSION-T2D simulator covers the time scale of minutes to max. 1 year. The MF-HOMA model originated at TNO and was built in the Marvel software tool (see Deliverable 4.2). In practice, this model is rather suited to create a forecast for a much longer time frame, i.e., months to as much as 6 years. Therefore, in the previous reporting period (Deliverable 4.3) the question was addressed if and how to integrate both models, and it was decided to keep both models and use the MF-HOMA model operated in a separate fashion, exclusively for the longer term forecast simulations. This long-term model is henceforth called the MT2D-Marvel model to avoid confusion with the MISSION-T2D simulator.

Keeping the low-aggregation and high-aggregation models separate implies that the models may be run more or less in a separate and independent fashion eventually facilitating cross-checking and sharing of ideas. As pointed out in the previous reporting period (Deliverable 4.3), to connect the models at the different time scales with each other, the MISSION-T2D simulator integrating the mechanisms of inflammation, metabolism, beta cell function etc., can be used in practice to run simulations for different *pre-diabetes subgroups* to derive personalised parameter values (strengths and speed) for the interactions between high-level variables in the MT2D-Marvel model, since these are expected to be different for different subgroups. On the other hand, long-term predictions by the MT2D-Marvel model could provide *anchoring points* for the MISSION-T2D simulator enabling also the more detailed predictions for the longer term. These aspects, requiring further exploration by Partners TNO and CNR, could not be addressed in the actual reporting period.

As per the choice made, the MISSION-T2D simulator is especially suited for validation by typical randomised controlled intervention studies (e.g., several weeks/months of high-fat diet versus low fat diet), whereas the MT2D-Marvel model is more suited for validation by long-running cohort studies such as, e.g., the Whitehall II cohort. The datasets collected in WP7 (Deliverable 7.2) will thus provide good and ample data for validation.

2.1.2 Integration of the metabolic and immune system models in the overall workflow

Work on WP4 in the present reporting period has mainly focused on the validation and refinement of the metabolic model embedded in the integrated workflow of the MISSION-T2D simulator as described in section 2.2. The report on integration of the models in the overall workflow has been already provided by CNR in deliverable D6.2 whereas details about the validation of the overall integrated architecture will be provided in deliverable D6.3 at PM36.

In section 2.3, a short account of progress on the MT2D-Marvel model will be given.

2.2 Progress on the MISSION-T2D simulator

2.2.1 Introduction

This section describes the activities in the past period to validate and refine the low aggregation level metabolic model (Dynamic E-MF model) embedded in the overall workflow of the MISSION-T2D simulator.

Figure 2.2.1 shows this overall workflow.

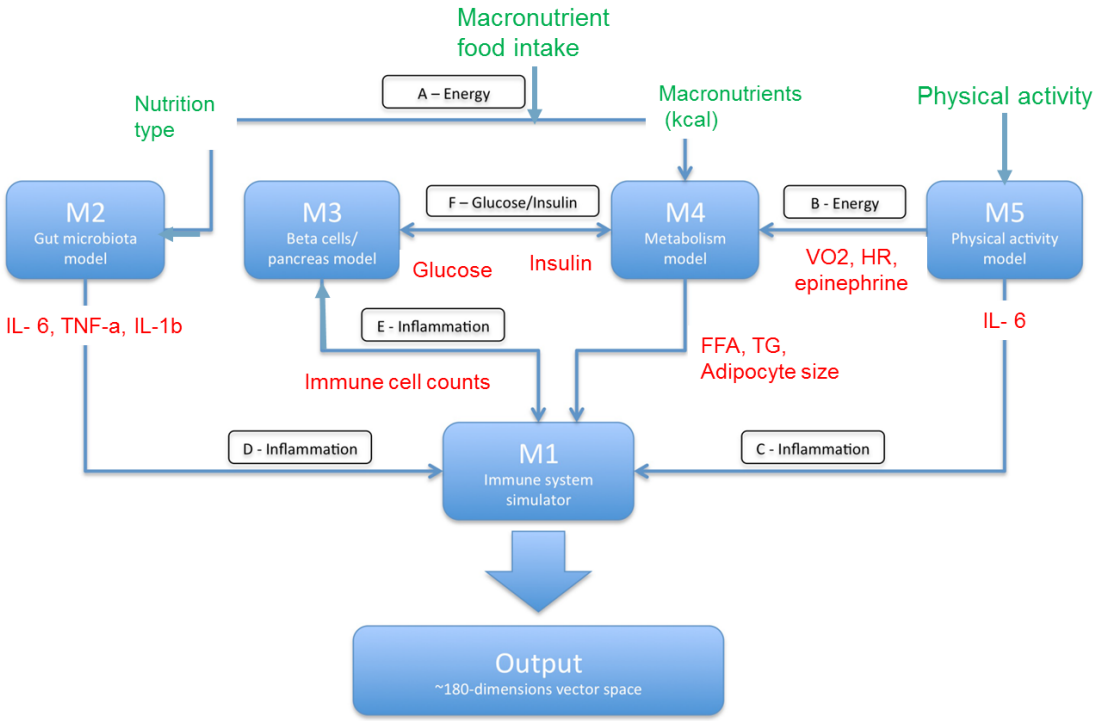

Figure 2.2.1. The module corresponding to the metabolic model (M4) is embedded in the overall workflow of the MISSION-T2D simulator. Symbols in red denote variables exchanged between the different modules.

The work focused primarily on validating the metabolic model predictions for various lifestyle intervention scenarios differing in calorie intake and physical activity. This required the implementation of various modifications to the model equations implemented in the model simulation software C code. To enable efficient and rapid interaction, partner TNO visited partner CNR for a week-long model validation and refinement session in week 40 of year 2015 during which much progress was made.

The interpretation and validation of simulated time courses for the 9 exchangeable and 13 non-exchangeable metabolites across 6 different tissues and an arterial blood compartment proved to be quite challenging, firstly by the sheer amount of information to consider, and secondly because of the fact that metabolic changes in all compartments occurred in a close interaction

with each other, which made it difficult to unambiguously identify the cause of any given inaccuracy of the predictions. Tools for rapid inspection of simulated time profiles had to be developed. Some relevant aspects of the work are described in the following sections 2.2.2 – 2.2.6.

2.2.2 Inspection of metabolic profiles

The metabolic model simulates time trajectories of 9 exchangeable metabolites and 13 non-exchangeable metabolites across 6 different tissue compartments and an arterial blood compartment. In the original model by Kim et al. (2007), the concentration changes of these metabolites are governed by the activities of 13-21 different metabolic fluxes in each tissue compartment (some fluxes are tissue-specific, e.g., liver has 21 fluxes; adipose 13). The corresponding reactions are modelled as enzyme kinetic equations with a total of 95 Vmax and more than 130 Km parameter values.

A total of 20 parameters are used to model the hormonal control, in addition to 14 modulation factors for fluxes regulated by the glucagon/insulin ratio (GIR), and 20 modulation factors for fluxes regulated by epinephrine. Furthermore, the metabolite dynamics are influenced by tissue weights and tissue bloodstream, with the latter varying dependent on work rate. The metabolite exchange between tissues and bloodstream (assumed to occur according to the perfect mixing assumption) is described using 47 different partition coefficients.

With this many parameters and complex regulatory phenomena, further complicated by the high degree of interaction across tissues, identification of causes of possible wrongly predicted metabolite time courses proved challenging. This section shows snapshots of the output of the simulation via scripts developed using the GNUplot software to help with the visualization of this complex task.

First, a visualization routine was developed that plots the concentrations of the exchangeable metabolites glucose (GLC), pyruvate (PYR), lactate (LAC), alanine (ALA), glycerol (GLR), free fatty acids (FA), and triglycerides (TG) in arterial blood or any desired tissue compartment. The plot is output to the screen and a screenshot can be rapidly imported into PowerPoint. Examples are shown in Figure 2.2.2.a and 2.2.2.b. This allows to quickly identify the tissue of origin in case of an abnormal arterial blood concentration profile of a given metabolite, by inspecting the tissue/blood concentration ratio, while taking into account the tissue-specific partition coefficient of that metabolite.

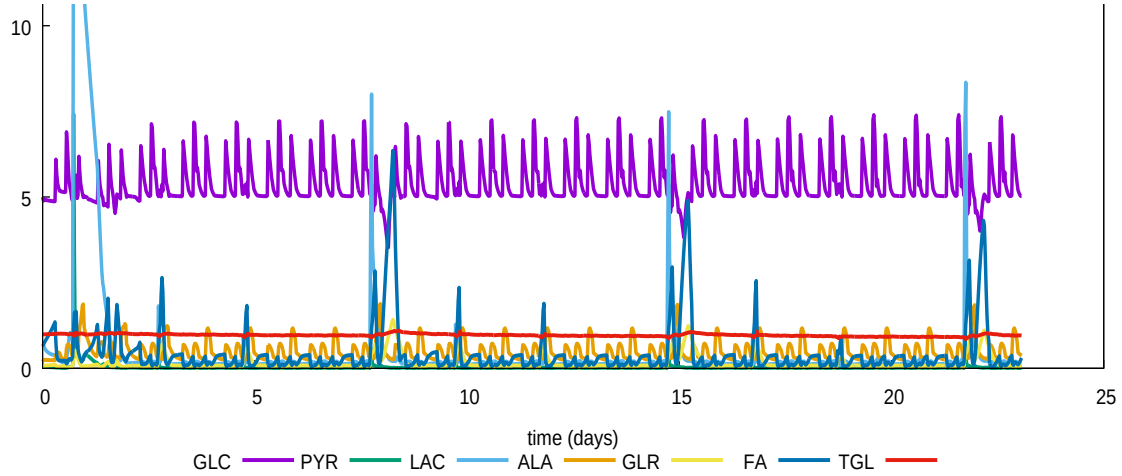

Figure 2.2.2.a. Example of a clipped screenshot showing simulated profiles of exchangeable metabolites in the arterial blood compartment, for a scenario with very low calorie diet (3 meals a day) and periodic exercise. The plot command file selectively exports profiles of glucose (GLC), pyruvate (PYR), lactate (LAC), alanine (ALA), glycerol (GLR), free fatty acids (FA), and triglycerides (TG).

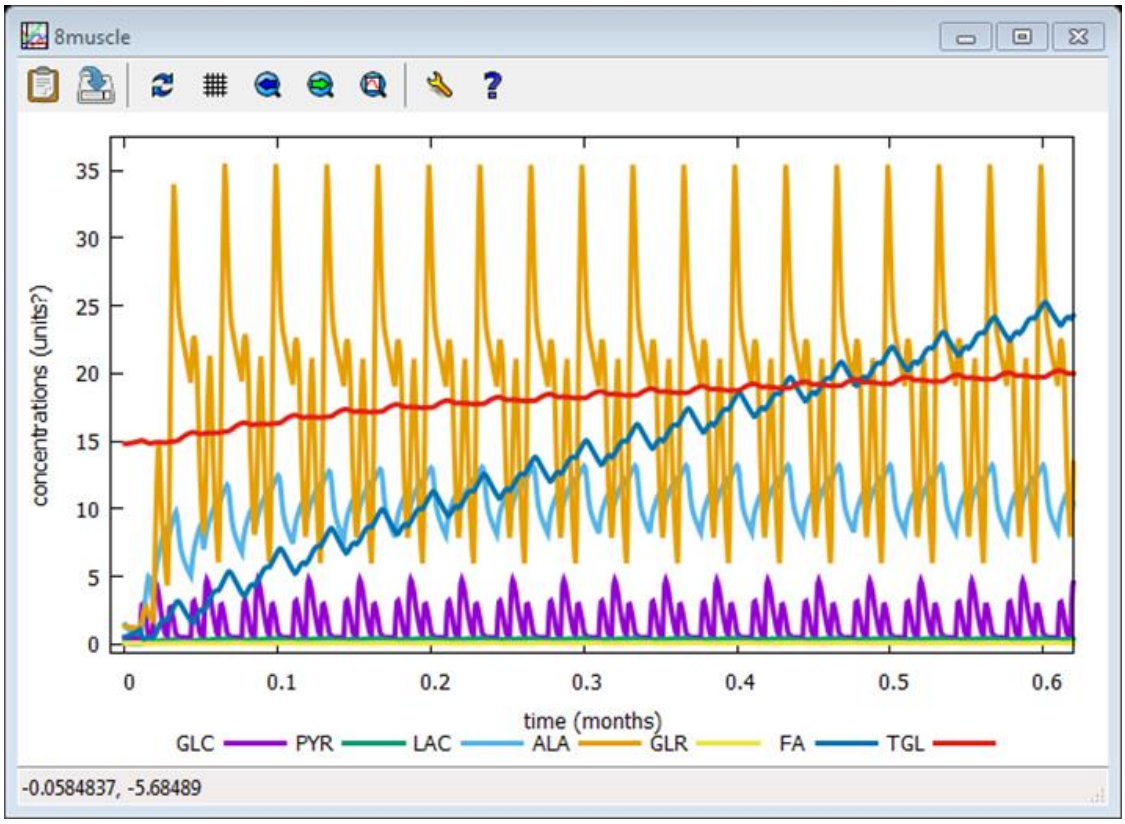

Figure 2.2.2.b. Example of a screenshot showing simulated profiles of exchangeable metabolites in the muscle compartment, for a scenario with high calorie diet (3 meals a day) without exercise. The plot command file selectively exports profiles of glucose (GLC), pyruvate (PYR), lactate (LAC), alanine (ALA), glycerol (GLR), free fatty acids (FA), and triglycerides (TG).

Secondly, a visualization routine was developed that plots the concentrations of the intracellular metabolites glucose-6-phosphate (G6P), glycerol phosphate (GRP), glyceraldehyde-3-phosphate (GAP), NAD+, NADH, ADP, ATP, Pi, creatine (CR), phosphocreatine (PCR), glycogen (GLY), Acetyl-coenzymeA (AcoA) and coenzymeA (CoA) for any desired tissue compartment. An example is shown in Figure 2.2.3.a and b. This allows for detailed inspections e.g., diagnosing a case where an intracellular metabolite concentration gets very close to zero, causing problems and eventual breakdown of the numerical integration routine. If such a case is diagnosed, it is typically a matter of a bad flux balance for that metabolite which requires, e.g., refinement of the Vmax parameter value for one or more of the enzyme reactions that produce or consume that metabolite. Once such a refinement has been made, however, simulations for all scenarios should be re-checked to validate that the refinement does not affect the accuracy of these.

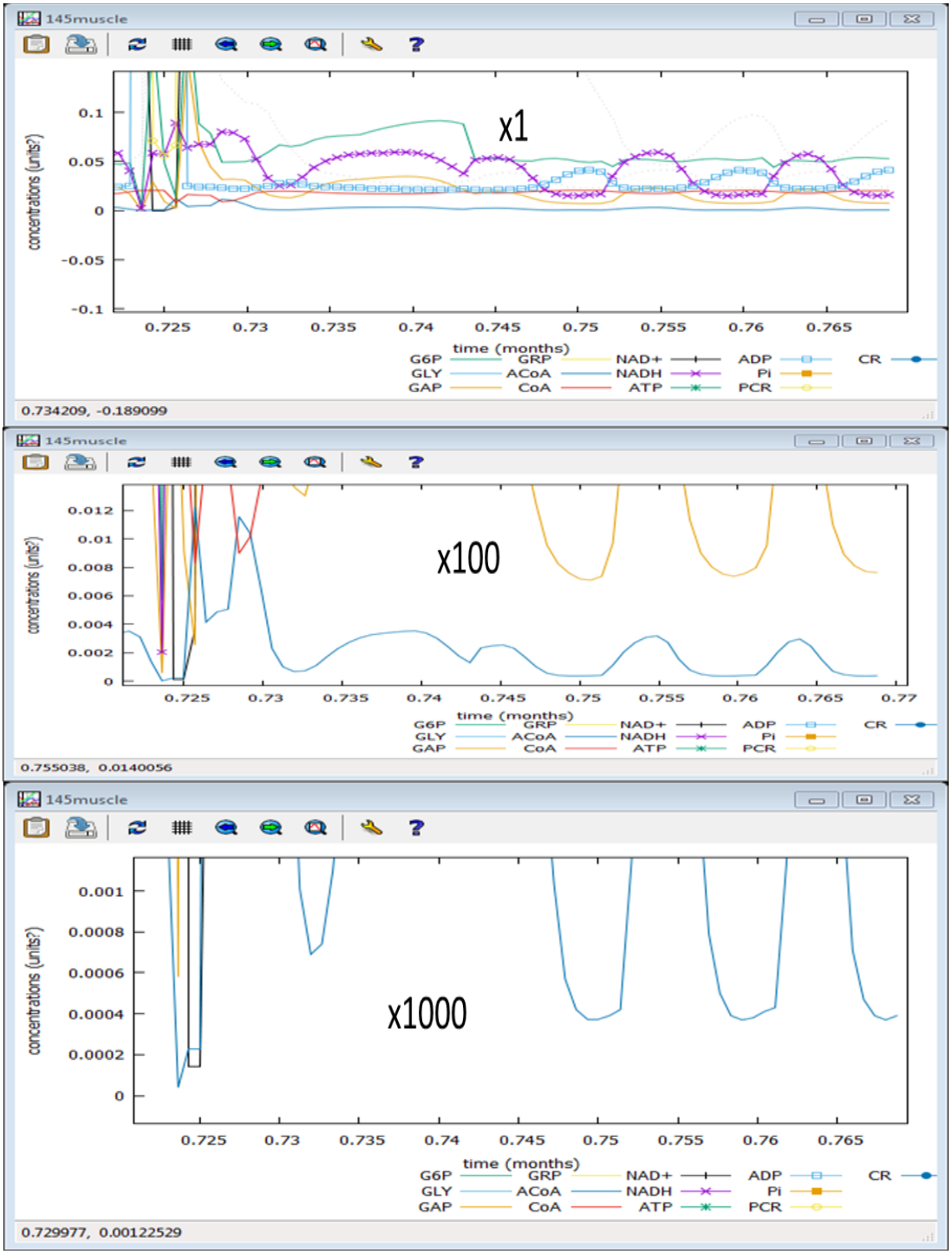

Figure 2.2.3.a. Example of screenshots showing simulated profiles of intracellular metabolites in the muscle compartment, for a scenario with very low calorie diet (3 meals a day) and periodic exercise. The plot was used to search for flux imbalances causing zero metabolite concentrations, by stepwise increasing the vertical scaling. The plot command file selectively exports profiles of glucose-6-phosphate (G6P), glycerol phosphate (GRP), glyceraldehyde-3-phosphate (GAP), NAD+, NADH, ADP, ATP, Pi, creatine (CR), phosphocreatine (PCR), glycogen (GLY), Acetyl-coenzymeA (AcoA) and coenzymeA (CoA).

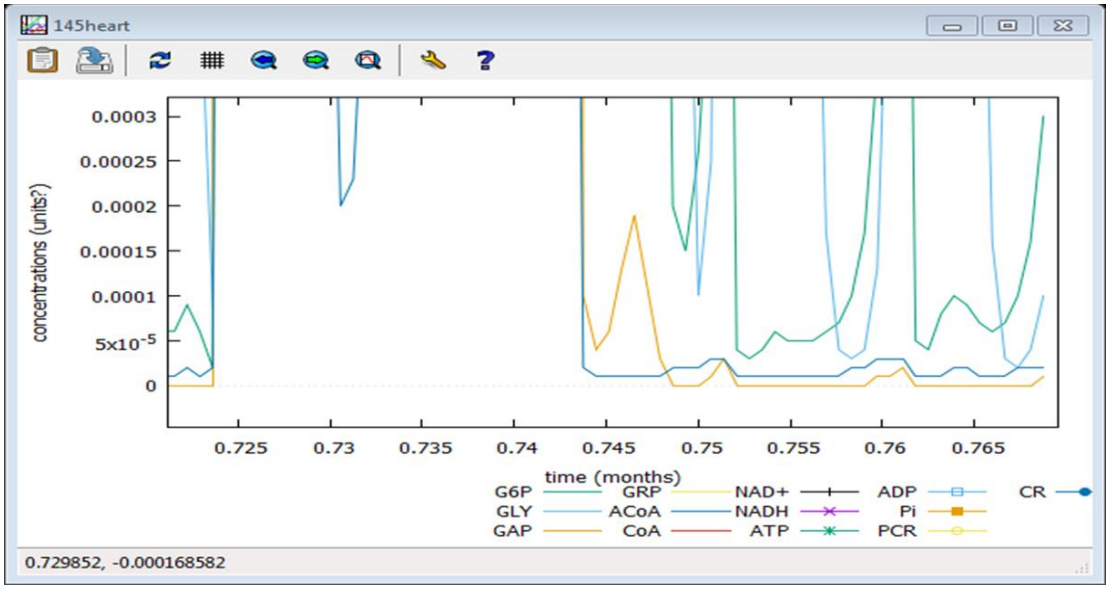

Figure 2.2.3.b. Example of a screenshot showing simulated profiles of intracellular metabolites in the heart compartment (same simulation as in figure 2.3.a.). This plot was used to diagnose the flux balance of GAP in the heart compartment as being the cause of breakdown of the integration algorithm shortly after simulation time t=0.77 months for this particular case (not shown).

Thirdly, a visualization routine was developed that plots the concentrations of the hormones glucagon, insulin and epinephrine (Ce). Such a plot can be used to rapidly inspect for which hormonal control signals are sensed by the metabolism is all the tissues. This information is helpful to decide whether issues with hormonal modulation of Vmax values may be at the root of problems with flux balances. An example is shown in figure 2.2.4.

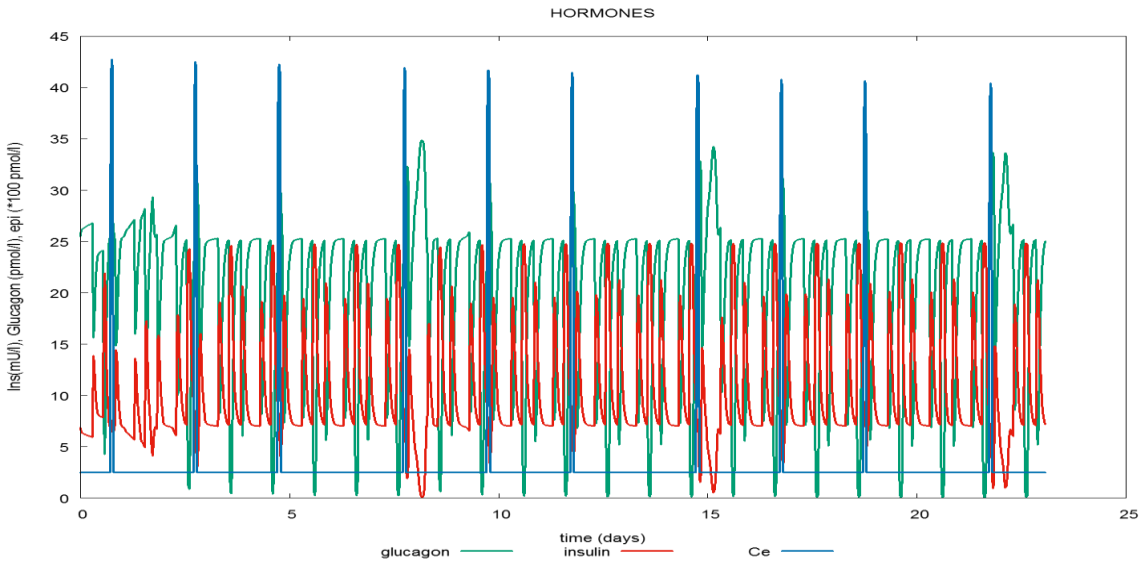

Figure 2.2.4. Example of a screenshot showing simulated profiles of the hormones glucagon, insulin and epinphrine (Ce) (same simulation as in figure 2.3).

2.2.3 Model software code aspects

Initially during the joint TNO-CNR model refinement sessions, computational speed would appear to be very slow in certain cases. Upon careful investigation, much of the problem appeared to arise from issues related to one or more intracellular metabolic pools becoming depleted, e.g., as a consequence of a mismatch in capacities of some of the metabolic pathway fluxes. *In such situations, the numerical integration routine dramatically reduces the integration step size leading to very slow progress and eventually a computational breakdown due to reaching a lower step size limit.* Most of these problems were resolved after a more extensive implementation of the control of metabolism by insulin as described in 2.2.5. The simulation of the meal and exercise effects on metabolism during a one-month period then required about 15 minutes of CPU time on a PC. With this speed, simulations of periods of up to one year have come into a manageable reach. Nevertheless, the simulations of scenario's with more extreme profiles of calorie intake and/or physical activity sometimes will still suffer from breaking down, requiring further detailed inspection and analysis using the tools described in 2.2.2.

2.2.4 Hormone controller aspects

The insulin-glucagon hormone controller was implemented as described by Kim et al. (2007), i.e., as an *integral rein controller*. This controller, although based on very simple principles, proved remarkably efficient and robust in controlling metabolic pathway activities under extreme nutritional challenges during the simulations. However, the controller locked itself frequently in situations where either insulin or glucagon decreased to 0, from where the controller was unable to recover. This situation was therefore prevented by imposing lower limits on glucagon and insulin in the simulation software such that the concentrations cannot go to 0. A further drawback was that the controller came with parameter settings that allowed ranges of concentrations of glucagon (0-35 pmol/L) and insulin (0-25 mU/L) that appear rather limited when compared to experimental values which may range at least to 80 pmol/L for glucagon and 100 mU/L for insulin (Schaller et al., 2013). Changing the parameters so as to allow for more realistic concentration ranges led to an unexplained strongly decreased robustness of the controller and therefore proved to be unfeasible in the end. As a possible improvement, to be able to compare calculated hormone concentrations to experimental values a simple range mapping (i.e., multiplication) can be programmed.

2.2.5 Refinement of the metabolic models

The metabolic model originally proposed by Kim et al. (2007) and its subsequent modifications for the purposes of MISSION-T2D in the previous reporting periods has been described in Deliverables 4.1 and 4.3. Briefly, the model was initially developed to only simulate the effects

on metabolism of a single 60 minutes period of moderate intensity exercise starting from an overnight fasting situation. Insulin, glucagon, and epinephrine are included to represent the main hormonal controllers of metabolism. In Deliverable 4.3 we described the implementation of additional equations for stomach emptying and for intestinal absorption of glucose, amino acids and triglycerides so as to allow the model to also describe the metabolism dynamics after feeding. Fermentation has not been considered at the current stage for lack of available models that would allow for easy integration in the current model.

The model validation in the overall workflow during the present reporting period showed that refinements in the following areas were required:

- Dynamic adaptation of fat mass
- Insulin-mediated glucose transport
- Dimensioning of amino acid fluxes
- Modulation of Vmax values by insulin
- Modification of glycogen dynamics

The refinements and their implementation will be described in more detail below. For reference of metabolic fluxes, figure 2.2.5 is used throughout this section 2.2.5.

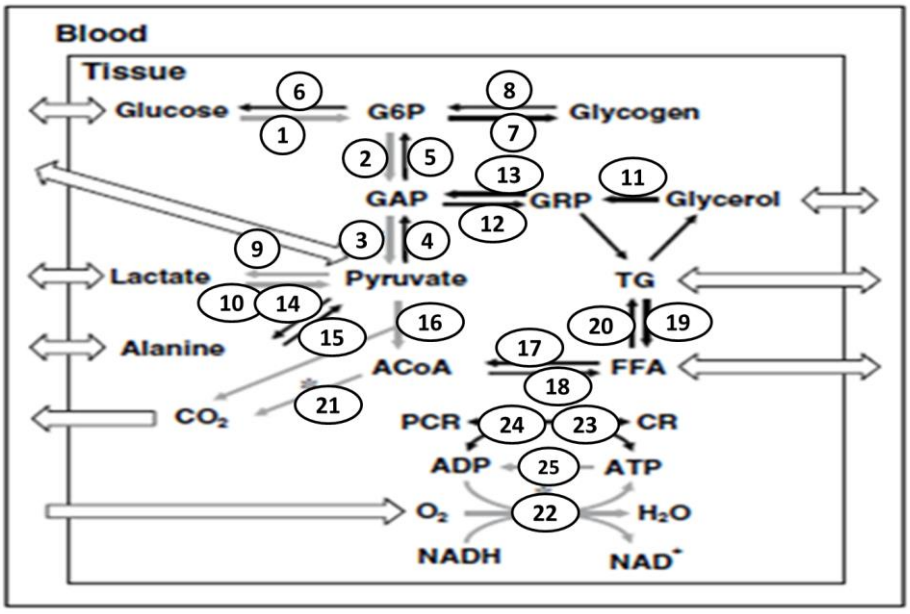

Figure 2.2.5. Intracellular metabolic fluxes with numbering used for referencing in section 2.2.5.

2.2.5.1 Dynamic adaptation of Fat Mass

Depending of the energy balance between calorie intake and calorie expenditure, body weight will change. For the purpose of the model, such changes were assumed to affect adipose tissue mass only. Therefore, in the model, depending on the daily food intake and exercise intensity, the volume of the adipose tissue compartment was updated once per day of simulated time course, according to the energy balance of that day. This update was done according to the

body weight dynamics model of Westerterp et al. (1995). Basically, this model calculates Fat Free Mass (FFM) and Fat Mass (FM) according to gender, body weight (BW), age and height:

FFM = -18.36 - 0.105 * Age + 34.009 * Height + 0.292 * BW (male);
FFM = -12.47 - 0.074 * Age + 29.392 * Height + 0.218 * BW (female).
FM = BW - FFM

Daily changes in FM and FFM are dependent on the daily energy balance (TDEB):

TDEB = TDEI – TDEE,

with TDEI the total daily energy intake and TDEE the total daily energy expenditure:

TDEE = REE + TEF + AEE,

where REE is the resting energy expenditure calculated (Mifflin et al, 1990) as:

REE = 10.0 * BW + 6.25 * 100*Height - 5 * Age + 5 (Female);
REE = 10.0 * BW + 6.25 * 100*Height - 5 * Age – 161 (Male),

TEF is the thermic effect of feeding:

TEF = TEF = 0.1 * TDEI (Westerterp et al 1995)

and AEE is the energy expenditure due to physical activity.

In practice, TDEB as above needs modification by two correction terms, i.e., in case of positive energy balance, the cost of conversion of excess calories into storage compounds (~10% of TDEB); and in case of negative energy balance, a reduction in REE depending on the size of AEE (equations Westerterp et al. 1995 not shown).
The ratio of mobilization or storage of energy between FM and FFM is not a constant. When TDEI exceeds TDEE, excess energy is stored as FM and FFM in a ratio that is a linear function of the body composition. The surplus is stored as FM and FFM in a mass ratio of 75:25 or in an energy ratio of 95:5. When TDEI is lower than TDEE, the energy deficit is covered from FM and FFM in a ratio that is a linear function of the relative size of the deficit (TDEI:TDEE) and of the body composition. In a situation where TDEI:TDEE is between 1 and 0.5 and where FM is higher than 7.5% and 15% of BM in males and females, respectively, the deficit is mobilized from FM and FFM in a mass ratio of 75:25 or in an energy ratio of 95:5. In the cases considered for the simulations, this situation mostly applies indicating that most of the weight dynamics

occur in the adipose tissue compartment. Therefore, as a first approximation, to accommodate the body weight changes arising from energy imbalance, the adipose tissue weight was equated to FM daily updated using the equations described above.

2.2.5.2 Insulin-mediated glucose transport

In the original model, hormonal control of metabolism was implemented as far as relevant only for the situation during fasting and physical exercise. Briefly, this includes the following cases (Kim, 2007):

- The reaction rates of glycogenolysis and all gluconeogenesis steps in liver are modulated by the glucagon/insulin ratio (GIR):

$$V_{x,i} = V_{x,i}^{0} \cdot \left(1.0 + \lambda_{x,i}^{G} \frac{(GIR(t) - GIR(0))^{2.0}}{\alpha_{x,i}^{G} + (GIR(t) - GIR(0))^{2.0}}\right) \quad (6)$$

- The reaction rates of glycogenolysis, glucose phosphorylation by hexokinase, lipolysis, and fatty acid oxidation in heart and skeletal muscle are modulated by the epinephrine concentration C_E:

$$V_{x,i} = V_{x,i}^{0} \cdot \left(1.0 + \lambda_{x,i}^{E} \frac{(C_E(t) - C_E(0))^{2.0}}{\alpha_{x,i}^{E} + (C_E(t) - C_E(0))^{2.0}}\right) \quad (7)$$

- Lipolysis in adipose and GI tissues is modulated by both epinephrine and insulin levels:

$$V_{x,i} = V_{x,i}^{0} \cdot \left(1.0 + \lambda_{x,i}^{G} \frac{(GIR(t) - GIR(0))^{2.0}}{\alpha_{x,i}^{G} + (GIR(t) - GIR(0))^{2.0}} + \lambda_{x,i}^{E} \frac{(C_E(t) - C_E(0))^{2.0}}{\alpha_{x,i}^{E} + (C_E(t) - C_E(0))^{2.0}}\right) \quad (8)$$

The main action of insulin, i.e., stimulating glucose transport into tissues, however was not implemented in the original model (Kim, 2007) which indeed does not feature any reaction corresponding to transmembrane glucose transport (i.e., action of the GLUT family of membrane transport proteins). Thus, initial simulations of meal intake effects on metabolism predicted an extreme accumulation of arterial glucose not seen in any of the datasets assembled in WP7. It was then tried to enlarge the flux capacity of the first step of glycolysis, i.e., the reaction glucose → Glucose-6-phosphate in an insulin-dependent manner similar to eq(6) of (Kim, 2007) shown above, but this did not cure the problem. Amplification by GIR of the other reactions in concert did not yield satisfactory results either.

We therefore needed to implement insulin-dependent amplification of glucose transport. Since inclusion of a new reaction in the software code would imply a rather large programming effort, we decided to modulate the blood-tissue partition coefficient of glucose (s[Glc]) as a simple alternative. This was implemented for 3 tissues: muscle, adipose, and liver, as follows:

$$s'[Glc] = s[Glc] \times 1/ (1+ 9 \times (Ci(t) - Ci0) / (800-Ci0) ,$$

where s'[Glc] is the modulated glucose partition coefficient, Ci(t) is the instantaneous insulin concentration, Ci0 is the basal insulin concentration, and 800 is an empirically optimized factor. This modulation factor ensures that the tissue-to-blood glucose concentration ratio is multiplied by a factor of up to 10 at high insulin concentrations. This effectively mimics an insulin-dependent increased transport capacity prevented unrealistic simulations of arterial glucose accumulation.

2.2.5.3 Re-dimensioning of amino acid fluxes

Initial simulations of meal intake scenarios showed large accumulation of the arterial alanine concentration. This was judged another consequence of the Kim et al. (2007) model not having been designed to describe the situation after feeding. In fact, only a single alanine-consuming reaction was included in the whole model, situated in the liver. This reaction (V15; see figure 2.2.5) was modified to be amplified by insulin (see 2.2.5.4). However, this resulted in large accumulation of pyruvate originating from the liver. This was then diagnosed as caused by the absence of the pyruvate dehydrogenase reaction (V16; see figure 2.2.5). It thus appeared that in the original model, the pathways for the citric acid cycle as well as de novo lipogenesis, had been disabled for the liver. After enabling V16, the problem could eventually be solved after further refinement of the modulation by hormones of several fluxes described in the next subsection.

2.2.5.4 Modulation of Vmax values by insulin

Upon simulation of meal intake scenarios, a series of issues emerged relating to unrealistic metabolite accumulation in various tissues as well as the arterial blood compartment. The main issues are discussed below, referring to fluxes by the numbering as depicted in figure 2.2.5.

Briefly, in all tissues except the brain, fluxes V1, V2 and V3 (i.e., the glycolytic pathway) needed amplification depending on the insulin-to-glucagon ratio, in order to metabolize meal-derived glucose sufficiently fast to prevent its accumulation.

Flux V4 (PYR → GAP) in the adipose tissue needed strong insulin-dependent amplification to generate sufficient glycerol for the synthesis of triglycerides from free fatty acids during scenarios of overfeeding.

Flux V7 (glycogen synthesis) needed insulin-dependent amplification, and flux V8 (glycogen hydrolysis) insulin-dependent inhibition as detailed in section 2.2.5.5.

Flux V16 (pyruvate dehydrogenase flux) needed insulin-dependent amplification in a strongly tissue-dependent manner to allow for sufficient pyruvate oxidation capacity and Acetyl-CoA supply for fatty acid synthesis capacity. In accordance, fluxes V18 and V20 needed strong insulin-dependent amplification in liver and adipose tissue for fatty acid and triglyceride synthesis capacity.
Flux V19 (triglyceride hydrolysis) needed insulin-dependent inhibition in heart, GI, and adipose tissue.
The flux modulations by the glucagon/insulin ratio as implemented by Kim et al. (2007) in the original model for simulation of metabolism during exercise were left unchanged.

Insulin-dependent Vmax amplification was done with a factor (F_{ampl}):

$$F_{ampl} = 1 + \mathbf{C_{ampl}} * igr*igr / ((Ci0)*(Ci0) + igr*igr),$$

with $\mathbf{C_{ampl}}$ a constant, and

$$igr = Ci/Cg – Ci0/Cg0,$$

where Ci and Cg are the instantaneous insulin and glucagon concentrations, and Ci0 and Cg0 the basal insulin and glucagon concentrations, respectively.
Insulin-dependent Vmax inhibition was done with a factor (F_{inhib}):

$$F_{inhib} = 2 * ((Ci0/Cg0)*(Ci0/Cg0)) / ((Ci0/Cg0)*(Ci0/Cg0) + igr*igr),$$

where igr, Ci0 and Cg0 are as defined above. The tissue-specific constants $\mathbf{C_{ampl}}$ implemented after the validation and refinement of the model are listed in Table 2.2.1.

Table 2.2.1. Tissue-specific constants $\mathbf{C_{ampl}}$ (see text) for insulin-dependent modulation of fluxes implemented after model validation and refinement. Flux numbers refer to figure 2.5.

Flux	Heart	Muscle	GI	Liver	Adipose
V1	2	5	5	5	5
V2	2	5	5	5	5
V3	5	5	5	5	2
V4					
V7	14.67	14.67		14.67	
V12					100
V15					
V16	3	3	3	1000	20
V18				1000	100
V20				1000	100

2.2.5.5 *Modification of glycogen dynamics*

Glycogen is an important energy storage metabolite, that allows to rapidly store excess glucose in case of energy overload (flux V7; figure 2.2.5), and allows to rapidly mobilize glucose reserves in case of energy depletion (via flux V8; figure 2.2.5). As such it behaves as a highly dynamic glucose energy buffer in vivo. Glycogen in the Kim (2007) model is present in liver, heart, muscle and brain. The largest flux capacity is in the muscle, the smallest in the brain. After modifying the model for meal intake scenarios in the previous reporting period, glycogen dynamics proved to be very slow and therefore this metabolite was virtually not functional as a dynamic energy buffer. This function was restored by implementing insulin-dependent amplification (factor C = 14.67; see 2.2.5.4) of reaction V7 (glycogen synthesis, i.e. glucose storage), and insulin-dependent reduction of the activity of reaction V8 (glycogen hydrolysis, i.e. glucose mobilisation) for heart, muscle, and liver (not for brain due to insignificant flux capacity). As the refinement caused glycogen concentrations to rise unrestrictedly during meal periods, the glucose storage flux was made to be strongly limited once glycogen reached a tissue-specific threshold value. This was done by additionally modulating flux V7 by a factor 1/(1+exp(GLY-GLYmax)), where GLY is the tissue glycogen concentration, and GLYmax the tissue-specific threshold set as 10 mM, 100 mM, 100mM, and 500 mM for brain, heart, muscle and liver, respectively, following literature data (Acheson et al. 1988, Obel et al. 2012).

2.2.6 Assessment of the present modeling state

The parameterization of the glucose-insulin hormone controller in the metabolic model is not yet such that it can be validated by matching with real plasma concentration values of insulin and glucagon, i.e., only relative comparisons of simulated and experimental data can be made. Amino acid and protein metabolism are modelled as if all amino acid nitrogen and carbon were concentrated in a single amino acid, i.e., alanine. Thus, no detailed predictions on amino acid metabolism are available for validation. Regarding food intake and digestion, the MISSION-T2D simulator can simulate the effects of multiple meals per day, of any desired composition in terms of total calories in carbohydrates, fat and protein. Regarding exercise, the MISSION-T2D simulator can simulate the effects of periods of exercise of “any” (within certain limits) realistic intensity and duration, placed in between meals. No gut microbial fermentation equations are included in the metabolic model i.e. the fiber component of the diet is not considered in the metabolite balances. Rather, the influence of the large intestine has been incorporated in the MISSION-T2D simulator by means of inflammation in relation to fermentation as resulting from the combination of microbiota composition and food composition (i.e., presence of fibers, model from WP2). The metabolic model embedded in the integrated simulator was thus refined to successfully simulate multi-organ metabolism dynamics for various lifestyle scenarios differing in food intake and exercise frequency and intensity. The model code can adequately handle intermediate profiles but was found liable to breakdown in scenarios where *extreme* nutritional

and/or physical exercise values are encountered. However, such breakdowns only occurred during later phases of computation, i.e., initial sections of the simulated profiles can still be used for evaluation.

2.3 Progress on the MT2D-Marvel model

2.3.1 Introduction

The high-level aggregation model of diabetes v1.0, integrating energy metabolism, glucose metabolism, inflammation and other system health variables, was introduced in deliverable 4.2. In deliverable 4.3, the personalisation of the model for specific diabetes subgroups resulting from TNO's diabetes subtyping tool, based on the functions of the pancreas, the liver and the muscle was described. In addition, results of sensitivity analysis and suggestions for metabolic-inflammatory signatures that can be used as potential risk predictors for diagnosis and coaching as derived from model simulations, were presented.

In the present reporting period, work was focused on validation of the MT2D-Marvel model against data from the Whitehall II cohort, described in section 2.3.2. This was previously decided to be done as an alternative to the originally planned validation in TNO's P4 field lab studies due to delays involved in establishing the latter.

Additionally, the work on integration of the model into a real e-health application was extended. After implementing the software infrastructure and interface with the Dutch e-health coaching application PatientCoach (discussed in a previous deliverable, D4.3), a Horizon2020 proposal under the call for proposals PHC28 was formulated to ensure a long-term application perspective as described in 2.3.3. Furthermore a manuscript "Multi-domain type 2 diabetes onset and early development forecasting model" was submitted to BMC Systems Biology.

2.3.2 Validation and refinement of the MT2D-Marvel model

Work was focused on validation of the MT2D-Marvel model against data from the Whitehall II cohort. The work and procedures to make the Whitehall II data model-ready, along with some pitfalls, are described separately in Deliverable 7.3. Summarizing, we were given data at only two phases (Phase S3 and S5) of the cohort, along with time-to-diabetes data. Serious inconsistencies in the latter prevented their use as a reference for the time dimension of diabetes development. In the MT2D Marvel model, time courses resulting from cause-effect relations are modelled as exponential curves with two parameters, i.e., strength and speed for each interaction (see D4.2). Since only data at two time points were available, no true time courses were available for model validation. Therefore, only one single effect parameter per interaction could be used, i.e., the speed parameter was fixed to the value pre-set in the model (see D4.2) and only the strength parameter was varied. Since the model contains a number of latent variables, it was decided to use Structural Equation Modelling (SEM). Since the time interval between the two cohort phases S3 and S5 was fixed, time was not a variable in the model

First, the model was adapted as follows. The processes discriminating between reversible and irreversible tissue damage were condensed into one single variable <tissue damage>. This model simplification was necessary to better align with the available data in the Whitehall II

dataset. It was verified with simulations that this simplification did not change the predicted overall kinetics of diabetes development. Furthermore, since published Whitehall II trajectory analyses show increase of inflammation and reduction of insulin sensitivity also in non-diabetic subjects (Tabak et al 2009, Carstensen et al 2010), Age was included as a control variable <Age> driving <tissue damage>. Simulations then showed increase of inflammation and insulin resistance also in the absence of overweight, in line with established diabetes risk models, that have Age as an independent risk factor (data not shown).

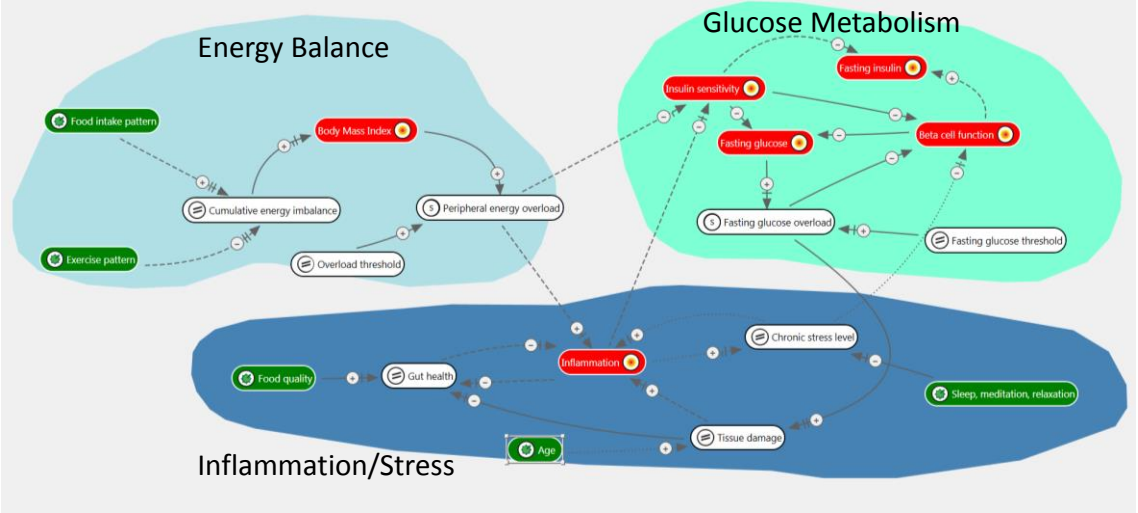

Figure 2.3.1. MT2D-Marvel model of type-2 diabetes development with reduced variable set to align with Whitehall II data availability. Age was included as additional control variable.

Until the reporting date, only preliminary results became available. The SEM modelling was set up in a stepwise fashion, initially separating the model subdomains. So far, only the model subdomains of Energy Balance and Glucose Metabolism could be considered. The equation set used for the Energy Balance subdomain was as follows:

```
WHIIdata.model<-'
CummEnergyImbalance ~ a1*FoodIntake + a2*Excercise
CummEnergyImbalance =~ a3*BMI
BMI ~ 0
PeripheralEnergyOLoad ~ a4*BMI
PeripheralEnergyOLoad ~ 0
'
```

The <Overload Threshold> was set to BMI=25. Data that showed clearly unphysiologic responses (e.g., increased food intake accompanied by weight loss) were excluded before SEM analysis. In this exploratory initial phase, the variables were used as such, i.e., not yet scaled to ranges as defined in D4.2. Results shown in Table 2.3.1 indicate that indeed fairly precise estimates of the model coefficients could be obtained for all the relations in the Energy Balance subdomain.

Table 2.3.1. Preliminary results of coefficient estimations for the Energy Balance subdomain using SEM.

```
Latent Variables:
                   Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
  CummEnergyImbalance =~
    BMI       (a3)    0.135    0.005   25.597    0.000    0.162    1.000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
  CummEnergyImbalance ~
    FoodIntak (a1)    6.124    0.425   14.396    0.000    5.099    0.549
    Excercise (a2)   -0.315    0.157   -2.007    0.045   -0.262   -0.043
  PeripheralEnergyOLoad ~
    BMI       (a4)    0.621    0.023   26.944    0.000    0.621    0.746

Covariances:
                   Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
  CummEnergyImbalance ~~
    PrphrlEnrgyOLd    0.020    0.004    5.124    0.000    0.020    0.293

Intercepts:
                Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
    BMI            0.000                              0.000    0.000
    PrphrlEnrgyOLd 0.000                              0.000    0.000
    CmmEnrgyImblnc 0.000                              0.000    0.000

Variances:
                Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
    BMI            0.000                              0.000    0.000
    PrphrlEnrgyOLd 0.005    0.000   15.748    0.000    0.005    0.258
    CmmEnrgyImblnc 1.000                              0.693    0.693
```

The equation set used for the Glucose Metabolism subdomain was:

```
 8  WHIIdata.model<-'
 9  InsulinSensitivity ~ a6*PeripheralEnergyOLoad
10  FastingInsulin ~ a8*InsulinSensitivity + a12*BetaCellFunction
11  BetaCellFunction ~ a9*InsulinSensitivity + a15*FastingGlucoseOLoad
12  FastingGlucose ~ a10*InsulinSensitivity + a13*BetaCellFunction
13  FastingGlucoseOLoad ~ a14*FastingGlucose
14  BetaCellFunction ~ 0
15  '
```

The <Fasting Glucose Threshold> was set to 6.0 mM. The results shown in Table 2.3.2 indicate that reasonably precise estimates of the model coefficients could be obtained for all the relations in the Glucose Metabolism subdomain except for coefficient a15 describing the interaction <Fasting Glucose Overload> => <Beta Cell Function>. This is probably due to the fact that data availability for glucose values > 6.0 mM was severely limited (N.B. all subjects receiving diabetic medication in any form including insulin, were excluded from analysis since such medication is expected to drastically change the interaction strengths in the model).

Table 2.3.2. Preliminary results of coefficient estimations for the Glucose Metabolism subdomain using SEM.

```
Regressions:
                     Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
  InsulinSensitivity ~
    PrphrEOL  (a6)     -0.513    0.075   -6.866    0.000   -0.513   -0.491
  FastingInsulin ~
    InslnSns  (a8)     -0.447    0.050   -9.002    0.000   -0.447   -0.487
    BtCllFnc (a12)      0.310    0.080    3.883    0.000    0.310    0.343
  BetaCellFunction ~
    InslnSns  (a9)     -0.371    0.142   -2.605    0.009   -0.371   -0.365
    FstngGOL (a15)      0.436    1.168    0.373    0.709    0.436    0.207
  FastingGlucose ~
    InslnSns (a10)     -0.283    0.054   -5.279    0.000   -0.283   -0.532
    BtCllFnc (a13)     -0.457    0.146   -3.136    0.002   -0.457   -0.874
  FastingGlucoseOLoad ~
    FstngGlc (a14)      0.880    0.053   16.596    0.000    0.880    0.969

Intercepts:
                  Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
    BetaCellFunctn   0.000                              0.000    0.000
    InsulinSnstvty  -0.054    0.019   -2.754    0.006   -0.054   -0.206
    FastingInsulin   0.027    0.015    1.833    0.067    0.027    0.113
    FastingGlucose   0.018    0.010    1.856    0.063    0.018    0.132
    FastingGlcsOLd  -0.004    0.003   -1.314    0.189   -0.004   -0.030

Variances:
                  Estimate  Std.Err  Z-value  P(>|z|)   Std.lv  Std.all
    InsulinSnstvty   0.051    0.007    7.517    0.000    0.051    0.759
    FastingInsulin   0.029    0.005    5.751    0.000    0.029    0.511
    BetaCellFunctn   0.080    0.065    1.220    0.222    0.080    1.137
    FastingGlucose   0.009    0.003    3.511    0.000    0.009    0.464
    FastingGlcsOLd   0.001    0.000    9.060    0.000    0.001    0.088
```

Analysis of the Inflammation/Stress subdomain of the model will be done shortly. Thereafter, the complete model with all 3 subdomains interlinked (i.e., feedback loops between subdomains duly considered) will be analysed. Here, a challenge lies in the high degree to which data are missing. To illustrate this fact, only around 30% of the subjects was found to have a complete data record at two time points for all variables in the combined Energy Balance and Glucose Metabolism subdomains and *not a single participant* out of the more than 10,000 in the cohort was found to have a complete data record at two time points for all variables in the full model i.e., the 3 subdomains combined. As a possible solution to overcome this difficulty as well as the issue of exclusion of unphysiologic data, a Bayesian approach may be included in the SEM analysis. After initial evidence for feasibility of the analysis is obtained, variables will be properly scaled and a true validation will be done, i.e., using cross-validation as a general technique.

2.3.3 H2020 POWER2DM – towards implementation of the MT2D-Marvel model for diabetes patient self-management support in practice.

The MT2D-Marvel model was included as one of several predictive models for diabetes development in a Horizon 2020 project proposal under the Call for proposal: H2020-PHC-2015-single-stage, within the Topic**: PHC-28-2015 - Self management of health and disease and decision support systems based on predictive computer modelling used by the patient him or herself**, thus contributing to MISSION-T2D impact and dissemination. The proposal was funded with approx. EUR 5 million upon excellent evaluation (15 points). The project is named POWER2DM - Predictive model-based decision support for diabetes patient empowerment (Project reference 689444) and started on Feb. 1, 2016. It will run for 42 months.

The main objective of POWER2DM is to develop and validate a personalized self-management support system (SMSS) for T1 and T2 diabetes patients that combines and integrates (1) a decision support system (DSS) based on leading European predictive personalized models for diabetes interlinked with predictive computer models, (2) automated e-coaching functionalities based on Behavioural Change Theories, and (3) real-time Personal Data processing and interpretation.

The DSS will be based on the complementary combination of proven predictive models for short-term plasma glucose prediction, medium term diabetes progression, and long-term risk scoring for diabetes complications. These models will be integrated in adaptive personalized behaviour change interventions to increase adherence of the patients to their care program and improve their interaction with health professionals. A cloud-based Data Integration Service, collecting and processing data from personal devices and EHR/PHR in real-time feeds the DSS. The results of the SMSS with respect to clinical parameters, awareness, acceptance and empowerment of the patient to participate in the care process will be evaluated in three studies in The Netherlands, Germany and Spain. Below is a schematic representation of the proposed architecture of the POWER2DM SMSS taken from the proposal.

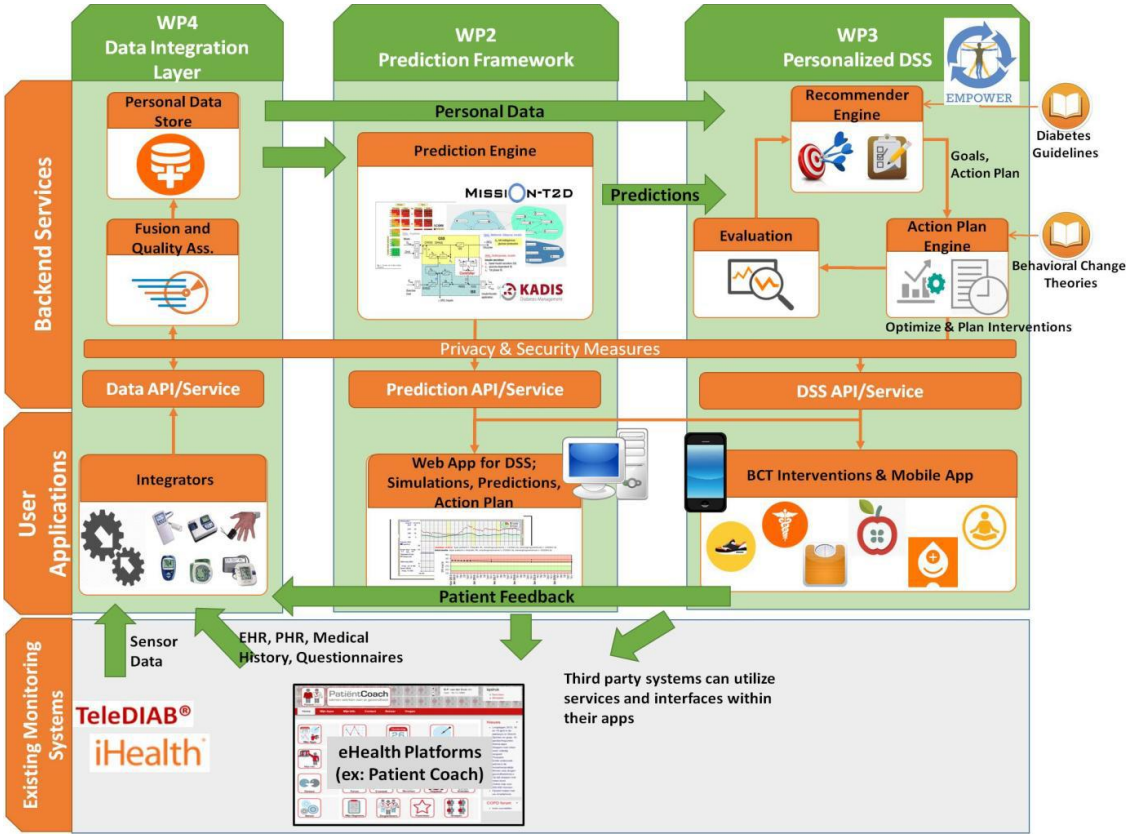

Figure 2.3.2. Proposed architecture for the POWER2DM Self Management Support System. The MT2D-Marvel model is one of the predictive models used in the Prediction Engine.

The deliverables of the POWER2DM project will increase self-management capabilities and participation of the patient in the care process, resulting in better self-control and management of the disease. This will lead to better glucose management, thereby preventing severe episodes and long-term complications. POWER2DM will reinforce the prevention sector in healthcare by raising the acceptance of SMSS based on DSS that use predictive models fed by data from personal devices. POWER2DM will challenge individuals towards more frequent and long-term use of personal devices for self-monitoring, boosting the development of these devices. POWER2DM will thereby make an essential step forward in empowering the patient, advancing prevention and decreasing disease burden and costs. The MT2D-Marvel model will play a prominent role in the project especially to support diabetes patients in reaching mid-term goals and, generally speaking, the experience gained during the MISSION-T2D project related to model building, data acquisition and analysis would certainly turn out useful. On the other hand the project POWER2DM, sharing much of the vision of MISSION-T2D, will certainly contributes to the dissemination of its results and its impact at large.

3 Deliverable Conclusions

The metabolic model embedded in the MISSION-T2D simulator (high level of detail, 1 minute - 1 year time scale) was refined and partially validated and as a result can successfully simulate multi-organ metabolism dynamics for various lifestyle scenarios differing in food intake and exercise frequency and -intensity.

The MT2D-Marvel model (low level of detail, month - 6 year time scale) was simplified and refined based on the available Whitehall II cohort data. Initial model subdomain analyses gave satisfactory results but challenges for overall validation remain.

4 Annexes

4.1 Annex 1. List of abbreviations used

BMI	Body Mass Index
DSS	Decision Support System
SEM	Structural Equation Model
SMSS	Self Management Support System
T2D	Type 2 Diabetes

5 Bibliography

Acheson KJ, Schutz Y; Bessard T, Anantharaman K, Flail J-P, and Jéquier E. (1988) Glycogen storage capacity and de novo lipogenesis during massive carbohydrate overfeeding in man. Am J Clin Nutr, 48, 240-247.

Carstensen M, Herder C, Kivimäki M, Jokela M, Roden M, Shipley MJ, Witte DR, Brunner EJ, Tabák AG. (2010) Accelerated increase in serum interleukin-1 receptor antagonist (IL-1Ra) starts 6 years before the diagnosis of type 2 diabetes: the Whitehall II prospective cohort study. Diabetes 59:1222-7.

Kim, J., G.M.Saidel, and M.E.Cabrera (2007). Multi-scale computational model of fuel homeostasis during exercise: effect of hormonal control. Ann. Biomed. Eng 35:69-90.

Mifflin MD, St Jeor ST, Hill LA, Scott BJ, Daugherty SA, Koh YO (1990) A new predictive equation for resting energy expenditure in healthy individuals. Am J Clin Nutr, 51, 241-247.

Obel LF, Müller MS , Walls AB, Sickmann HM, Bak LK, Waagepetersen HS and Schousboe A. (2012) Brain glycogen - new perspectives on its metabolic function and regulation at the subcellular level. Frontiers in neuroenergetics, 4, 1-15.

Schaller, S., Willmann, S., Lippert, J., Schaupp, L., Pieber, T.R., Schuppert, A., and Eissing, T. (2013) A Generic Integrated Physiologically based Whole-body Model of the Glucose-Insulin-Glucagon Regulatory System. CPT: Pharmacometrics & Systems Pharmacology 2, e65.

Tabak, A.G., Jokela, M., Akbaraly, T., Brunner, E.J., Kivimäki, M., Witte, D.R. (2009). Trajectories of glycemia, insulin sensitivity and insulin secretion preceding the diagnosis of type 2 diabetes: the Whitehall II study. Lancet 373, 2215–2221. doi:10.1016/S0140-6736(09)60619-X.

Westerterp K., Donkers J.H.H., Fredrix E.W.h., Boekhoudt, P. (1995) Energy intake, physical activity and body weight: a simulation model. Br J Nutr, 73, 337-347.