G.A. nº 270086

Collaborative Project of the 7th Framework Programme

WP4: Vertical model integration

D4.1: Assessment of software tools used for mechanistic modelling at each level of the body

V0.5 29/07/2011

www.Synergy-COPD.eu

Document Information

Project Number	270086	Acronym	Synergy-COPD
Full title	Modelling and simulation environment for systems medicine (Chronic obstructive pulmonary disease -COPD- as a use case)		
Project URL	http://www.Synergy-COPD.eu		
EU Project officer	Amalia I.VLAD		

Deliverable	Number	4.1	Title	Assessment of software tools used for mechanistic modelling at each level of the body.
Work package	Number	4	Title	Vertical model integration

Date of delivery	Contractual	PM06	Actual	PM06
Nature	Prototype ☐ Report ☑ Dissemination ☐ Other ☐			
Dissemination Level	Public ☑ Consortium ☐			

Document responsible	Dr. Josep Roca		Email	jroca@clinic.ub.es
	Partner	IDIBAPS	Phone	(+34) 932 275 400, ext. 2698

Authors	Name	Partner
Main author(s)	Isaac Cano	IDIBAPS
Co-author(s)	Josep Roca	IDIBAPS
	Vitaly Selivanov	IDIBAPS
	Igor Marin	IDIBAPS
	Marta Cascante	IDIBAPS
	Peter D. Wagner	External Advisor
	Luigi Ceccaroni	BDIGITAL
	Kelly Burrowes	UOXF.BL
	Dieter Maier	BIOMAX
	David Gomez	KI

Abstract (for dissemination)	This document focuses on characterising the mechanistic models used in the scope of the Synergy-COPD project. The characterisation of the models consists of providing the “minimum information” needed for each model to support the vertical integration. The proposed model annotation schema aims at providing a concise description of the biological processes encoded by the model: a) main attributes and parameters of the model are annotated

	using unique identifiers and ontologies, b) input and output attribute specification (i.e., units, source code name, etc.) and c) execution environment. All the model annotations are imported into a knowledge base so that there is going to be a central repository of model annotations that is going to be used to automate the vertical integration of the mechanistic models in a mid-term basis. In addition, the knowledge base potentially automates the semantic detection of the linking attributes between models.
Key words	Mechanistic modelling, model annotation schema, integrative architecture.

Version Log			
Issue Date	**Version**	**Author**	**Change**
01/07/2011	0.1	All partners involved	First draft considering the input of the partners
11/07/2011	0.2	All partners involved	First complete draft sent to the partners for review
18/07/2011	0.3	Isaac Cano, Josep Roca	Final draft text considering input from the partners sent to overview.
27/07/2011	0.4	Isaac Cano, Josep Roca	Final text ready for the overview from all partners.
28/07/2011	0.5	All partners involved	Final text ready for submission.

Index

1. EXECUTIVE SUMMARY 6

2. INTRODUCTION 7

3. DESCRIPTION OF THE MODELS 10

3.1 MODEL 6: OXYGEN TRANSPORT AND UTILIZATION 10
3.1.1 Abstract 10
3.1.2 Model implementation 12
3.1.2.1 Programming Language 12
3.1.2.2 Modular description of the model 12
3.1.2.2.1 Aim of the module 12
3.1.2.2.2 Algorithm and parameters 14
3.1.2.2.3 Inputs and outputs 15
3.2 MODEL 3: SPATIAL HETEROGENEITIES OF LUNG VENTILATION AND PERFUSION 17
3.2.1 Abstract 17
3.2.2 Model implementation 18
3.2.2.1 Programming Language 18
3.2.2.2 Modular description of the model 18
3.2.2.2.1 Aim of the model 18
3.2.2.2.2 Algorithm and parameters 19
3.2.2.2.3 Inputs and outputs 20
3.3 MODEL 7: CELL BIOENERGETICS, MITOCHONDRIAL RESPIRATION AND REACTIVE-OXYGEN-SPECIES GENERATION 22
3.3.1 Abstract 22
3.3.2 Model implementation 23
3.3.2.1 Programming Language 23
3.3.2.2 Modular description of the model 24
3.3.2.2.1 Electron Chain Module 24
3.3.2.2.1.1 Aim of the electron chain module 24
3.3.2.2.1.2 Algorithm and parameters 24
3.3.2.2.1.3 Inputs and outputs 24
3.3.2.2.2 TCA cycle module 27
3.3.2.2.2.1 Aim of the TCA cycle module 27
3.3.2.2.2.2 Algorithm and parameters 27
3.3.2.2.2.3 Inputs and outputs 27

4. PRELIMINARY ARCHITECTURE FOR INTEGRATING THE MODELS 29

4.1 PRELIMINARY ARCHITECTURE FOR INTEGRATING THE MECHANISTIC MODELS 29
4.1.1 Conceptual Architecture for integrating mechanistic models 29
4.1.1.1 Synopsys 29
4.1.1.2 Description of the approach 30
4.2 PRELIMINARY MANUALLY INTEGRATED MODEL 33
4.2.1 The point of connection 33
4.2.2 The mechanism of connection 34
4.2.3 Conversion of units in the connected models 38
4.3 ATTRIBUTES MAPPING OF THE TWO MAIN MODELS 38

5. REFERENCES 40

LIST OF KEY WORDS/ABBREVIATIONS 41

6. APPENDICES 42

6.1 APPENDIX 1 -- MINIMAL INFORMATION STANDARDS 42

List of figures

FIGURE 1: *SYNERGY-COPD INITIALLY CONSIDERED MODELS.* 7
FIGURE 2: *GRAPHICAL REPRESENTATION OF THE MODEL ANNOTATION SCHEMA PROPOSED FOR THE DESCRIPTION OF THE SYNERGY-COPD MECHANISTIC MODELS.* 9
FIGURE 3: *THIS ILLUSTRATION SHOWS THE FIVE EQUATIONS (RIGHT) THAT CORRESPONDS TO THE FIVE MAIN FUNCTIONS (LEFT) THAT THIS MODEL CONSIDERS AS AN INTEGRATED SYSTEM.* 11
FIGURE 4: *(A) CT IMAGE PROVIDING PATIENT DATA. (B) ACINAR DATA POINTS CREATED WITHIN THE LUNG VOLUME. (C) PREDICTION OF ALVEOLAR VENTILATION (VA) AND BLOOD FLOW IS MADE AT EACH OF THE ACINAR POINTS IN (B) BASED ON KNOWN CORRELATIONS OF VA, Q, AND PATIENT MEASURED VALUES.* 18
FIGURE 5: *SCHEME FOR MITOCHONDRIAL RESPIRATION AND LINKED PROCESSES SIMULATED IN THE MODEL. TWO REACTIONS LEAD FROM PYRUVATE TO SUCCINATE AND FURTHER TRANSFORMATION TO OXALOACETATE REDUCE NAD+ TO NADH. THE LATTER IS USED BY COMPLEX I TO GENERATE A TRANSMEMBRANE ELECTROCHEMICAL PROTON POTENTIAL (DMH+) AND REDUCE UBIQUINONE (Q) TO UBIQUINOL (QH_2), OXIDATION OF WHICH BY COMPLEX III ALSO CONTRIBUTES TO DMH+. COMPLEX III REDUCES CYTOCHROME C, OXIDATION OF WHICH BY COMPLEX IV AND REDUCTION OF MOLECULAR OXYGEN TO H_2O IS ALSO COUPLED TO DMH+ GENERATION. OXIDATION OF SUCCINATE TO FUMARATE BY COMPLEX II IS COUPLED TO THE REDUCTION OF UBIQUINONE AND THUS FUELS COMPLEX III. THE PRODUCT OF ELECTRON TRANSPORT, DMH+, IS CONSUMED FOR ATP SYNTHESIS.* 22
FIGURE 6: *M6+M7 BLACK-BOX EXAMPLE.* 32
FIGURE 7: *OUTPUT DEFINITIONS FOR EACH BOX.* 32
FIGURE 8: *HEMOGLOBIN SATURATION CURVE.* 36
FIGURE 9: *IN THIS FIGURE RED POINTS ARE TAKEN FROM THE SATURATION CURVE AND THE GREEN LINE IS CALCULATED USING EQ.5.* 37
FIGURE 10: *EXAMPLE OF A SMALL CURATED AND ANNOTATED MODEL* 45

1. Executive summary

The current version of the deliverable is planned to be further refined such that a new release will be available by the end of August 2011. We would like, however, to highlight the high intrinsic value of the current format of D4.1 in the two domains described below. We believe that it perfectly justifies the current edition of the document:

Biomedical domain

The main objective of the deliverable is to report on the classical models being used for vertical integration in the project. It is of note, however, that in the process of elaboration of the deliverable, the consortium has generated two types of unexpected contributions that are enriching the scientific knowledge in the area.

First, the original mechanistic models were originated in three different biomedical areas that often show a certain degree of fragmentation: a) classical physiologists addressing oxygen transport from the atmosphere to the cell; b) bioengineers modelling of spatial heterogeneities in lung ventilation and perfusion having impact of arterial oxygenation; and, c) specialists on biochemistry addressing cell bioenergetics and reactive oxygen species generation at mitochondrial level. In this regard, the deliverable represents a substantial effort of conceptual integration of this area of knowledge.

Second, different interactions within the consortium during this first phase of the project have prompted an evolution in some of the mechanistic models. As a consequence, we are planning a refined version of some of the oxygen transport modules to enhance the representation of functional heterogeneities both at lung and at tissue level, to be reported in D4.2.

Bioinformatics domain

The deliverable exhibits remarkable achievements in the preparation of the strategies for vertical integration to be further developed in D4.2. The initial description of such strategies and the effort carried out in terms of annotation of the models are reflecting those achievements alluded to above. Last but not least, the report of the preliminary steps toward semantic integration sheds light on the approach taken to address interactions between mechanistic and probabilistic modelling, one of the main components of the simulation environment to be developed within the project lifetime.

2. Introduction

It is stated in de description of work that the title of this document is “Assessment of software tools used for mechanistic modelling at each level of the body”. This title can have several interpretations and from all the possible interpretations we have considered this document to focus on characterising the mechanistic models used in the scope of the Synergy-COPD project.

One other possible interpretation could be to use this document to describe the state of the art of software tools used for mechanistic modelling but we consider, together with the Synergy-COPD consortium, that it is primarily important to produce a description of the mechanistic models considered in this project for all the partners to have a reference document for better understanding the mechanistic models.

The description of work of the Synergy-COPD considers five models: **M1**) *Central and peripheral O_2 transport and utilization,* **M2**) *Pulmonary gas exchange*, **M3**) *Spatial heterogeneities of lung ventilation and perfusion*, **M4**) *Skeletal muscle bioenergetics* and **M5**) *Mitochondrial reactive oxygen species (ROS) generation*. **Figure 1** graphically shows the five initially considered models and their corresponding level of the body.

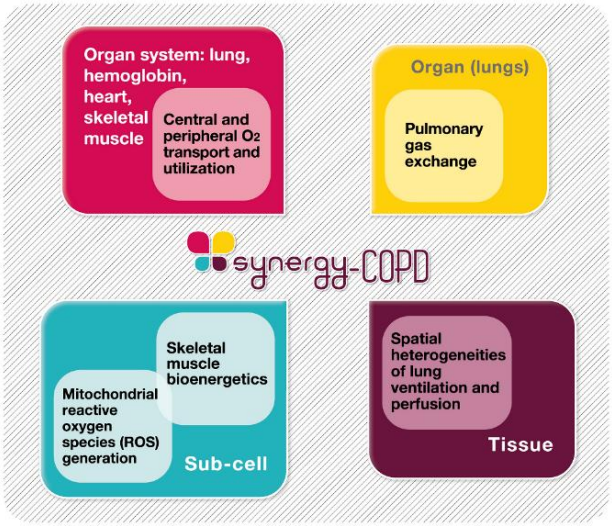


Figure 1: *.Synergy-COPD initially considered models.*

However, **M1**-**M2** and **M4**-**M5** have already been mechanistically integrated by their main authors: Peter D. Wagner (external advisor) and Vitaly Selivanov (IDIBAPS) respectively, and accordingly, the Synergy-COPD project now considers the following three main models, as an evolution and integration of five previously-considered models:

- M6. ***Oxygen transport and utilization*** (M1 + M2)
- M3. ***Spatial heterogeneities of lung ventilation and perfusion***

- M7. *Cell Bioenergetics, mitochondrial respiration and reactive-oxygen-species generation* (M4 + M5)

The ability to construct an integrative model of the above models is highly dependent on the ontological annotation (including metadata) associated with model components. Minimum information specifications exist for experimental measurements in specific areas and the Minimum Information about a Biomedical or Biological Investigation (MIBBI) web site at [mibbi.org] (accessed 18 July 2011) has been established to provide a portal for these and for model annotation. Examples of minimum information standards for biological modelling are *minimal information required in the annotation of biochemical models* (MIRIAM) and *minimum information about a simulation experiment* (MIASE). Check appendix 1 for more information about the above minimum information standards. A description of available standards and decision criteria for adoption of standards within Synergy-COPD is provided in Deliverable 3.1, Section 1.

In order to provide the minimum information required for the integration of the mechanistic models, we have considered the above minimum information specifications as the basis of the proposed *model annotation schema* (MAS). The model annotation schema depicted in see **Figure 2** aims at providing a description of the model that considers the following items:

- **Abstract:** a concise description of the biological processes encoded by the model and sub-models. For instance, model M7 (described in Section 3.3) models the cell bioenergetics, mitochondrial respiration and reactive-oxygen-species generation. However this model is the result of integrating two sub-models that respectively encode the *Electron Chain (see Section 3.3.2.2.1)* and the *TCA cycle (see Section 3.3.2.2.2)* sub-models.
- **Model Implementation:** description of the software tools used for implementing the ordinary differential equations that encode the modelled biological processes. For instance, model M7 is implemented using the C++ programing language and standard C++ programing libraries.
- **Model constituents:**
    - **Identifier:** all the main model constituents (i.e., variables, parameters, biological entities from molecular to higher levels) are annotated, if possible, using unique identifiers. For instance, one of the main constituents of model M7 is the Oxygen (O_2) compound, which is annotated using the KEGG[1] object identifier C00007.
    - **Ontological annotation:** in addition to the unique identifier annotation, all model constituents are ontologically annotated, if possible, with external ontologies, such us the Gene Ontology[2] or the Medical Subject Heading (MeSH) NLM controlled

[1] **KEGG objects** are biological entities from molecular to higher levels. Each object (except for GENES) is identified by the KEGG object identifier, consisting of a five-digit number prefixed by an upper-case alphabet, such as K05032and D00336, or prefixed by a 2-4 letter code for PATHWAY and BRITE, such asmap00010 and br08301.

[2] The **Gene Ontology (GO)** project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases.

vocabulary thesaurus. A full list of available ontological annotations is provided in Section 2.4.2 of deliverable 3.1.

- **Input/output specification:** for those model constituents that potentially can be linking points between models it is crucial to state their role as input or output when being integrated with other models.
- **Predefined measurement units:** for those model constituents that can be measured either quantitative or qualitative it is specified the units in which the model assumes the mode constituent to be measured. This is crucial when integrating the different models for model consistency purposes.

- **Source code name:** for model simulation purposes, all model constituents are defined in terms of their source code name so that Synergy-COPD programmers can relate each model constituent with its corresponding name in the source code of the model implementation.
- **Execution environment:** A description of the initial conditions needed for simulating the model is provided in order to be able to simulate the model under specific conditions.

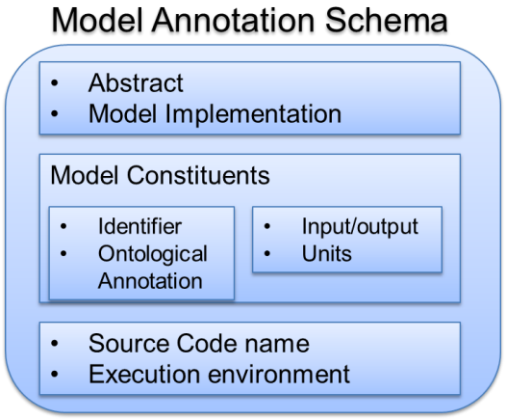


Figure 2: *Graphical representation of the model annotation schema proposed for the description of the Synergy-COPD mechanistic models.*

The annotation of the models provided in Section 3 follows the proposed Model Annotation Schema, which may be updated during the execution of the Synergy-COPD project in order to provide a description of the mechanistic models useful along the life-time of the project. In addition, all the model annotations are imported into a knowledge base so that there is a central repository of model annotations that facilitates the detection of linking attributes between models. See Section 3 of deliverable 3.1 for further details on the Synergy-COPD knowledge base.

The proposed architecture to automate the integration of mechanistic models is presented in Section 4.1. One of the key characteristics of the proposed architecture is the assumption that the different computational models can be encoded in their own programming language or in any public, standardized, machine-readable format. This is that the proposed architecture does not require encoding the mechanistic models in a common intermediate format like the Systems Biology Markup Language (SBML) even though it is generally preferable for standardization purposes.

In addition, Section 4.2 provides the description of the manual integration of the *Central and Peripheral O_2 Transport and Utilization and Pulmonary Gas Exchange* and the *Skeletal Muscle Bioenergetics and Mitochondrial Reactive Oxygen Species Generation* models. The manual integration is going to be used as a first prototype in order to be able to start building the simulation engine and web interface on top of it. Once the architecture for integrating the models is available it will substitute the monolithic integration of the mechanistic models.

3. Description of the models

The aim of this Section is to create a minimal but sufficient description of each model that is used for integration the models at a semantic level (i.e., mapping of attributes), which is tackled in Section 4.

As stated in Section 0 Synergy-COPD now considers three mechanistic models:

- M6. ***Oxygen transport and utilization***
- M3. ***Spatial heterogeneities of lung ventilation and perfusion***
- M7. ***Cell Bioenergetics, mitochondrial respiration and reactive-oxygen-species generation***

These models are described based on the model annotation schema proposed in Section 0.

3.1 Model 6: Oxygen transport and utilization

3.1.1 Abstract

Short introduction of the model:

- **What does it model?**

  This model predicts O_2 (Oxygen) transport from air to mitochondria, and then couples this to O_2 utilisation by the mitochondria. The model answers the question:

Given the transport capacity of the lungs, heart, blood and muscles (characterised by numerical values for the major transport variables), how much O_2 can be supplied to the tissues, and what are the partial pressure of oxygen (PO_2) values at each step?

This model considers the O_2 transport/utilisation pathway as an integrated system.

Four main **components**:

- ✓ Lungs
- ✓ Heart
- ✓ Circulation
- ✓ Muscle

Five main **functions**:

- ✓ Ventilation
- ✓ Alveolar diffusion
- ✓ Circulatory perfusion
- ✓ Muscle diffusion
- ✓ Mitochondrial metabolism

- Schematic description of the biological process.

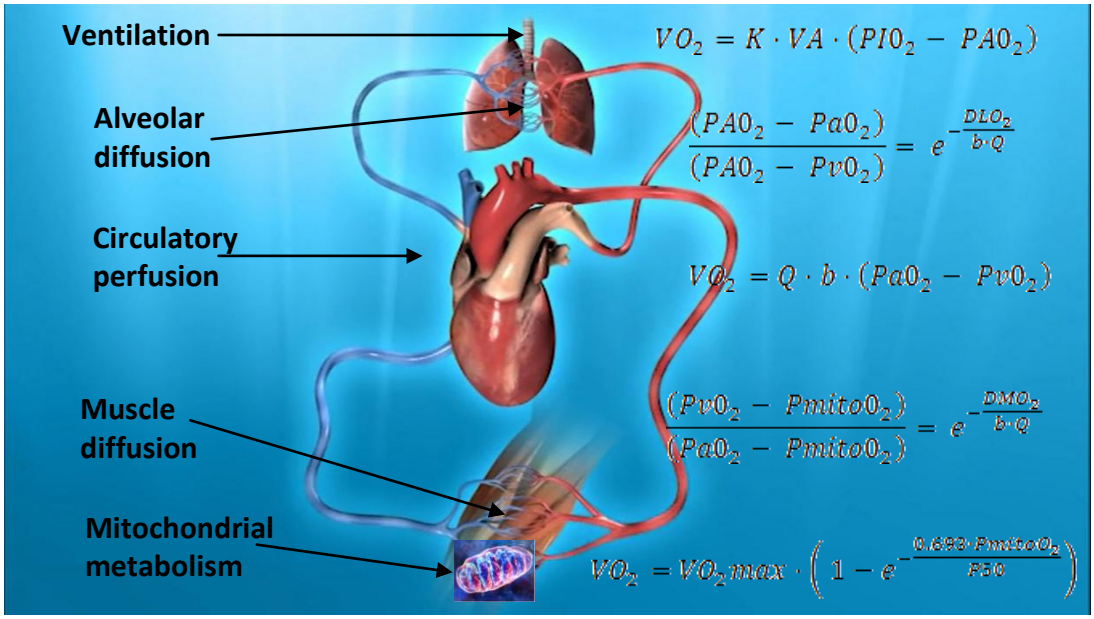


Figure 3: *This illustration shows the five equations (right) that corresponds to the five main functions (left) that this model considers as an integrated system.*

3.1.2 Model implementation

3.1.2.1 Programming Language

The source code of the model is written in the FORTRAN 77 programming language so the only requirement is the corresponding FORTRAN compiler for your platform:

- **In Unix-like systems** → fort77 or gfortran
- **In windows systems** → Microsoft FORTRAN 5 or above.

3.1.2.2 Modular description of the model

3.1.2.2.1 Aim of the module

The central and peripheral O_2 transport and utilisation and pulmonary gas exchange model is encoded as a unique module so the modular description of the model corresponds to the description of the whole model.

The biological rationale of this model is the following:

O2 transport is accomplished by 4 organs/tissues - lungs, heart, blood and muscles. Transport through this system involves four sequential processes or steps:

1. **VENTILATION TO BRING O_2 FROM AIR TO THE ALVEOLAR GAS**

$$VO_2 = K \cdot VA \cdot (PIO_2 - PAO_2)$$

2. **DIFFUSION OF O2 FROM ALVEOLAR GAS INTO CAPILLARY BLOOD**

$$\frac{(PAO_2 - PaO_2)}{(PAO_2 - PvO_2)} = e^{-\frac{DLO_2}{b \cdot Q}}$$

3. **CIRCULATORY TRANSPORT OF O2 FROM LUNGS TO TISSUE MICROVESSELS**

$$VO_2 = Q \cdot b \cdot (PaO_2 - PvO_2)$$

4. DIFFUSION OF O2 FROM TISSUE MICROVESSELS TO MITOCHONDRIA

$$\frac{(Pv0_2 - Pmito0_2)}{(Pa0_2 - Pmito0_2)} = e^{-\frac{DMO_2}{b \cdot Q}}$$

Modelling is based on well-established principles of mass conservation for O_2 at every step, using correspondingly well-established transport equations. There is one equation for each of the above 4 steps.

A fifth equation brings in mitochondrial metabolism. It expresses the rate of mitochondrial O_2 consumption as an exponential, two parameter function of mitochondrial PO2. The two parameters are the P50 and VO_2max of the exponential relationship:

5. MITOCHONDRIAL METABOLISM

$$VO_2 = mitoVO_2max \cdot \left(1 - e^{-\frac{0.693 \cdot PmitoO_2}{P50}}\right)$$

Implicit in the above is knowing the oxygen–hemoglobin (HbO_2) dissociation curve in a quantitative sense. This means that if we know the partial pressure of Oxygen (PO_2), we know the concentration of Oxygen ($[O_2]$).

As should be clear, we have 5 equations and 5 unknowns. This means a unique solution can be obtained. This model finds the solution for a desired set of input variables.

The assumption of a linear oxygen–hemoglobin (HbO_2) dissociation curve is only presented for explanation purposes, but the actual solution applied in this computational model use the real O_2 and CO_2 dissociation curves and makes no linear approximation to either. This also means that O_2-CO_2 interaction is accounted for. However, this makes the solutions possible only via numerical analysis. Closed form, algebraic solutions are not feasible.

There are some important assumptions/approximations:

1. O_2 transport is in a steady state - all input and output variables are constant in time.
2. The lungs are currently assumed to be homogeneous.
3. The tissues are likewise assumed to be homogeneous.

Technical note:

Both lung and muscle diffusing capacity are constant lumped-parameter variables in this version because that is how they are usually measured. In reality, they contain a true diffusion component plus an element that recognises that chemical combination with Hb is not instantaneous. In addition, the algorithm uses an arbitrary value for volume of blood in the lung/tissue capillaries (called VC). While VC is technically needed to compute transit time, the final result depends only on DLO_2/Q, or DMO_2/Q, and is independent of choice of VC. In other words, choosing a large VC gives a long transit time for a given blood flow rate, but the time course is precisely slowed down in proportion so that endcapillary and mean capillary data are unaffected.

3.1.2.2.2 Algorithm and parameters

This model has a bisection search algorithm that finds the unique mitochondrial PO_2 that satisfies simultaneously the first four equations, O_2 transport model from mouth to mitochondria and the mitochondrial respiration curve (equation 5).

The logic is to cycle the system between lungs and tissues until convergence and stability of outcome variables is established. In the lungs, starting with a trial venous and alveolar PO_2, the arterial PO_2 is determined by calculating the end capillary value according to the diffusion equation (eqn. 2). This output arterial PO_2 then becomes the input arterial PO_2 for tissue O_2 diffusion assuming a mitochondrial PO_2 to start. This yields a venous PO_2, and then we cycle back to the lung with this new venous PO_2 and so on.

Ultimately, stability is reached and at that point, VO_2 will be the same no matter which of the 5 equations is used to compute it.

3.1.2.2.3 Inputs and outputs

The independent (**INPUT**) variables are:

Variable	Description	Units	Input/Output	Ontological Annotation
PIO2	partial oxygen pressure in inhaled air	mmHg	input	Partial;Oxygen;Pressure
VA	Alveolar Ventilation	l/min	input	Alveolar;Ventilation
DLO2	lung diffusion capacity	ml/min/mmHg	input	Lung;Diffusion;Capacity
Q	cardiac output	l/min	input	Heart
b	slope of the O2-hemoglobin dissociation curve	ml/ml/mmHg	input	GO:0005833 hemoglobin complex;SBO:0000180 dissociation
DMO2	muscle diffusion capacity	ml/min/mm Hg	input	Capacity;Diffusion;Muscle
mitoVO2max	maximal mitochondrial oxygen consumption when oxygen is in excess	ml/min	input	GO:0005739 mitochondrion;Consumption;Oxygen
P50	partial oxygen pressure at which speed of mitochondrial oxygen consumption is half maximal	mmHg	input	Pressure;Partial;Oxygen

The **default** model parameter values are the following:

Parameter	Description	Units	Default value
PVO2	Mixed venous PO_2 for lung diffusion computation	mmHg	20.0
PVCO2	Mixed venous PCO_2 for lung diffusion computation	mmHg	75.0
PALVO2	Alveolar PO_2 for lung diffusion computation	mmHg	100.0
PALVCO2	Alveolar PCO_2 for lung diffusion computation	mmHg	35.0
VA	Default ventilation	l/min	85.0
PB	Barometric pressure	mmHg	760.0
FIO2	Inspired O_2 fraction		0.2093
PIO2	Corresponding inspired PO_2		$\mathbf{PIO2 = FIO2 * (PB - 50.0)}$
FICO2	Inspired CO_2 fraction		0.0
FAO2	Alveolar O_2 fraction		$\mathbf{FAO2 = PALVO2/(PB - 50.0)}$
FACO2	Alveolar CO_2 fraction		$\mathbf{FACO2 = PALVCO2/(PB - 50.0)}$
REXP	assumed respiratory exchange ratio or RQ (=VCO2/VO2) at VO_2max it should be reduced at lower exercise intensities. It is held constant.		1.15

O2SOL	O2 physical solubility in water/blood	ml/dl/ mmHg	0.003
HB	Hemoglobin	gm/100 ml	15.0
HRCIT	hematocrit (per cent)	gm/100 ml	3.0*HB
TEMP	body temperature	degrees centigrade	38.0
P50	O2 hemoglobin dissociation curve P50	Torr3	26.8
APH	PH of blood equilibrated with PCO2 of 30	Torr	7.30
APCO2	corresponding PCO2	30 Torr	30.0
BPH	PH of blood equilibrated with PCO2	60 Torr	7.10
BPCO2	corresponding PCO2	60 Torr	60.0
QT	total tissue or pulmonary blood flow	l/min	20.0
VC	capillary blood volume	ml	5.0*QT
DMO2	membrane diffusing capacity for O2	ml/min/Torr	
DT	time step interval for numerical integration	seconds	0.001
ITHETA			0
BPCO2			60.0
DLNGO2			60.0
DMUSO2			80.0
THALF	half time of the CO2 chemical reactions, diffusion and its chemical reactions in blood	seconds	0.15

Here are the constants for the THETA-O2 relationship: should one want to model diffusion incorporating chemical reaction rates as stated, we do not do this because experimental data are for total DLO2 incorporating this effect already:

- **B1**=36.1633
- **B2**=-1.5551
- **B3**=0.0234
- **B4**=-1.15173E-4

[3] The torr is a non-SI unit of pressure with the ratio of 760 to 1 standard atmosphere, chosen to be roughly equal to the fluid pressure exerted by a millimetre of mercury, i.e., a pressure of 1 Torr is approximately equal to 1 mmHg. [source wikipedia]

The dependent (**OUTPUT**) variables are:

Variable	Description	Units	Input/Output	Ontological Annotation
VO2	consumed oxygen	l/min	output	Oxygen;Consumption
PAO2	alveolar partial oxygen pressure	mmHg	output	Pressure;Oxygen;Partial; Alveolar
PaO2	arterial partial oxygen pressure. CaO2 is the [O2] at PaO2	mmHg	output	Pressure;Oxygen;Partial; Arterial Blood Gas Measurement;Arterial
PvO2	venous partial oxygen pressure. CvO2 is the [O2] at PvO2	mmHg	output	Pressure;Oxygen;Partial; Venous Blood
PmitoO2	mitochondrial partial oxygen pressure	mmHg	output	GO:0005739 mitochondrion;Oxygen; Partial;Pressure atio

3.2 Model 3: Spatial heterogeneities of lung ventilation and perfusion

3.2.1 Abstract

Concise description of the model:

- What does it model?

  It derives anatomically-based patient specific geometric models of the lung surface, airways and blood vessels in order to predict the regional distribution of ventilation and blood flow within the lung. The anatomically-based models can be used to increase our understanding of how regional perturbations to the structure or function of the airway and vascular trees and the state of health of the functional tissue affect gas exchange.

  The current modelling process is somewhat complicated (details below) therefore during the course of the Synergy-COPD project we plan to develop a simplified modelling system, based on these more detailed models which will be more easily integrated into the Synergy-COPD modelling platform. A detailed model will be created and analysed for one of each of the three COPD phenotypes being explored within Synergy-COPD through the PAC-COPD database. For additional subjects the approach will be to use CT data from a given patient to describe the model geometry and tissue density – enabling creation of acinar points within the model and an allocation of the tissue properties at that location (i.e. is emphysema present at a given acinar location). Given this imaging input as well as information about the patient's pulmonary artery pressure and/or cardiac output and tidal ventilation volume, the model will be used to predict the regional distribution of alveolar ventilation and blood flow at each acinus within the lung. This simplified approach that will be adopted is illustrated schematically in the figure below.

- Schematic description of the biological process.

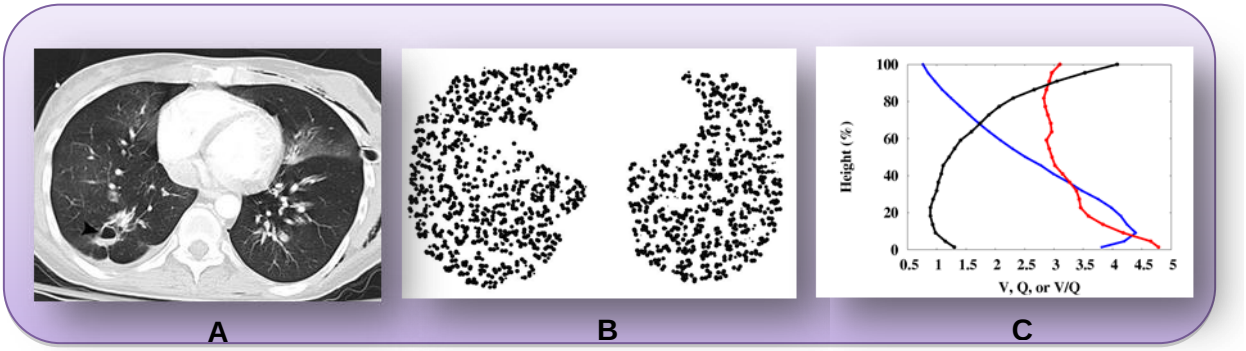


Figure 4: *(A) CT image providing patient data. (B) Acinar data points created within the lung volume. (C) Prediction of alveolar ventilation (VA) and blood flow is made at each of the acinar points in (B) based on known correlations of VA, Q, and patient measured values.*

3.2.2 Model implementation

This Subsection aims at specifying the technical details of the model implementation and modularity. It is supposed that biological processes can be sliced into layers or modules so the models have different functional modules.

3.2.2.1 Programming Language

The programming language used to code the model and the minimum requirements for compiling/running are:

- **Compiler version:** g77
- **Required libraries:** The current software is called CMISS [www.cmiss.org] (accessed 19 July 2011) – which is implemented in Fortran 77 and requires several Perl and solver libraries.
- **Supported operating systems:** Linux

3.2.2.2 Modular description of the model

3.2.2.2.1 Aim of the model

There are four fairly independent simulations applied to produce the desired outcome of ventilation/perfusion matching within a realistic lung geometry. These are described below.

1. ***Model geometry***

We derive patient-specific geometric models from high-resolution CT data – initially one for each of the three different COPD phenotypes from the PAC-COPD database. We use segmentation software from the University of Iowa (PASS) to segment the lobe surfaces. Central airways and blood vessels are manually extracted from the CT slice data. We apply a 3D volume filling branching algorithm (Tawhai et al, 2004) to construct models of the branching structure of the airways, arteries and veins. We apply appropriate diameter ratios from measured data to each of the 1D branching models.

2. Tissue mechanics

Ideally we have CT data at two different lung volumes (the surface deformation provide boundary conditions) to enable us to predict patient-specific lung deformation. We apply a finite deformation model (Tawhai et al, 2009), assuming that lung tissue is a continuum, to predict the stresses and strains acting on the branching structures embedded within the lung volume model.

3. Ventilation

The ventilation model determines the time-averaged (static) distribution of ventilation during normal breathing throughout the conducting airways. The model includes air flow in the conducting airways coupled to lumped representations of compliant terminal acinar units, such that the ventilation distribution is governed by local tissue density and elastic recoil pressure, airway resistance and acinar compliance. The 1D Poiseuille flow equation, including an energy loss correction factor at the bifurcations, in combination with conservation of mass at each junction are used to predict the distribution of ventilation within the 1D conducting airway tree.

4. Blood flow

Arterial and venous blood flow is modelled as Poiseuille flow (Clark et al., 2011), capillary blood flow is modelled as sheet flow (Fung and Sobin, 1969), and intra-acinar (extra-capillary) blood flow is modelled as a Poiseuille flow in a symmetric ladder structure (Clark et al., 2010a). Distension of vessels and hydrostatic effects are also incorporated in the model. Predictions of flow, pressure, deformed vessels radius and sheet height, capillary blood volume and red blood cell transit times are all calculated within the vascular circuit.

The ultimate outcome from this model will be a regional description of alveolar ventilation (VA) and pulmonary blood flow (Q). These values will be calculated at ~30,000 acini within the lung volume. These are the inputs that will be fed into *Model 6: Oxygen transport and utilisation*.

3.2.2.2.2 Algorithm and parameters

The proposed simplified method to predict regional VA and Q information is as follows:

1. Extract information from a given patient's CT scan – we need the following information:

- Lung volume.
- Shape and tissue density at two different lung volumes.
- Diameters of the trachea and main pulmonary blood vessels.

This data can be measured from CT and used as input to this model. From this information we can estimate the locations of acinar units for which we want to predict VA and Q properties for.

2. Ideally we would apply values of cardiac output and/or pulmonary arterial pressure for a given patient as input to the model.

3. Ideally measurements of a patient's tidal volume would also be set as an input parameter. In the absence of this a mean population value can be used.

4. Using information about the known distribution of ventilation and blood flow and how this is affected by cardiac output, tidal volume and regional tissue density derived from CT scans – we will predict the heterogeneity and distribution of VA and Q at each acinar location. This information will be fed to Model 6.

3.2.2.2.3 Inputs and outputs

The main **inputs** are:

- ✓ CT data at two different lung volumes (lung volume, shape, tissue density, trachea and main vessel diameters)
- ✓ Cardiac output and/or pulmonary arterial pressure
- ✓ Tidal ventilation volume

The main **outputs** are:

- ✓ Ventilation (VA) at each acinus (alveolar sac)
- ✓ Blood flow (Q) at each acinus (alveolar sac)

The following table describes the main variables of the model:

Variable	Name in model	Unit	comment
Lung Volume		l	NB/ This isn't used in calculations but needs to be known (therefore doesn't have a 'name in model')
lung shape			This is the lung volume mesh so the data is a unit free vector.
tissue density	ρ	Hounsfield	Both average and regional values if possible. The regional information will depend on the resolution of the image data – but would at most be per voxel.
Tissue compliance	CA	L/mmHg	Derived from change in tissue density over two lung volumes
diameter trachea	D_{airway}	mm	Large airway diameters derived from imaging and a diameter ratio applied for distal airways
diameter vessels	D_{artery}, D_{vein}	mm	Large vessel diameters derived from imaging and a diameter ratio applied for distal vessels.
Cardiac output	Q_{RV}	ml/min	If measured this could be applied in the model, otherwise we will set an average value.
pulmonary arterial pressure	PAP	mmHg	If measured this could be applied in the model, otherwise we will set an average value.
Tidal ventilation volume	Vt	L	If measured this could be applied in the model, otherwise we will set an average value.
Ventilation alveolar sac	VA	ml/min	This value predicted at each acinus within the lung.
Blood flow alveolar sac	Q	ml/min	This value predicted at each acinus within the lung.

3.3 Model 7: Cell Bioenergetics, mitochondrial respiration and reactive-oxygen-species generation

3.3.1 Abstract

Mitochondrial respiration, dealing with transfer of unpaired electrons, may produce Reactive Oxygen Species (ROS) such as O_2^- and subsequently H_2O_2 as side products. ROS are chemically very active and can cause oxidative damage to cellular components. The production of ROS in the mitochondrial respiratory chain (RC), normally low, are primary signals that modulate cellular adaptation to environment, and are also destructive factors that damage cells under the conditions of hypoxia/re-oxygenation, i.e., a sharp increase of ROS levels is incompatible with cell survival.

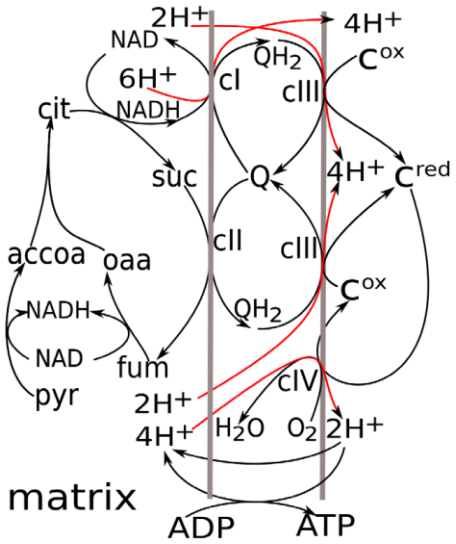


Figure 5: *Scheme for mitochondrial respiration and linked processes simulated in the model. Two reactions lead from pyruvate to succinate and further transformation to oxaloacetate reduce NAD+ to NADH. The latter is used by complex I to generate a transmembrane electrochemical proton potential (DmH+) and reduce ubiquinone (Q) to ubiquinol (QH_2), oxidation of which by complex III also contributes to DmH+. Complex III reduces cytochrome c, oxidation of which by complex IV and reduction of molecular oxygen to H_2O is also coupled to DmH+ generation. Oxidation of succinate to fumarate by complex II is coupled to the reduction of ubiquinone and thus fuels complex III. The product of electron transport, DmH+, is consumed for ATP synthesis.*

Respiration at the level of mitochondria is considered as delivery of electrons and protons from NADH or succinate to oxygen through a set of transporters constituting the respiratory chain. This model takes into account the important role of ROS, this is, the mechanisms and determinants of ROS production by extending o previous rule-based model of complex III in order to account for electron transport in the whole RC coupled to proton translocation, transmembrane electrochemical potential generation, TriCarboxylic Acid (TCA) cycle reactions, and substrate transport to mitochondria. It fits respiratory electron fluxes measured in rat brain mitochondria fuelled by succinate or pyruvate and malate, and the dynamics of NAD+ reduction by reverse electron transport from succinate through complex I.

Mathematical modelling have been used to analyse experiments with isolated brain mitochondria aimed to study the mechanism of underlying processes governing the formation of free radicals that can transfer an unpaired electron to oxygen-producing superoxide and thus can initiate the generation of ROS. Mitochondrial complex III can operate in two distinct steady states at the same micro-environmental conditions, producing either low or high levels of ROS. Here, this property of bistability was confirmed for the whole RC. Furthermore, associations between measured ROS production and computed individual free radical levels in complexes I and III have been established.

The mitochondrial respiration chain (RC) is highly related with metabolism since several metabolites and proteins are involved in both processes, furthermore the ATP produced in RC is consumed in many metabolic reactions and other ones produces ATP. To elucidate how RC and metabolism are related it is necessary an expansion of ROS model that takes into account metabolic reaction. To simulate such reactions a system of ODE's will be used. In order to simulate energetic metabolism, those equations do not only consider the kinetic of the reactions and the stoichiometry of the metabolites but also ATP, NADH and NADPH consumption and production. ROS model expansion is going to be done throughout the project and is going to include glycolysis, glycogen production, pentose phosphate pathway and TCA.

3.3.2 Model implementation

3.3.2.1 Programming Language

The software, which constructs the large Ordinary Differential Equations (ODE) system that describes the dynamics of multiple redox states of respiratory complexes based on the rules of electron transitions, and the shell performing simulations and their visualization, fitting experimental data, and statistical analysis of model parameters is written in C++.

- **Compiler version**

  Currently we use GNU C++ compiler g++ 4:4.4.3. However, the source code of the model does not have any restrictions for using different C++ compiler.

- **Required libraries**

  Standard C++ libraries

- **Supported operating systems**

  Currently we use Linux Ubuntu 10.4, but in fact it can be recompiled for other operative systems.

3.3.2.2 Modular description of the model

The ROS model can be seen as the join of two sub-models: the *Electron Chain* sub-model and the *TCA cycle* sub-model. The following subsections aims at describing the two sub-modules.

3.3.2.2.1 Electron Chain Module

3.3.2.2.1.1 Aim of the electron chain module

The aim of the electron transport chain module is to study the mechanics of electron and proton transport in mitochondria by the simulation of experimental data taking into account the details of all elementary steps of this process. Such simulations will provide a deep insight into the operation of this important cellular energy transformation necessary for ATP synthesis. The detailed model description of reactive oxygen species (ROS) production as a side product of electron transport under various conditions will give insight into the possible role of these chemically active compounds in the development of diseases under various stress conditions.

3.3.2.2.1.2 Algorithm and parameters

Electron transport proceeds through redox reactions between electron carriers subsequently carrying electrons from NADH or succinate to oxygen as a final electron acceptor. The carriers occupy fixed positions in respiratory complexes and the combinations of redox states of carriers define the redox states of whole complexes. Each carrier can be oxidized (0) or reduced (1). In this way the redox states of a whole complex can be represented by various combinations of “zeros” and “ones”. Such representation is convenient for algorithmic construction of a system of Ordinary Differential Equations (ODE) that describes the dynamics of all possible redox states of the complexes as a result of redox reactions between the carriers. The algorithm automatically constructs the ODE based on a set of rules that restrict the possible interactions between carriers. It is described in more details in our recent publication [8].

3.3.2.2.1.3 Inputs and outputs

Note: In the *electron chain module* and in general in the *skeletal muscle bioenergetics and mitochondrial reactive oxygen species generation* model all the variables con both play the role of input or output variables. This is because of the ODE system and the iterative process of fitting the experimental data. It is worth to mention that all the variables considered on the electron chain module are expressed in mg of mitochondrial protein.

Coenzymes/metabolites

Variable	Source code name	KEGG code	Description
Ubiquinone	nq	C00399	CoQ oxidized form
Ubiquinol	nqh	C00390	CoQ reduced form
ADP	nadp	C00206	Adenosine diphosphate
ATP	atp	C00002	Adenosine triphosphate
NAD+	nnad	C00003	Oxidized Nicotinamide adenine dinucleotide
NADH	nadh	C00004	Reduced Nicotinamide adenine dinucleotide
O2-	nros	C00704	superoxide anion
O2	nox	C00007	oxygen
H+	nho	C00080	external H+
H+	nhi	C00080	internal H+

Reactions

Variable	Source code name	KEGG code	KEGG Type	KEGG pathway	Description
NADH:acceptor oxidoreductase	fmnred	R00281	M00142	hsa00190	Complex I
NADH:ubiquinone oxidoreductase	redox1	R02166	M00142	hsa00190	Complex I
ubiquinone_dissociation	qpdiss1	R02163	M00142	hsa00190	Complex I
ubiquinol_binding	qpbind1	R02163	M00142	hsa00190	Complex I
Succinate dehydrogenase-SDHA	fsdh	R02164	M00148	hsa00190	Flavoprotein subunit of complex II
Succinate dehydrogenase-SDHB	fsdh	R02164	M00148	hsa00190	Iron-sulfur subunit of complex II
Succinate dehydrogenase-SDHC	fsdh	R02164	M00148	hsa00190	(Integral membrane protein CII-3) (QPs-1) (QPs1) (Succinate dehydrogenase complex subunit C) (Succinate-ubiquinone oxidoreductase cytochrome B large subunit)
Succinate dehydrogenase-SDHD	fsdh	R02164	M00148	hsa00190	(CybS) (CII-4) (QPs3) (Succinate dehydrogenase complex subunit D) (Succinate-ubiquinone oxidoreductase cytochrome b small subunit) (Succinate-ubiquinone reductase membrane anchor subunit)
iron-sulfur	shiftFeS	R02161	M00152	hsa00190	Complex III
cytochrome b	shiftFeS	R02161	M00152	hsa00190	Complex III
c1 → Cyt	c1c	R02161	M00152	hsa00190	Complex III
FeS → c1	shift1	R00082	M00152	hsa00190	Complex III
bl → bh	shift1	R02161	M00152	hsa00190	Complex III
	bhqn	R02161	M00152	hsa00190	Complex III
→ QH2_p	qphbind	R02161	M00152	hsa00190	Complex III
→ Q_n	binqn	R02161	M00152	hsa00190	Complex III
Q_p →	qpdiss	R02161	M00152	hsa00190	Complex III
QH2_n →	dissqn	R02161	M00152	hsa00190	Complex III
ATP Synthase	fatps	R00086	M00158	hsa00190	Complex V
ATPase	fatpa	R00087	M00158	hsa00190	F-ATPase
Hi-Ho exchange	flk		hsa00190	hsa00190	Exchange by diffusion

Enzymes

Variable	Source code name	UNIPROT code	KEGG Type	KEGG Pathway	Description
NADH dehydrogenase [ubiquinone] flavoprotein	fc1	P49821	M00142	hsa00190	C I: NADH dehydrogenase [ubiquinone] flavoprotein 1
NADH dehydrogenase [ubiquinone] flavoprotein	fc1	P19404	M00142	hsa00190	C I: NADH dehydrogenase [ubiquinone] flavoprotein 2
NADH dehydrogenase [ubiquinone] flavoprotein	fc1	P56181	M00142	hsa00190	C I: NADH dehydrogenase [ubiquinone] flavoprotein 3
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	Q91YT0	M00142	hsa00190	C I: iron-sulfur clusters 1
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	O75306	M00142	hsa00190	C I: iron-sulfur clusters 2
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	O75489	M00142	hsa00190	C I: iron-sulfur clusters 3
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	Q02375	M00142	hsa00190	C I: iron-sulfur clusters 4
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	O43920	M00142	hsa00190	C I: iron-sulfur clusters 5
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	Q9M9B4	M00142	hsa00190	C I: iron-sulfur clusters 6
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	O75251	M00142	hsa00190	C I: iron-sulfur clusters 7
NADH dehydrogenase [ubiquinone] iron-sulfur	qnp1	O00217	M00142	hsa00190	C I: iron-sulfur clusters 8
ND2 subunit	qnp1	P03891	M00142	hsa00190	C I: quinone-binding site
Succinate dehydrogenase-SDHA	fsdh	P31040	M00148	hsa00190	Flavoprotein subunit of complex II
Succinate dehydrogenase-SDHB	fsdh	P21912	M00148	hsa00190	Iron-sulfur subunit of complex II
Succinate dehydrogenase-SDHC	fsdh	Q99643	M00148	hsa00190	(Integral membrane protein CII-3) (QPs-1) (QPs1) (Succinate dehydrogenase complex subunit C) (Succinate-ubiquinone oxidoreductase cytochrome B large subunit)
Succinate dehydrogenase-SDHD	fsdh	O14521	M00148	hsa00190	CybS ,CII-4, Qps3. Succinate dehydrogenase complex subunit D. Succinate-ubiquinone oxidoreductase cytochrome b small subunit. Succinate-ubiquinone reductase membrane anchor subunit
Iron-sulfur	BC1	P47985	M00152	hsa00190	No Quinone binding. Complex III: Ubiquinol-cytochrome c reductase iron-sulfur subunit
Cytochrome bl	BC1	P00156	M00152	hsa00190	No Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. Low potential, heme 1, b562
Cytochrome bh	BC1	P00156	M00152	hsa00190	No Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. High potential, heme 2, b566
Cytochrome c1	BC1	Q7GXY8	M00152	hsa00190	No Quinone binding. Complex III: cytochrome C1
Iron-sulfur	BC1qn	P47985	M00152	hsa00190	N-side Quinone binding. Complex III: Ubiquinol-cytochrome c reductase iron-sulfur subunit
Cytochrome bl	BC1qn	P00156	M00152	hsa00190	N-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. Low potential, heme 1, b562
Cytochrome bh	BC1qn	P00156	M00152	hsa00190	N-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. High potential, heme 2, b566
Cytochrome c1	BC1qn	Q7GXY8	M00152	hsa00190	N-side Quinone binding. Complex III: cytochrome C1
Iron-sulfur	pBC1q	P47985	M00152	hsa00190	P&N-side Quinone binding. Complex III: Ubiquinol-cytochrome c reductase iron-sulfur subunit
Cytochrome bl	pBC1q	P00156	M00152	hsa00190	P&N-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. Low potential, heme 1, b562
Cytochrome bh	pBC1q	P00156	M00152	hsa00190	P&N-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. High potential, heme 2, b566
Cytochrome c1	pBC1q	Q7GXY8	M00152	hsa00190	P&N-side Quinone binding. Complex III: cytochrome C1
Iron-sulfur	qhpBC1	P47985	M00152	hsa00190	P-side Quinone binding. Complex III: Ubiquinol-cytochrome c reductase iron-sulfur subunit
Cytochrome bl	qhpBC1	P00156	M00152	hsa00190	P-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. Low potential, heme 1, b562
Cytochrome bh	qhpBC1	P00156	M00152	hsa00190	P-side Quinone binding. Complex III: Ubiquinol-cytochrome-c reductase complex cytochrome b subunit. High potential, heme 2, b566
Cytochrome c1	qhpBC1	Q7GXY8	M00152	hsa00190	P-side Quinone binding. Complex III: cytochrome C1
ATP Synthase	fatps	O75964	M00158	hsa00190	Complex V
ATPase	fatpa	P00846	hsa00190	hsa00190	F-ATPase

3.3.2.2.2 TCA cycle module

3.3.2.2.2.1 Aim of the TCA cycle module

The aim of TCA cycle module is to describe the dynamics of production of NADH and succinate, which are the substrates of respiratory chain, and in this way are the points of linking with the model of mitochondrial respiration. The linked respiration-TCA model can be used for the investigation of relations between TCA metabolite composition, which is essentially changes under disease conditions, and oxidative stress. This linked model will contribute to the understanding of distortions at biochemical levels that underlie systemic pathologies.

3.3.2.2.2.2 Algorithm and parameters

This module consists of a system of Ordinary Differential Equations (ODE) describing the dynamics of metabolite transformations in the biochemical reactions of TCA cycle. The classical methods of kinetic modelling were used in constructing this module. This system of ODE is described in more detail in our recent publication [8].

3.3.2.2.2.3 Inputs and outputs

Note: In the *TCA cycle* module and in general in the *skeletal muscle bioenergetics and mitochondrial reactive oxygen species generation* model all the variables con both play the role of input or output variables. This is because of the ODE system and the iterative process of fitting the experimental data. It is worth to mention that all the variables considered on the TCA cycle module are expressed in mg of mitochondrial protein.

Metabolites

Variable	Source code name object	KEGG code	Description
succinate	nsuc	C00042	
fumarate	nfum	C00122	
pyruvate	npyr	C00022	
citrate	ncit	C00158	
NAD+	nnad	C00003	Oxidized Nicotinamide adenine dinucleotide
NADH	nadh	C00004	Reduced Nicotinamide adenine dinucleotide

Reactions

Variable	Source code name object	KEGG code	KEGG Type	KEGG pathway	Description
Succinate dehydrogenase-SDHA	fsdh	R02164	M00148	hsa00020	Flavoprotein subunit of complex II
Succinate dehydrogenase-SDHB	fsdh	R02164	M00148	hsa00020	Iron-sulfur subunit of complex II
Succinate dehydrogenase-SDHC	fsdh	R02164	M00148	hsa00020	(Integral membrane protein CII-3) (QPs-1) (QPs1) (Succinate dehydrogenase complex subunit C) (Succinate-ubiquinone oxidoreductase cytochrome B large subunit)
Succinate dehydrogenase-SDHD	fsdh	R02164	M00148	hsa00020	(CybS) (CII-4) (QPs3) (Succinate dehydrogenase complex subunit D) (Succinate-ubiquinone oxidoreductase cytochrome b small subunit) (Succinate-ubiquinone reductase membrane anchor subunit)
Succinyl CoA synthetase	ftca	R00432	hsa00020	hsa00020	The GTP form is the one more commonly used in the human citric acid cycle
Aconitase	ftca	R01324	hsa00020	hsa00020	
Isocitrate dehydrogenase	ftca	R00709	hsa00020	hsa00020	
Alpha-ketoglutarate dehydrogenase	ftca	R01700	hsa00020	hsa00020	
Fumarase	fuoa	R01082	hsa00020	hsa00020	
Malate dehydrogenase	fuoa	R00342	hsa00020	hsa00020	
Citrate synthase	fcs	R00351	hsa00020	hsa00020	
Malic Enzyme	fme	R00214	hsa00020	hsa00020	NAD-dependent malic enzyme, mitochondrial

Enzymes

Variable	Source code name object	UNIPROT code	KEGG Type	KEGG pathway	Description
Succinate dehydrogenase-SDHA	fsdh	P31040	M00148	hsa00020	Flavoprotein subunit of complex II
Succinate dehydrogenase-SDHB	fsdh	P21912	M00148	hsa00020	Iron-sulfur subunit of complex II
Succinate dehydrogenase-SDHC	fsdh	Q99643	M00148	hsa00020	(Integral membrane protein CII-3) (QPs-1) (QPs1) (Succinate dehydrogenase complex subunit C) (Succinate-ubiquinone oxidoreductase cytochrome B large subunit)
Succinate dehydrogenase-SDHD	fsdh	O14521	M00148	hsa00020	(CybS) (CII-4) (QPs3) (Succinate dehydrogenase complex subunit D) (Succinate-ubiquinone oxidoreductase cytochrome b small subunit) (Succinate-ubiquinone reductase membrane anchor subunit)
Succinyl CoA synthetase-	ftca	P53597	hsa00020	hsa00020	The GTP form is the one more commonly used in the human citric acid cycle
Aconitase	ftca	A2A274	hsa00020	hsa00020	
Isocitrate dehydrogenase	ftca	P48735	hsa00020	hsa00020	
Alpha-ketoglutarate dehydrogenase	ftca	Q02218	hsa00020	hsa00020	
Fumarase	fuoa	P07954	hsa00020	hsa00020	
Malate dehydrogenase	fuoa	P40926	hsa00020	hsa00020	
Citrate synthase	fcs	O75390	hsa00020	hsa00020	
Malic Enzyme	fme	P23368	hsa00020	hsa00020	NAD-dependent malic enzyme, mitochondrial
Pyruvate monocarboxylate transporter	fpyrt	P53985	Transporter	hsa00020	
Pyruvate monocarboxylate transporter	fpyrt	O15427	Transporter	hsa00020	
Pyruvate monocarboxylate transporter	fpyrt	O15403	Transporter	hsa00020	
Pyruvate monocarboxylate transporter	fpyrt	O60669	Transporter	hsa00020	
Pyruvate monocarboxylate transporter	fpyrt	O15374	Transporter	hsa00020	

4. Preliminary Architecture for integrating the models

The aim of this Section is to establish the basis of the architecture for integrating the mechanistic models (Section 4.1). Even though, the project also considers a manual integration of the mechanistic models (Section 4.2) in order to rapidly provide an integrated model and to create the simulation environment and the web interface on top of it. In this way, when the architecture for integrating the models become available then the manually integrated model will be replaced by the new architecture.

4.1 Preliminary Architecture for integrating the mechanistic models

The aim of this Section is to establish the basis of the architecture for integrating the mechanistic models (Section 3). Even though, the project also considers a manual integration of the mechanistic models (Section 4.2) in order to rapidly provide a prototype integrated model and to create the simulation environment and the web interface on top of it. In this way, when the architecture for integrating the models become available then the manually integrated model will be replaced by the new architecture.

4.1.1 Conceptual Architecture for integrating mechanistic models

4.1.1.1 Synopsys

Our aim is to provide a method to integrate two of the mechanistic models (M1 [See Section 3.1] and M3 [See Section 3.3]) while they retain their independence as much as possible. To this end we describe a modularity approach that requests the different models to be provided in specific formats. The next sections describes the approach, while the final sections specifies details on the integration and overviews the complexities and outputs from this integrative approach.

The concepts provided here will guide us in the integration approach, however detailed numerical analysis will be performed and specifics will be updated. Deliverable 6.3, that is "*First prototype of the simulation environment*", will provide extended specifics of the integration; however the mathematical and modularity concepts are described here. We aim to use a similar concept to the one proposed in Cytosolve [7].

4.1.1.2 Description of the approach

Definition of the integration problem

We consider a set of models that needs to be integrated. In our specific case every model, which is considered a system of ODEs (ordinary differential equations), has the following attributes:

- **Coding**: the coding language used to run a simulation of the model.
- **Parameters**.
- **State variables**.
- **Initial conditions**: the initial values of the state variables.
- **Inputs**: those parameters and initial conditions for state variables that need to be instantiated in order to run a simulation.
- **Outputs**: those are the state variables which values are relevant and interesting for the researchers.

In our specific case we consider two models that are defined as:

Model 6:

- **Description**: Central and peripheral O2 transport and utilization and pulmonary gas exchange.
- **Code**: Fortran 77.
- **Details**: the code returns the steady state of a system, given a set of input parameters and initial conditions.

Model 7:

- **Description**: Skeletal muscle bioenergetics and mitochondrial reactive oxygen species generation.
- **Code**: C++
- **Details**: we are interested in the implementation that returns the steady state of the system, given a set of input parameters and initial conditions.

The specific inputs and outputs of each model is described in a different section of the deliverable.
Our aim is to integrate both models in such a way that: 1) they remain as separate codes, and 2) they are able to exchange useful information in order to find the steady state of both models simultaneously. This concepts is described in Cytosolve where models are described as black-boxes, where and input and output is defined for each box; however the basic assumptions of both approaches are different as it will be shown later.

Approach for the integration problem

Similarly to the approach proposed in Cytosolve, we consider a set of black-boxes (models) where and input and output is defined. Given a time t, each black-box returns an output based on the input given an integrative step of h; therefore the new situation of the system after time t+h. The idea behind Cytosolve considers models that evolve over the same time scale and it generates an architecture that is able to work with an arbitrary number of models.

The integration of M6 and M7 is different as the time scales of both models are different; however we retain the concept of black-box and extend the architecture scheme to models with different scales. In our case we are, by now, interested in finding steady state conditions of a system (collection of black-boxes), which allows us to use certain numerical algorithms for integration. The generic approach is described as follows:

a. A number *n* of black-boxes is defined.
b. A transfer matrix is defined: that is the pass of an output from a black-box A to black-box B, by transforming the units to adequate scales.
c. A graphical black-box interaction is defined, where two boxes shared a directed link if the output of the first appears in the output of the second.
d. We define an algorithmic approach using the black-boxes interaction map and time scales to compute a steady state for all the black-boxes at the same time.

Note that (d) reflects that integrations cannot be made blindly, but it needs expert supervision. Despite this drawback, the architecture still considers separate black-boxes, so no modification of the black-boxes´ code is under consideration in the case of an update.

Now we show an example of the architecture based on the integration of models M6 and M7 (First defined in the Synergy-COPD description of work and now reviewed in Section 0). Therefore we consider two black-boxes M6 and M7. **Figure 6** shows the relation scheme between them; this figure shows the situation where both links exist. However we will consider two options.

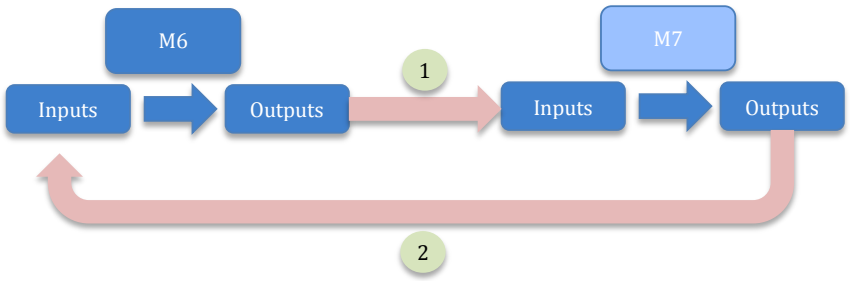


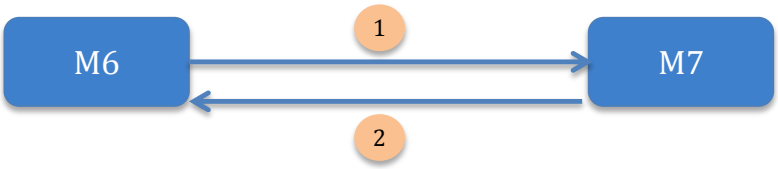
Figure 6: *M6+M7 black-box example.*

OPTION 1: ONLY LINK (1) IN FIGURE 6 EXISTS.

M6 output is used in M7 input. However as there is no feedback loop between both models, the steady states of both models can be computed by first computing M6´s and secondly M7´s. In this case we request that authors will provide a version where the output of the black-boxes is the steady state of a model given the inputs.

OPTION 2: BOTH LINKS, (1) AND (2), IN FIGURE 6 EXIST.

M6 output is used in M7 input, but in this case M7 output is used again in M6 input. We need to device a method that computes the steady state of both methods simultaneously. In this case we request that authors will provide the following versions as black-boxes: a) M6: computes the steady state given the input, b) M7: computes a time step of size h in the integrative process given an input. (This is equivalent to say that it runs a simulation for a time interval h). **Figure 7** extends this idea graphically.



1 M6´s output is the steady of the model. This output is sent to M7

2 M7´2 output is the new state of the system after a time interval h of simulation.

Figure 7: *Output definitions for each box.*

If we iterate this architecture, M6 will be updating its steady state based on M7 trajectory, and by doing this the trajectory of M7 can find a steady state if it exists. By trajectory we denote the values of the state variables over time.

Conclusions and Challenges

Model integration is not a new problem; however the most widely used solution has been to consider the manual integration of the different models. However this approach has certain drawbacks such as the lack of scalability and the complexities in updatability. We aim to use an integrative approach that overcomes these drawbacks.

We have developed a proposal that keeps the models as modules, where those modules are black-boxes; black-boxes are organized within an architecture that computes the steady state for all models simultaneously. However it requires the authors of each code to develop specific versions of their code. In the case at hand, M6 will not be modified largely, however M7 will need to be provided in such a way that run a simulation of time interval h.

The architecture provides a way to find a numerical solution of two non-independent set of ODEs, however there can be situation where: no steady state is reached by this method (as it can iterate between two models), and the bi-stability of M3 can also be identified here. We will consider the study of the numerical problems associated

4.2 Preliminary manually integrated model

The aim of this Section is to describe the manual integration of the two main models: the mitochondrial respiration and ROS production model (Section 3.3) with the model of oxygen delivery from lungs to mitochondria (Section 3.1) at a functional level. The decisions made here may or may not have an effect on the architecture for integrating the models because it should be used to describe the manual integration that will be useful for the early deployment of a first prototype. Hence the integrative architecture is going to be core of the final output of de project.

4.2.1 The point of connection

The electron flux through *complex III* to *cytochrome c* and further to *complex IV* reduces oxygen forming H_2O. This flux in the mitROS (mitochondrial ROS) model depends on electrochemical proton gradient, NADH and *succinate* as electron sources, and oxygen as substrate-acceptor.

Oxygen concentration as a one of the mitROS model variables depends on the consumption and delivery of oxygen from lungs through the circulatory system. Equation 0 is the differential equation describing oxygen levels in mitochondria, which is one of the ordinary differential equations that are constructed in the function “distr(...)” of the mitROS model source code to describe mitochondrial electron transport:

$$\frac{dO_2}{dt} = delivery - \frac{e\_flow}{4} \quad (0)$$

The rate of oxygen delivery (**delivery** in eq. 0) depends on the oxygen levels in mitochondria and it is recalculated at every time step of the mitROS model numerical solution.

4.2.2 The mechanism of connection

The function describing oxygen delivery (**delivery** in eq. 0) represents the Central and peripheral O2 transport and utilisation and pulmonary gas exchange model, which accounts for:

1. The rate of oxygen uptake in lungs (**vO_2** in eq. 1), which is proportional to the **rate of ventilation** (**k_v** from eq. 1) and the difference in concentrations of inhaled and exhaled oxygen, which in turn are proportional to the respective partial pressures in atmosphere (C_{atm} in eq. 1) and alveoli (C_A in eq. 1):

$$vO_2 = K_v \cdot (C_{atm} - C_A) \quad (1)$$

2. The alveolar diffusion, described by the following differential equation:

$$\frac{dP}{dt} = k_A \cdot (P_A - P)$$

Here **k_A** is the diffusion coefficient; **P_A** and **P** are the oxygen partial pressures in an alveolus and in a given point of lung blood circulation. This differential equation has the following analytic solution:

$$\frac{-d(P_A - P)}{(P_A - P)} = k_A \cdot dt$$

$$d(\ln|P_A - P|) = -k_A \cdot dt$$

Integration of this equation from **P_v** to **P_a** and from **t=0** to **t=T** (respective transit time) gives:

$$\frac{(P_A - P_a)}{(P_A - P_v)} = e^{-k_A \cdot T} \quad (2)$$

3. The diffusion from blood to tissue, which is described similar to alveolar diffusion:

$$\frac{dP}{dt} = \mathrm{k_m} \cdot (\mathrm{P} - \mathrm{P_m})$$

Here P_m is oxygen pressure in mitochondria. The corresponding analytical solution is the following:

$$\frac{d(\mathrm{P} - \mathrm{P_m})}{(\mathrm{P} - \mathrm{P_m})} = \mathrm{k_m} \cdot dt$$

$$d(\ln|\mathrm{P} - \mathrm{P_m}|) = \mathrm{k_m} \cdot dt$$

$$\frac{(\mathrm{P_v} - \mathrm{P_m})}{(\mathrm{P_a} - \mathrm{P_m})} = e^{\mathrm{k_m} \cdot \mathrm{T}} \quad (3)$$

4. The rate of oxygen diffusion in tissue, which is defined by the rate of blood flow (V_b) and the difference between arterial and venous oxygen concentrations:

$$\mathrm{jO_2} = (\mathrm{C_A} - \mathrm{C_V}) \cdot \mathrm{V_b} \quad (4)$$

It should be noted that if the eq. 4 refers to the whole organism, jO_2 is the total oxygen consumption, equal to VO_2 (oxygen supply by lungs, eq. 1). However, if by diffusion in tissue we mean an individual organ, then that organ will take only a part of total blood flow and, respectively, VO_2 and jO_2 will not be equal.

The fifth equation of Peter model describes mitochondrial metabolism, which we are going to be replaced with the mitROS model. This will eliminate the fifth equation and solve the other four equations to get the rate of oxygen delivery in eq. 0, as it is defined by respiration and blood circulation.

This rate of oxygen delivery is jO_2 defined by eq. 4. The other tree variables defined from solution of eq. 1-4, are P_A, P_a, P_v. This solution would allow linking mitochondrial respiration with the parameters of oxygen transport so that clinical data can be analysed.

Before solving C_A , C_a and C_v must be expressed through P_A , P_a, and P_v. Partial pressures and concentrations of oxygen in atmosphere and alveoli are linearly dependent, so that eq. 1 appears as:

$$VO_2 = k'_v(P_{atm} - P_A) \quad \textbf{(1a)}$$

where $\mathbf{k'_v}$ includes a constant of conversion between partial pressure and concentration. Since in blood hemoglobin is the main transporter of oxygen, in eq .4 the oxygen concentrations in arteries and veins are defined by hemoglobin saturation curve:

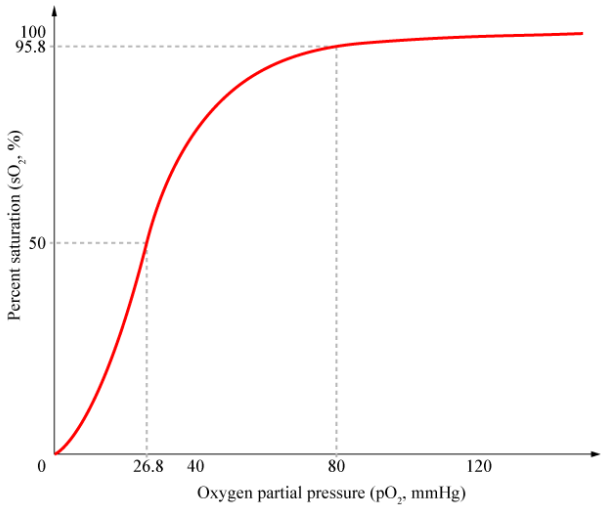


Figure 8: *Hemoglobin saturation curve.*

To be able to mathematically solve the equations the dependence between the oxygen partial pressure and concentration must be simulated by an analytical function, which admits the solution of the equations. It is sufficient that the function approximates the curve in the regions of possible physiological change of partial oxygen pressure. Since the oxygen pressure in veins never drops below **20 mmHg** and in arteries never goes above **100 mmHg**, we found a function $\boldsymbol{f(P)}$ that approximates the saturation curve in the region 20-100 mm Hg:

$$\boldsymbol{f(P) = f0 + v \cdot \frac{(P - x0)}{(k_m + (P - x0))}} \quad \textbf{(5)}$$

where **P** is partial pressure of oxygen in blood, $\boldsymbol{f0 = 0.29}$, $\boldsymbol{v = 0.83}$, $\boldsymbol{k_m = 15.0 mmHg}$ and $\boldsymbol{x0 = 20.0 mmHg}$. This function reproduces the haemoglobin saturation curve as follows:

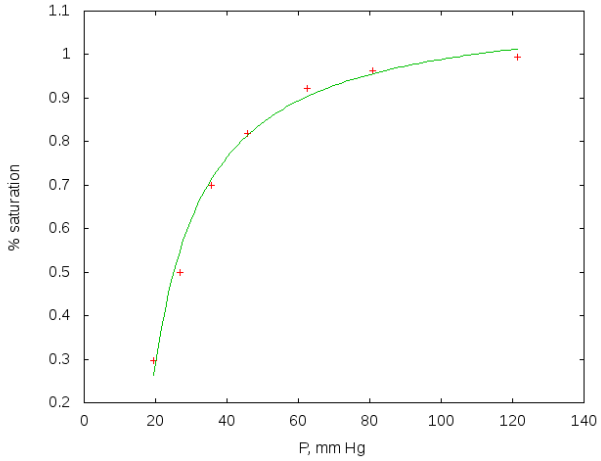


Figure 9: *In this Figure red points are taken from the saturation curve and the green line is calculated using eq.5.*

Using this curve, oxygen concentration (**C**) at a given partial pressure (**P**) can be expressed as follows:

$$\mathrm{C(P)} = \mathrm{C}_{max} \cdot f(\mathrm{P})$$

where C_{max} is oxygen concentration at maximal saturation of hemoglobin, e.g. at hemoglobin content of 150 g per litre of blood C_{max} is 203 mL of O_2/L of blood. Thus, eq. 4 can be converted:

$$\mathrm{jO}_2 = \mathrm{C}_{\max} \cdot v \cdot \left(\frac{(P_a - x0)}{k_m + (P_a - x0)} - \frac{(P_v - x0)}{k_m + (P_v - x0)} \right) \cdot V_b \qquad \textbf{(4a)}$$

After expressing oxygen concentrations through partial pressures we have the system of four equations (1a, 2, 3, 4a) with four variables (P_A, P_a, P_v, jO_2). The analytic solution of these equations is:

$$P_A = \frac{(P_{atm} - vO_2)}{k_v} \qquad \textbf{(6)}$$

$$P_a = \frac{k_v \cdot ((k_l - 1) \cdot P_{atm} + k_l \cdot (k_t - 1) \cdot P_m) + vO_2 - k_l \cdot vO_2}{(k_l \cdot k_t - 1) \cdot k_v} \qquad \textbf{(7)}$$

$$P_v = \frac{-k_v \cdot P_m + k_t \cdot (k_v \cdot ((k_l - 1) \cdot P_{atm} + P_m) + vO_2 - k_l \cdot vO_2)}{(k_l \cdot k_t - 1) \cdot k_v} \qquad \textbf{(8)}$$

$$jO_2 = C_{max} \cdot k_m \cdot (k_l \cdot k_t - 1) \cdot k_v \cdot v \cdot v_b \cdot \left(\frac{1}{k_m \cdot (k_l \cdot k_t - 1) \cdot k_v + k_v \cdot (x0 - P_m) + k_t \cdot (vO_2 - k_l \cdot vO_2 + k_v \cdot ((k_l - 1) \cdot P_{atm} + P_m - k_l \cdot x0))} + \frac{1}{k_m \cdot (k_v - k_l \cdot k_t \cdot k_v) + (k_l - 1) \cdot vO_2 + k_v \cdot (P_{atm} - k_l \cdot P_{atm} - x0 + k_l \cdot (P_m - k_t \cdot P_m + k_t \cdot x0))} \right) \qquad \textbf{(9)}$$

Where $k_l = e^{-k_A \cdot T}$ (from eq.2) and $k_t = e^{k_m \cdot T}$ (from eq.3) and **Pm** is oxygen partial pressure in mitochondria. At every time moment oxygen concentration in mitochondria is defined from the numerical solution provided by the ROS model. Oxygen pressure is defined from oxygen concentration taking into account that oxygen solubility in tissues is 0.03 mL O2/L/mm Hg.

Thus, at each step of the numerical solution of the ROS, when using the value of O_2 in mitochondria the program calls a function that return the value of **jO_2** (in accordance with eq. 9) that is used in eq. 0 as "**delivery**" to calculate the derivative of O_2 in mitochondria.

At the same time the same function gives the values of **P_A**, **P_a** and **P_v** in accordance with eq. 6-8. Initially it is necessary to provide some parameters **k_l**, **k_t** and **k_v**, but then they can be found specifically for each patient's state by fitting the measured values **P_A**, **P_a**, **P_v** and **vO_2**.

4.2.3 Conversion of units in the connected models

ROS model deals with intracellular processes which rates are expressed per mg of mitochondrial protein. The oxygen delivery is described at the levels of whole organism and organs. Special attention is needed to correlate the quantities of oxygen calculated from both models.

To convert electron flow calculated by ROS model into oxygen consumption by a specific muscle the following factors were taken into account:

1. To reduce O_2 to H_2O four electrons are needed.
2. 1 g of muscle tissue contains 53 mg of mitochondrial protein (Vinnakota et al, 2004). Thus, $\text{e-flow nmol/min/mg mitochondrial protein} = O_2 - flow \cdot 53\text{nmol}/4/\text{min/g tissue}$.
3. The mass of quadriceps muscle is 2.3 kg, thus:
   $\text{e-flow nmol/min/mg mitochondrial protein} = O_2 - flow \cdot 53 \cdot \frac{2300}{4}\text{nmol/min/muscle}$
   $= O_2 - flow \cdot 53 \cdot \frac{2.3}{4}\mu\text{mol/min/muscle}$.
4. If the volume is expressed in mL (2300 mL), this flux will give the changes in O2 concentration expressed in µmol/mL=mM.

4.3 Attributes mapping of the two main models

The semantic annotation of the mechanistic models provided in Section 3 aims at creating an semi-automated method to determine the semantic connections between models; i.e. which variables are shared by the different models that are semantically equal (see D3.1 for details).

In order to consider two variables semantically equal we have defined the following condition:

- The two set of ontological annotation must be similar:
  - They must refer to the same biological element (e.g., O_2)
  - The biological element must be located at the same biological region (i.e., same Gene Ontology annotation)
  - Ontological descriptions of the variables (e.g. *Arterial Blood Pressure*) must be inferable onto each other e.g. *Arterial Blood Pressure* is a *Blood Pressure* therefore two variables described with *Arterial Blood Pressure* and *Blood Pressure* respectively would be flagged as similar for manual evaluation of complete equivalence.
- Both variables must be expressed in the same units or a suitable conversion method must exist (i.e., the architecture for integrating the models is going to provide the unit conversion).
- Optionally the Input/Output status of the variable can be taken into consideration i.e. one variable must be an output from a model point of view whilst the other variable must be an input for a different model.

At this stage of the project we consider to semantically map the attributes of the two main models described in Section 3.1 and Section 3.3. The final outcome of the mapping process is presented in ***Table 1*** whist the details of the method used to find the correct semantic mapping is going to be provided in deliverable 3.1.

Source Code Names	Model	Description	Units	Input/Output	Ontological Annotation
PmitoO2	**Gas Exchange**	mitochondrial partial oxygen pressure	mm Hg	output	**GO:0005739 mitochondrion; Oxygen [C00007]**
nox	**ROS model**	oxygen	mg of mitochondrial protein	input	**GO:0005739 mitochondrion; Oxygen [C00007]**

Table 1: *Semantic mapping of the Central and Peripheral O_2 Transport and Utilization and Pulmonary Gas Exchange (**PmitoO2** source code name) and the Skeletal Muscle Bioenergetics and Mitochondrial Reactive Oxygen Species Generation Gas Exchange and ROS models (**nox** source code name).*

5. References

Ref. 1 Clark AR, Tawhai MH, Hoffman EA and Burrowes KS. ***The interdependent contributions of gravitational and structural features to perfusion distribution in a multiscale model of the pulmonary circulation***. J Appl Physiol 110:943-955, 2011.

Ref. 2 Tawhai MH, Hunter P, Tschirren J, Reinhardt J, McLennan G, Hoffman EA. ***CT-based geometry analysis and finite element models of the human and ovine bronchial tree***. J Appl Physiol. 97(6):2310-21, 2004.

Ref. 3 Burrowes KS, Hoffman EA, Tawhai MH. ***Species-specific pulmonary arterial asymmetry determines species differences in regional pulmonary perfusion.*** Ann Biomed Eng. 37(12):2497-509, 2009.

Ref. 4 Fung YC, Sobin SS. ***Theory of sheet flow in lung alveoli***. J Appl Physiol. 26(4):472-88, 1969.

Ref. 5 Clark AR, Burrowes KS, Tawhai MH. ***Contribution of serial and parallel microperfusion to spatial variability in pulmonary inter- and intra-acinar blood flow***. J Appl Physiol. 108(5):1116-26, 2010.

Ref. 6 A.R. Clark, M.H. Tawhai, E.A. Hoffman and K.S. Burrowes. ***The interdependent contributions of gravitational and structural features to perfusion distribution in a multiscale model of the pulmonary circulation***. J Appl Physiol 110:943-955, 2011.

Ref. 7 V. A. Shiva Ayyadurai and C. Forbes Dewey, Jr., CytoSolve: ***A Scalable Computational Method for Dynamic Integration of Multiple Molecular Pathway Model.*** Cell Mol Bioeng. 4(1): 28-45, 2011.

Ref. 8 Selivanov VA, Votyakova TV, Pivtoraiko VN, Zeak J, Sukhomlin T, et al. ***Reactive Oxygen Species Production by Forward and Reverse Electron Fluxes in the Mitochondrial Respiratory Chain***. PLoS Comput Biol 7(3), 2011.

List of Key Words/Abbreviations

ATP	Adenosine TriPhosphate
CMISS	Continuum Mechanics, Image analysis, Signal processing and System Identification
CT	x-ray Computed Tomography
Hb	Haemoglobin
Hb	HemogloBin
MAS	Model Annotation Schema
MIASE	Minimum Information About a Simulation Experiment
MIBBI	Minimum Information about a Biomedical or Biological Investigation
MIRIAM	Minimal Information Required In the Annotation of biochemical Models
mitROS	mitochondrial ROS
NADH	ubiquinona oxidorreductasa
ODE	Ordinary Differential Equation
PO2	Partial pressure of Oxygen
RC	Respiration Chain
ROS	Reactive Oxygen Species
SBML	Systems Biology Markup Language
TCA	TriCarboxylic Acid

6. Appendices

6.1 Appendix 1 – Minimal Information Standards

The two standards of primary relevance to the Virtual Physiological Human network of excellence[4] as well to the Synergy-COPD project in the context of models are Minimal Information Required In the Annotation of Models (MIRIAM) - See [biomodels.net/miriam/] (accessed 19 July 2011) and Minimal Information About a Simulation Experiment (MIASE) - See [biomodels.net/miase/] (accessed 19 July 2011)). Their key points are summarised here.

MIRIAM

The MIRIAM Standard describes information about the reference correspondence, attribution annotation, and external resource annotation.

Reference correspondence

- The model must be encoded in a public, standardized, machine-readable format.
- The model must comply with the standard in which it is encoded.
- The model must be clearly related to a single reference description. If a model is composed from different parts, there should still be a description of the derived/combined model.
- The encoded model structure must reflect the biological processes listed in the reference description.
- The model must be instantiated in a simulation: all quantitative attributes have to be defined, including initial conditions.
- When instantiated, the model must be able to reproduce all results given in the reference description within an epsilon (algorithms, round-up errors).

Attribution annotation

- The model has to be named.
- A citation of the reference description (by complete citation, unique identifier, or unambiguous URL) must be associated with the model, allowing identification of the mode's authors.
- The name and contact information of model creators (i.e. those that created this encoding of the model) must be associated with the model.
- The date and time of creation and last modification should be specified. A change history is useful but not required.
- The model should be linked to a precise statement about the terms of distribution. MIRIAM does not require "freedom of use" or "no cost" (although these are recommended by the VPH).

External resource annotation

[4] See [http://toolkit.vph-noe.eu/component/docman/doc_download/3-g02-toolkit-model-characterisation-guideline-v10]

- The annotation must enable users to unambiguously relate a piece of knowledge to a model constituent.
- The referenced information should be described using a triplet {data type, identifier, qualifier}:
    - The data type should be written as a Uniform Resource Identifier (URI).
    - The identifier is analysed within the framework of the data type.
    - Data type and identifier can be combined in a single URI, such as: urn:miriam:dataType:identifier. For example: urn:miriam:uniprot:P62158.
    - Qualifiers (optional) should refine the link between the model constituent and the piece of knowledge: “has a”, “is version of”, “is homolog to”, etc.
- The community has to agree on a set of standard valid URIs. A database and the associated API (Web Services) have been developed at the EBI to provide the generation and interpretation of URIs.

Figure 10: shows an example annotated model following the MIRIAM guidelines.

MIASE

MIASE Guidelines list the common set of information a modeller needs to provide in order to enable the execution and reproduction of a numerical simulation experiment, derived from a given set of quantitative models. They fulfil the MIRIAM requirement for reproducible simulation results (Rules for Reference Correspondence, No 6), and are composed of the three following parts described below: Information about the models to use, Information about the simulation steps, and Information about the output.

Information about the models to use

All models used in the experiment must be identified, accessible, and fully described.

- The description of the simulation experiment must be provided together with the models necessary for the experiment, or with a precise and unambiguous way of accessing those models.
- The models required for the simulations must be provided with all governing equations, parameter values and necessary conditions (initial state and/or boundary conditions).
- If a model is not encoded in a standard format, then the model code must be made available to the user. If a model is not encoded in an open format or code, its full description must be provided, sufficient to re-implement it.
- Any modification of a model (pre-processing) required before the execution of a step of the simulation experiment must be described.

Information about the simulation steps

A precise description of the simulation steps and other procedures used by the experiment must be provided.

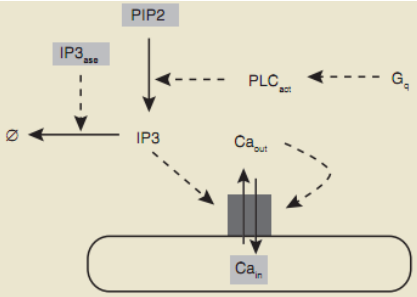
- All simulation steps must be clearly described, including the simulation algorithms to be used, the models on which to apply each simulation, the order of the simulation steps, and the data processing to be done between the simulation steps.
- All information needed for the correct implementation of the necessary simulation steps must be included, through precise descriptions, or references to unambiguous information sources.
- If a simulation step is performed using a computer program for which source-code is not available, all information needed to reproduce the simulation, and not only repeat it, must be provided, including the algorithms used by the original software and any information necessary to implement them, such as the discretization and integration methods.
- If it is known that a simulation step will produce different results when performed in a different simulation environment or on a different computational platform, an explanation of how the model has to be run with the specified environment/platform in order to achieve the purpose of the experiment must be given.

Information about the output

All information necessary to obtain the desired numerical results must be provided.

- All post-processing steps applied on the raw numerical results of simulation steps in order to generate the final results have to be described in detail. That includes the identification of data to process, the order in which changes were applied, and also the nature of these changes.
- If the expected insights depend on the relation between different results, such as a plot of one against another, the results to be compared have to be specified.

Creators	Joe User (juser@eden.com), Anne Other (aother@eden.com)
Creation date	01 January 2000
Last modification	31 May 2005



Constituent	Data type	Identifier	Qualifier	Meaning
Model	http://www.pubmed.gov/	0000000		
	http://www.ncbi.nlm.nih.gov/Taxonomy/	9606		*Homo sapiens*
	http://www.geneontology.org/	GO:0007204	IsVersionOf	Positive regulation of cytosolic [Ca^{2+}]
	http://www.geneontology.org/	GO:0051279	IsVersionOf	Regulation of release of sequestered Ca^{2+} into cytoplasm
	http://www.genome.jp/kegg/pathway/	hsa04020	IsPartOf	Calcium signaling pathway, *H. sapiens*
	http://www.genome.jp/kegg/pathway/	hsa04070	IsPartOf	Phosphatidylinositol signaling system, *H sapiens*
Compartment ER	http://www.geneontology.org/	GO:0005790		Smooth endoplasmic reticulum
Reactant Ca_{in}	http://www.ebi.ac.uk/chebi/	CHEBI:29108		$Calcium^{2+}$
Cytoplasm	http://www.geneontology.org/	GO:0005737		Cytoplasm
Reactant Ca_{out}	http://www.ebi.ac.uk/chebi/	CHEBI:29108		$Calcium^{2+}$
Reactant IP3	http://www.ebi.ac.uk/chebi/	CHEBI:16595		1D-myo-inositol 1,4,5-tris (dihydrogen phosphate)
Reactant PIP2	http://www.ebi.ac.uk/chebi/	CHEBI:18348		1-phosphatidyl-1D -myo-inositol 4,5-bisphosphate
Reactant IP3R	http://www.uniprot.org/	Q14643	HasVersion	Inositol 1,4,5-trisphosphate receptor type 1
	http://www.uniprot.org/	Q14571	HasVersion	Inositol 1,4,5-trisphosphate receptor type 2
	http://www.uniprot.org/	Q14573	HasVersion	Inositol 1,4,5-trisphosphate receptor type 3
Reactant PLC_{act}	http://www.uniprot.org/	Q9NQ66	IsVersionOf	PIP2 phosphodiesterase β-1
Reactant PLC_{tot}	http://www.uniprot.org/	Q9NQ66		PIP2 phosphodiesterase β-1
Reactant $IP3_{ase}$	http://www.uniprot.org/	Q14642		Type I inositol-1,4,5-trisphosphate 5-phosphatase
Reactant G_q	http://www.uniprot.org/	Q6NT27		Guanine nucleotide binding protein Gq
Reaction $Ca_{release}$	http://www.geneontology.org/	GO:0005220		IP3-sensitive calcium-release channel activity
	http://www.geneontology.org/	GO:0008095	IsVersionOf	IP3 receptor activity
Reaction $IP3_{production}$	http://www.geneontology.org/	GO:0004435	IsVersionOf	Phosphoinositide phospholipase C activity
	http://www.ebi.ac.uk/intenz/	3.1.4.11	IsVersionOf	Phosphoinositide phospholipase C
Reaction $IP3_{degradation}$	http://www.ebi.ac.uk/intenz/	3.1.3.56	IsVersionOf	Inositol-polyphosphate 5-phosphatase
Reaction $PLC_{activation}$	http://www.geneontology.org/	GO:0007200		G-protein signaling coupled to IP3 second messenger

$$k_1 = k_2 = k_3 = 1\ s^{-1}$$

$$Km = 10^{-7}\,M, Km_2 = 10^{-8}\,\mathrm{M}, Km_3 = 2.10^{-6}\,M$$

$$K_A = 10^{-11}, m = 4, n = 3, \alpha = 0.001$$

$$[Ca_{in}] = [IP3R] = [PLC_{tot}] = [PIP2] = [IP3_{ase}] = 0.001\,M$$

$$[G_q] = 0.01\,M,\ [Ca_{out}] = [IP3] = [PLC_{act}] = 0\,M$$

$$\frac{d[Ca_{out}]}{dt} = \frac{k_1[IP3R] * ([Ca_{in}] - [Ca_{out}])}{Km_1 + |[Ca_{in}] - [Ca_{out}]|} * \frac{[IP3]^m}{K_A + [IP3]^m}$$

$$\frac{d[IP3]}{dt} = \frac{k_2[PLC_{act}] * [PIP2]}{Km_2 + [PIP2]} - \frac{k_3[IP3_{ase}] * [IP3]}{Km_3 + [IP3]}$$

$$\frac{d[PLC_{act}]}{dt} = \frac{[G_q]^n}{\alpha + [G_q]^n} * [PLC_{tot}]$$

The model describes the release of calcium from the endoplasmic reticulum, regulated by cytoplasmic calcium and the Inositol 1,4,5-trisphosphate produced in response to G-protein–coupled receptor activation. Note that although working, this model is only meant to provide a large number of example annotations.

Figure 10: *Example of a small curated and annotated model*