Deliverable No. 2.5

Clinical relevance of the CHIC project – Describing the integrated workflows of the scenarios from a clinical perspective

Grant Agreement No.:	600841
Deliverable No.:	D2.5
Deliverable Name:	Clinical relevance of the CHIC project – Describing the integrated workflows of the scenarios from a clinical perspective
Contractual Submission Date:	31/12/2015
Actual Submission Date:	30/12/2015

Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	**CHIC**
Project Full Name:	Computational Horizons In Cancer (CHIC): Developing Meta- and Hyper-Multiscale Models and Repositories for In Silico Oncology
Deliverable No.:	D2.5
Document name:	Clinical relevance of the CHIC project – Describing the integrated workflows of the scenarios from a clinical perspective
Nature (R, P, D, O)[1]	R
Dissemination Level (PU, PP, RE, CO)[2]	PU
Version:	1
Actual Submission Date:	30/12/2015
Editor: Institution: E-Mail:	Norbert Graf USAAR graf@uks.eu

ABSTRACT:

This deliverable describes the viewpoint of clinicians as end-users of the developed hypermodels. The main intention is to guarantee usability and clinical relevance. For each of the 4 hypermodels (nephroblastoma, glioblastoma, non-small cell lung cancer and prostate cancer) it is outlined how a clinician will use the developed hypermodel as a tool for clinical decision support. These descriptions serve not only as a basis for the development of the tools but also for the demonstrators required for the upcoming reviews.

KEYWORD LIST:

Clinical relevance, integrated workflows of hypermodel scenarios, demonstrators

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 600841.

[1] **R**=Report, **P**=Prototype, **D**=Demonstrator, **O**=Other

[2] **PU**=Public, **PP**=Restricted to other programme participants (including the Commission Services), **RE**=Restricted to a group specified by the consortium (including the Commission Services), **CO**=Confidential, only for members of the consortium (including the Commission Services)

MODIFICATION CONTROL			
Version	**Date**	**Status**	**Author**
1.0	19/07/2015	Draft	Norbert Graf
2.0	24/07/2015	Draft	Caterina Guiot
3.0	20/10/2015	Draft	Joost Dejaegher, Lien Solie
4.0	01/11/2015	Draft	Norbert Graf
4.1	27/11/2015	Draft	Lien Solie
4.2	29/11/2015	Draft	Norbert Graf
4.3	02/12/2015	Draft	Norbert Graf
4.4	11/12/2015	Draft	Elias Neri, Georgios Stamtakos, Caterina Guiot, Norbert Graf
4.5	14/12/2015	Draft	Rainer Bohle, Norbert Graf
4.6	15/12/2015	Draft	Dimitra Dionysiou, Nikos Tousert, Eleni Georgiadi, Ioannis Karatzanis, Stelios Sfakianakis, Simone Bnà, Markos Antonopoulos, Eleni Kolokotroni
5.0	19/12/2015	Prefinal	Norbert Graf
6.0	28/12/2015	Final	Norbert Graf, Eleni Georgiadi, Georgios Stamatakos, Lien Solie

List of contributors

- Caterina Guiot,
- Joost Dejaegher
- Lien Solie
- Marc Stauch
- Elias Neri
- Georgios Stamatakos
- Rainer Bohle
- Dimitra Dionysiou
- Nikos Tousert
- Eleni Georgiadi
- Ioannis Karatzanis
- Stelios Sfakianakis
- Simone Bnà
- Markos Antonopoulos
- Eleni Kolokotroni

Contents

1 Executive Summary

One major outcome of the CHIC project is to show the clinical relevance of the developed hypermodels and how they will and can be used beyond the lifetime of the project. To achieve this goal two fundamental steps are necessary. The first step is to define the hypermodels with clinicians as the drivers. This includes a strong iterative process during the development with frequent feedback between clinicians and basic scientists, IT-people and modellers to guarantee that the hypermodels are clinically driven and relevant. This iterative process is already installed and shown in figure 1.1.

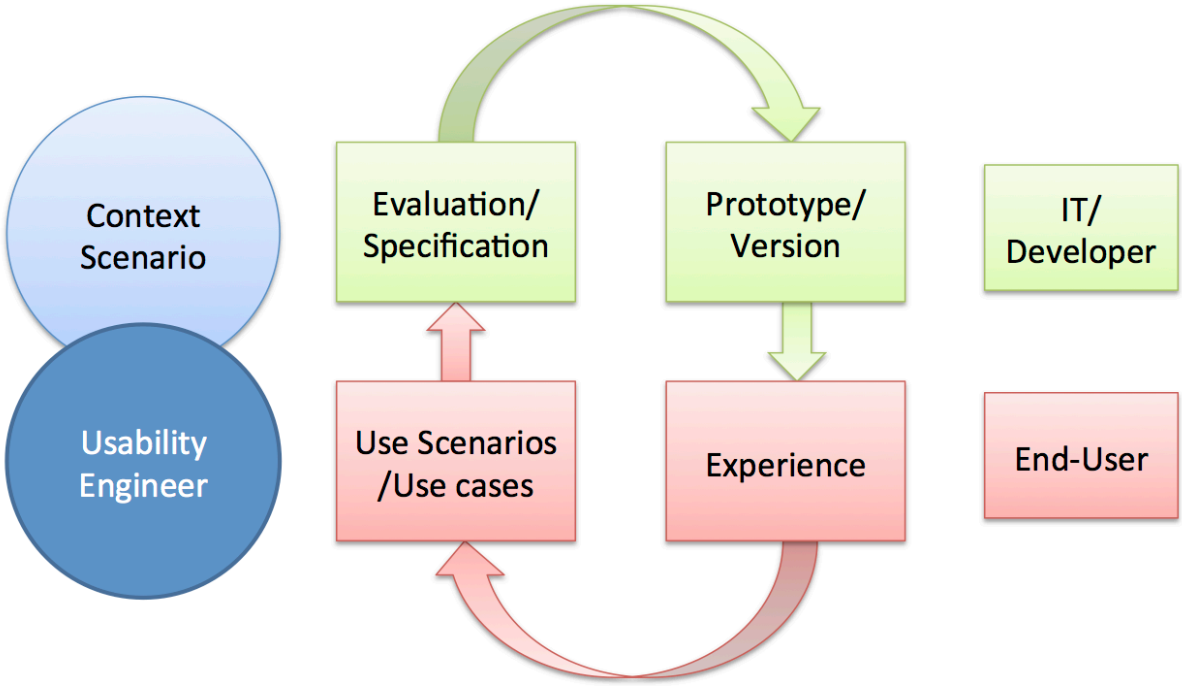


Fig. 1.1: Iterative process between developers and end-users of hypermodels.

The second step is to enrol clinicians and end-users early in the project to govern the development of hypermodels so that these hypermodels will be accepted by the scientific community and the en-users including clinicians.

This document is only dealing with the first step. It describes from a clinical perspective for the four cancer types (nephroblastoma, glioblastoma, non-small cell lung cancer and prostate cancer) how the scenarios are translated into an integrated clinically usable tool. This document is based on the clinical perspective. Basic science, IT and modelling follow this guideline to render the developed hypermodels and tools usable, executable and relevant for clinicians and other end-users.

2 Introduction

2.1 Purpose of this document

The purpose of this document is to describe the viewpoint of clinicians as end-users of the developed hypermodels. The main intention is to guarantee usability and clinical relevance. For each of the 3 main hypermodels (nephroblastoma, glioblastoma, non-small cell lung cancer) and the additional model of prostate cancer it is outlined how a clinician will use the developed (hyper)model as a tool for clinical decision support. These descriptions serve not only as a basis for the development of the tools but also for the demonstrators required for the upcoming reviews. Separate chapters will address the three different main hypermodels and the case of prostate cancer. In D2.2 the scenarios for these hypermodels are described in detail. This document delineates the interaction between the end-user (clinician) and the developed integrated tools step by step. In addition it helps to govern the iterative process in creating the tools, as the developers do know what to build to deliver integrated tools and end-users are able to test and give feedback for each of the described steps. Such an approach will guarantee usability and clinical relevance.

For each hypermodel a short overview of all clinical relevant steps of the specific scenario is given below. These steps are describing main points of interaction of a clinical end-user with the hypermodel. Points of intersection with basic scientists, IT-people and modellers are given in short descriptions that will make these hypermodels executable and usable for end-users. All tools developed by basic scientists, IT-people and modellers, as well as the ethical and legal framework at these intersection points will guarantee a smooth run of the hypermodel execution by presenting a closed workflow to the end-user with easy and understandable interactions. The first step of data creation and collection lies outside of CHIC. Nevertheless, the document describes the raw data that are needed for running the hypermodel and that need to be uploaded to the CHIC data repository.

In the following chapters at each step the necessary developments by basic scientists and IT-people are mentioned. In addition the legal and ethical considerations are shown that need to be fulfilled. These necessary developments are described from a clinical perspective and written in cursive blue letters as separate sections.

3 Nephroblastoma hypermodel

3.1 Short outline of the nephroblastoma hypermodel

The nephroblastoma scenario is described in detail in D2.2. A summary of the advanced scenario is provided in figure 5.13 of deliverable D2.2 and an update is given below in figure 3.1. All children with nephroblastoma receive pre-operative chemotherapy based on imaging studies alone. Around 10% of patients do not respond to pre-operative chemotherapy. For these patients primary surgery would be beneficial. The hypermodel therefore shall answer the question: Will a given nephroblastoma in a patient respond to pre-operative chemotherapy by tumour shrinkage, yes or no? After clinical validation of the hypermodel the tool will be used prospectively to guide the initial treatment of children with nephroblastoma.

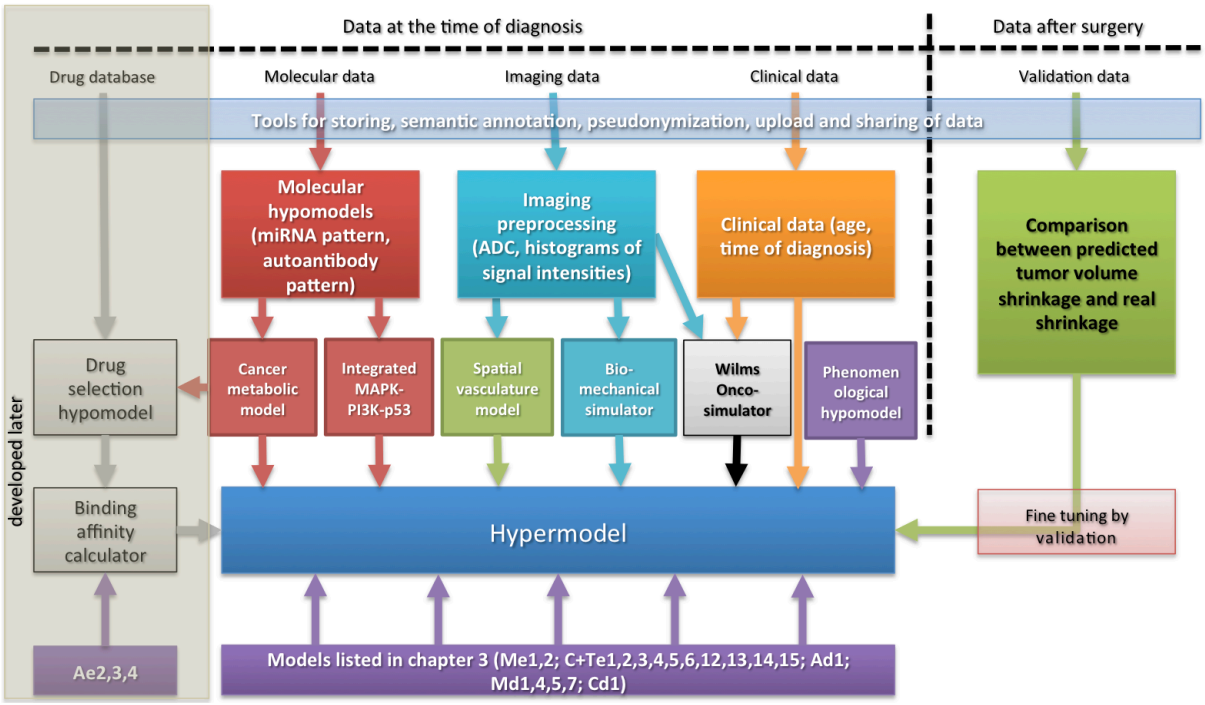


Fig. 3.1: Schematic view of the advanced nephroblastoma scenario described by the linkage of hypomodels composing the advanced nephroblastoma hypermodel

From a clinical perspective the following steps describe the nephroblastoma scenario:

1. Steps outside of the CHIC project:
    a. A specific patient with nephroblastoma is enrolled in a prospective clinical trial
    b. Clinical, imaging and molecular data are collected at the time of diagnosis
    c. Clinical, imaging and molecular data are collected after pre-operative chemotherapy before surgery
2. Steps within the CHIC project:

    All the functionalities should be provided by an end-user application named CRAF (Clinically Relevant Application Framework). The CRAF application communicates with other CHIC components in order to deliver to the clinician the desired output. The clinician interacts only

with CRAF. This includes login to the system (CRAF), selection of the cancer type (nephroblastoma), selection of the clinical question (This leads to the selection of a hypermodel), election of the patient by the clinician, (If the patient does not exist, clinician should upload through CRAF segmented images and clinical data of the new patient), running of the hypermodel by the clinician through CRAF, and visualization of results and storing in in silico trial repository. In addition the clinician should be able to retrieve the results through CRAF, without interacting with other components. Relevant steps are:

a. Data upload to the CHIC data repository
    i. A tool is needed for upload of the different data types, for guaranteeing data safety and security, to handle semantic interoperability
b. Data post-processing
    i. Tumour segmentation using DrEye needs to be done and the segmented data needs to be uploaded
c. Selection of hypomodels or a hypermodel
    i. The clinician is able to select a hypermodel composed of different hypomodels.
d. Execution of the hypermodel
    i. After the selection of the hypermodel, the CRAF checks if all data are available
    ii. If not all data are available the missing data needs to be uploaded to the CHIC data repository (step 2.a.i.) or missing data can be inserted manually
    iii. If all data are available the hypermodel will be automatically executed
e. Visualization, reporting and storage of the results
    i. Results of the hypermodel will be visualized, reported and stored in the CHIC platform linked to the specific patient
f. Validation of the results in single patients and fine-tuning of the hypermodel
    i. After the end of the pre-operative chemotherapy new data are uploaded to the CHIC data repository to allow the validation and fine tuning of the hypermodel

A detailed description of all these steps is given in the following sections of this chapter. Figure 3.3 gives a schematic overview of the clinical nephroblastoma scenario based mainly on the different data that need to be integrated and used by the hypermodel.

3.2 Gross topologies of the three multimodeller hypermodels

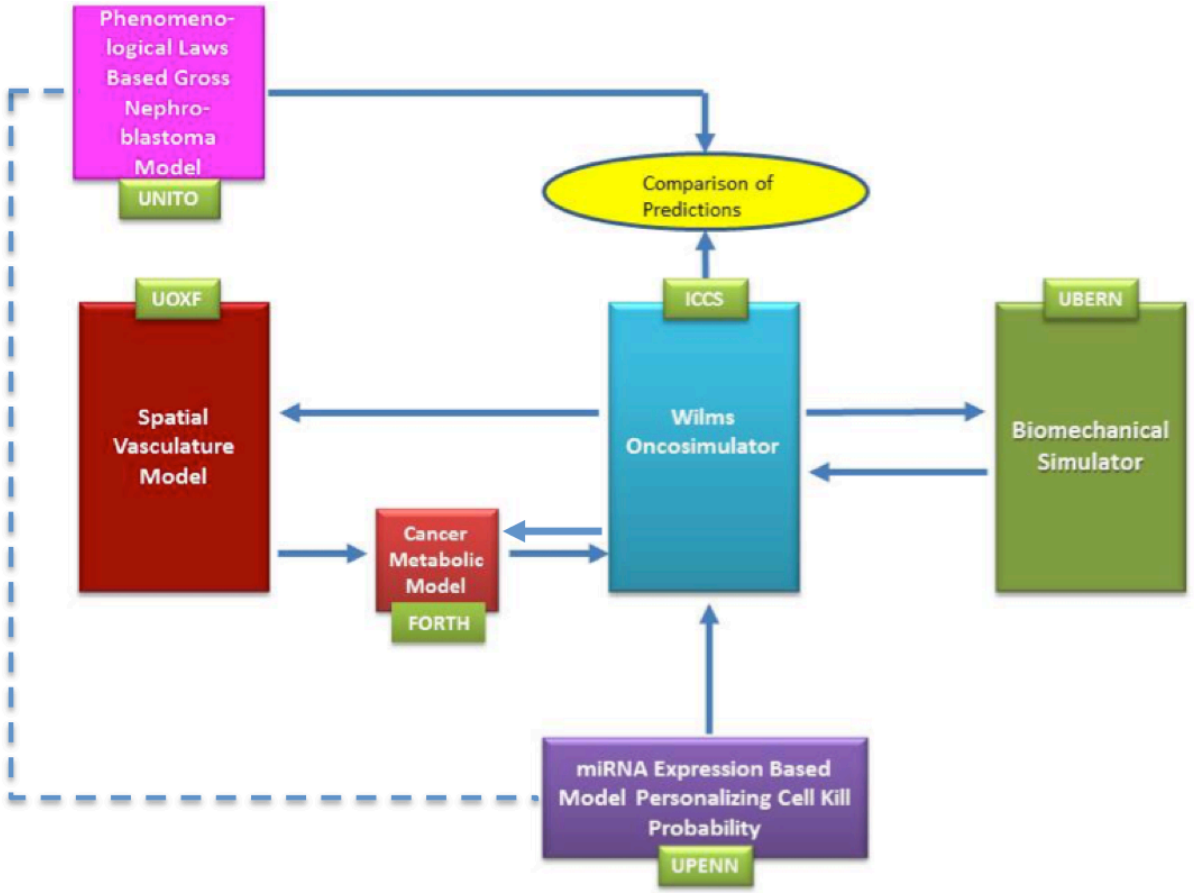


Fig. 3.2: High level topology of the CHIC multimodeller hypermodel of nephroblastoma (Wilms tumour). Basic science view.

3.3 Steps of interactions of the clinician with the integrated nephroblastoma hypermodel

A short overview of all clinical relevant steps of the nephroblastoma scenario is given above. These steps are describing main points of interaction of a clinical end-user with the hypermodel. Points of intersection with basic scientists, IT-people and modellers are given in the descriptions below that will make these hypermodels executable and usable for end-users. All tools developed by basic scientists, IT-people and modellers, as well as the ethical and legal framework at these intersection points will guarantee a smooth run of the hypermodel execution by presenting a closed workflow to the end-user with easy and understandable interactions. The first step of data creation and collection lies outside of CHIC. Nevertheless it describes the raw and mha data, which result from the production of metaimages by annotation of the images. Both raw and mha data are needed for running the hypermodel and that need to be uploaded to the CHIC data repository.

At each step the necessary developments by basic scientists and IT-people are mentioned. In addition the legal and ethical considerations are shown that need to be fulfilled. These necessary developments are important to guarantee a smooth execution of the hypermodel by end-users. They are described from a clinical perspective and written in cursive blue letters as separate sections.

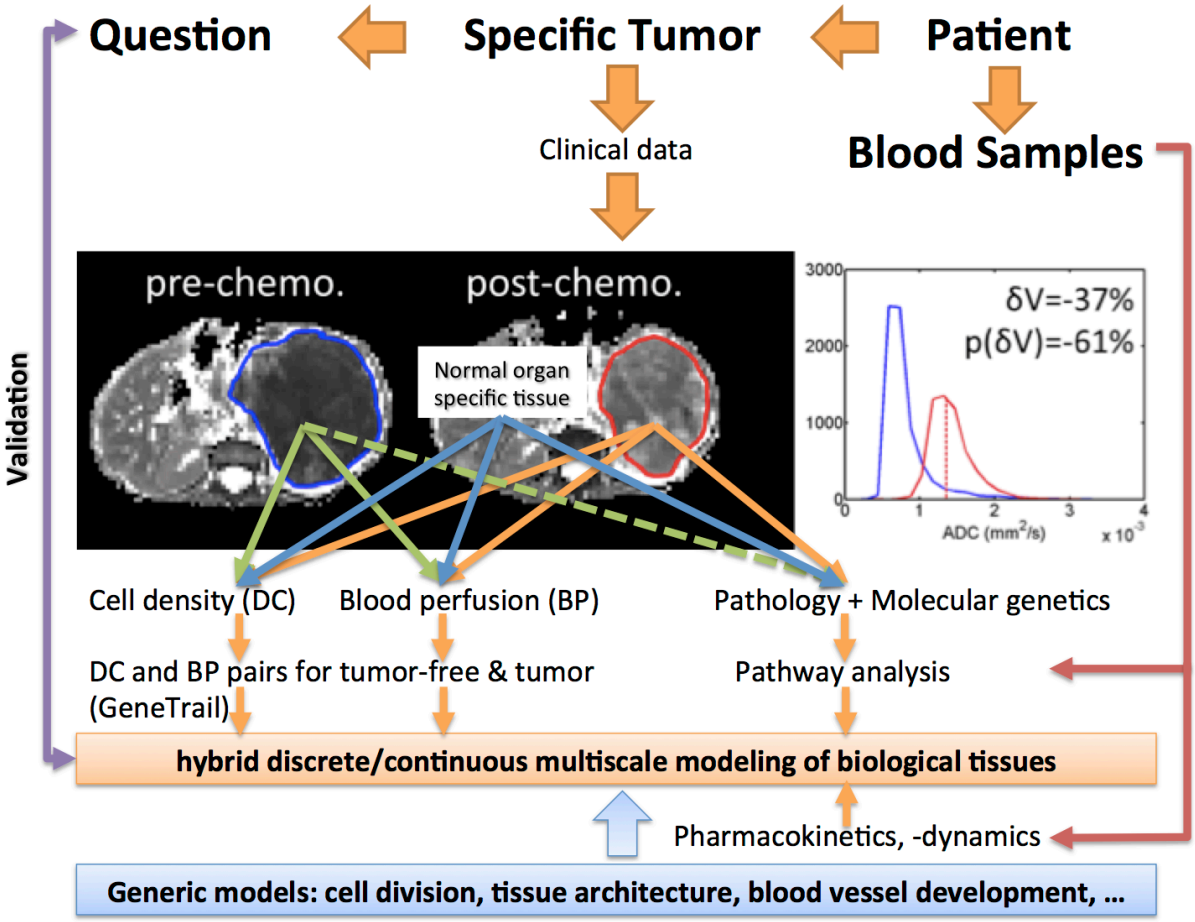


Fig. 3.3: Schematic view of the nephroblastoma scenario from a clinical perspective

3.3.1 Steps outside the CHIC platform

3.3.1.1 Data creation and collection

A new patient with a nephroblastoma will be recruited for participation in the SIOP nephroblastoma trial and study. After the end of the SIOP 2001 study the next one will be the UMBRELLA study, starting in 2016. For the nephroblastoma hypermodel the following data are needed:

1. Clinical data
2. Imaging data
3. Molecular data
4. External data (e.g. drug database, literature, KEGG database, etc.)

Clinical data will be stored in ObTiMA[3] (Ontology based Trial Management Application) as the clinical trial management system (CTMS). In ObTiMA a first pseudonymization will take place. ObTiMA will also store data of miRNA expression levels from the blood at diagnosis. Imaging data are stored as DICOM files in the local PACS (Picture Archiving and Communication System) of the treating

[3] http://obtima.org

hospitals. They are transferred with the first pseudonym to the study centre for nephroblastoma at USAAR and stored on a local DICOM server. For the CHIC project only MRI data of the different modalities are used, including (T1, T1 with contrast, T2 and DWI (Diffusion Weighted Imaging)). All external data that are used for the nephroblastoma hypermodel are defined by the different hypomodels to guarantee their scientific correct development by basic scientists (e.g. the KEEG database for the molecular hypomodel; the CHIC security framework; etc.).

3.3.2 Steps within the CHIC platform

If a clinician as an end-user will execute a hypermodel he needs primarily to enter the CHIC platform. For that purpose the IT people developed a portal in CHIC as a front-end, alongside with a set of standalone tools such as the Data Upload Tool and the CRAF[4] that support the same authentication mechanisms and unified user experience. The clinician needs to be registered in CHIC and receives credentials to enter the CHIC platform. On the portal or using CRAF he/she can easily select the hypermodel he/she wants to run. After selection he/she is guided through all steps to successfully execute the hypermodel. He/she can stop the execution at any step any time and start the execution of another model or for another patient.

Legal and ethical considerations:

The legal and ethical framework needs to be in place including necessary signed contracts as given by the legal WP. Where the validation process requires the clinician to upload and/or process patient data in personal form, the specific informed consent of the patient or his representative will be obtained. In addition the end-user will not interact with any tool guaranteeing data safety and data security that is developed by CHIC. These tools are all integrated in tools processing the data, like the upload tool and the CRAF.

Necessary developments by basic scientists and IT-people:

The CRAF needs to be in place including a roles and rights management with a log-in application and the easy selection of hypermodels to be executed. After the selection of the hypermodel the end-user will be guided automatically through all steps of the executable workflow.

3.3.2.1 Data upload

The first step is to run the data upload tool. For that purpose the data upload tool will ask to select a patient and the different data types to make an upload of data for this patient to the CHIC data repository possible. In the nephroblastoma scenario the clinical data and the miRNA data from blood are stored in ObTiMA. The imaging data are stored on a local DICOM server. All data are already pseudonymized for the first time. ObTiMA also guarantees semantic interoperability of the data. Through the CRAF application the clinician will then select a question that needs to be answered. The selection of this question results in a selection of a hypermodel.

Concrete steps are:

The user uses ObTiMA for the selection of clinical and molecular data of a specific patient, he/she has primarily chosen. The end-user can select, which of the data he/she wants to download from ObTiMA into the CHIC data repository. After selection of this data he/she starts the download -

[4] Clinical Research Application Framework

upload process within ObTiMA by clicking a button on the download page in ObTiMA. If other molecular data besides miRNA data are needed they need to be uploaded via the upload tool. The end-user then selects the imaging studies of the specific patient by the upload tool. After selection DrEye will open and the imaging data of this patient automatically uploaded into DrEye. The end-user needs to render the tumour in all modalities before the imaging data and the tumour segmentation can be uploaded to the CHIC data repository.

All data not stored in ObTiMA will not be selected by the end-user. These data are automatically selected by the hypomodels without interaction by the end-user.

Necessary developments by basic scientists and IT-people:

Single sign on (SSO) between CHIC and ObTiMA and integration of ObTiMA in the CHIC security framework is needed at a later step of the project. In ObTiMA the patient selected in the portal needs to be automatically found and the end-user will be directed to the page for download in ObTiMA. A button for download and upload to the CHIC data repository needs to be available within this page in ObTiMA. The end-user will be signed off after clicking this button, ObTiMA will then automatically close and the end-user is redirected to the upload tool, to select the imaging data of the patient. The Clinical Research Application Framework ("CRAF") is the central component to support the clinicians to perform CHIC-enabled clinical research in their premises. To this end, its user interface needs to be simple and smooth by hiding the complexity of the CHIC platform. At the same time, CRAF coordinates the functionality of other CHIC components including repositories that are also highly important for the clinicians to gain access to the CHIC services, such as the Data Upload tool for uploading patient data to the CHIC cloud, and the Visualization and image processing tools (e.g. DrEye). Such tools are in charge of running the hypermodel with the selected patient data, the model repository, the in silico trial repository, the clinical data repository and the hypermodelling editor which allows the clinician to connect hypomodels.

3.3.2.2 Post-processing of data

Regarding the nephroblastoma hypermodel imaging data needs to be post-processed to gain data of the tumour volume and the tumour composition. This post-processing of the data is done within DrEye. The imaging data are automatically opened in DrEye as described in 3.2.2.1. After tumour segmentation is finished the imaging studies, the tumour segmentation data and further statistical data of the tumour composition, done by DrEye, the end-user will store these data in the clinical data repository through CRAF. The end-user through the CRAF tool, will select the post-processed data to be uploaded and their uploading will start.

As soon as a completely automatic segmentation tool for nephroblastoma is available this tool will be integrated. This will change the workflow in a way that DrEye will be integrated after the automatic segmentation process. The end-user will select the imaging studies and the tool for automatic segmentation of the tumour. This tool will then call DrEye after finishing the segmentation process and upload the images with the segmentation for reviewing by the clinician. After the end of this review process the workflow continues as described above.

No other data do need post-processing. The interpretation of the miRNA data will be done in the molecular hypomodel.

Necessary developments by basic scientists and IT-people:

By selecting the imaging data of the patient DrEye will be opened with these specific imaging data. After the end-user has finished the segmentation process and stored these data in the clinical data

repository through CRAF. By using DrEye the imaging data as well as the segmentation data and further statistical data will be uploaded to the CHIC data repository by the CRAF and the Data upload tools.

3.3.2.3 Selection of hypomodels/hypermodel

After uploading of all data of a specific patient the end-user is guided to select hypomodels or directly a hypermodel. For that purpose different hypomodels are shown to the end-user that are related to the nephroblastoma scenario. Information describing the hypomodels is given, so that the user can easily find the hypomodel he will use. These are besides pre-processing of imaging and clinical data, the biomechanical model, the cancer metabolic model, the spatial vasculature model, the Wilms oncosimulator, the integrated MAPK-PI3K-p53 network and the phenomenological hypomodel. The end-user can select independently each alone or together with others or all hypomodels or a hypermodel. The composition of the hypermodel from the different hypomodels is predefined and depends on the input and output data of each hypomodel. The clinician is guided to the Hypermodelling Editor, where he/she can freely design a new hypermodel for the specific case. As soon as the so composed hypermodel is created the end-user will be informed and he can execute the hypermodel. Validation will compare the results given by the hypermodel for nephroblastoma and the real data for fine-tuning the hypermodel. Validation can only be done, if results from the hypermodel are available and post-processed imaging data after pre-operative chemotherapy are uploaded to the CHIC data repository for the specific patient.

On the other hand it is possible to use directly a composed hypermodel for the selected patient. This hypermodel is always composed of all hypomodels as described in the following sections.

Necessary developments by basic scientists and IT-people:

The different hypomodels, as mentioned above, need to be finalised from the basic scientific viewpoint. By selecting different hypomodels their linkage to compose a functioning hypermodel needs to be done automatically in a later stage of the project. The data of the specific patient can be retrieved automatically from the CHIC data repository. If this is not possible after a second pseudonymization the CHIC data repository needs to be divided into a clinical data repository and a research data repository. This means that the upload of specific data of a patient by the data upload tool or from ObTiMA should not perform a second pseudonymization on the data. To get these data available in the research data repository a tool needs to be build that will download the data from the CHIC clinical data repository to the CHIC research repository. In addition access rights to the CHIC clinical data repository needs to be restricted to clinicians that are allowed to run the hypermodel in the clinical setting even if the results of the hypermodel is still research.

If validation is done the workflow has to check if results of the hypermodel are available and post-processed data of imaging studies after pre-operative chemotherapy are uploaded to the CHIC data repository. Otherwise the validation tool cannot be selected. In that case the end-user will be asked to primarily run the hypermodel for this patient and if necessary to post-process the imaging DICOM files after preoperative chemotherapy. This will be done accordingly to the described process for the imaging studies at diagnosis.

Legal and ethical considerations:

If a clinical data repository is needed this implies legal considerations to guarantee data safety and privacy.

3.3.2.3.1 Subcellular Hypomodels

The end-user will not be able to interfere with the molecular hypomodels. He just can select it for usage. Basic scientists create the hypomodels. There are different possibilities to use the miRNA data within the molecular hypomodel. The first one is to characterize tumour proliferation or tumour necrosis or apoptosis by the expression levels of specific miRNAs that are related to proliferation, necrosis or apoptosis. As a result cell kill probability is calculated. For that purpose data from literature are needed that show such correlations between miRNA expression and proliferation or apoptosis and necrosis. The second possibility is to use the expression levels of the miRNAs to find deregulated molecular pathways that interfere with proliferation, necrosis or apoptosis. For that reason data from KEGG are needed and will be used in this hypomodel. The Wilms Oncosimulator, the cancer metabolic model and the phenomenological hypomodel are further hypomodels that are important for the composition of the nephroblastoma hypermodel.

Necessary developments by basic scientists and IT-people:

The molecular hypomodel are finalised for usage. Necessary external data are uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

3.3.2.3.1.1 Metabolic Hypomodel

The end-user will not be able to interfere with this hypomodel. The cancer metabolic model uses the expression levels of the miRNAs to find deregulated metabolic molecular pathways that interfere with proliferation, necrosis or apoptosis.

Necessary developments by basic scientists and IT-people:

The metabolic hypomodel is finalised for usage.

3.3.2.3.2 Super-cellular Hypomodels

3.3.2.3.2.1 Oncosimulator

The end user may interfere with this hypormodel and set multiple kinetic parameters. The oncosimulator simulates the free growth and the response of nephroblastoma tumours to chemotherapeutic schemes. This model is linked to metabolic hypomodel, vasculature hypomodel, biomechanical hypomodel.

Necessary developments by basic scientists and IT-people:

The oncosimulator (hypomodel) needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

3.3.2.3.2.2 Biomechanical Simulator

The end-user will not be able to interfere with the biomechanical simulator. He just can select it for usage. Basic scientists create the hypomodel. They do need segmentation data of the tumour and the surrounding tissue.

Necessary developments by basic scientists and IT-people:

The biomechanical simulator (hypomodel) needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

3.3.2.3.2.3 Spatial Vasculature Model

The end-user will not be able to interfere with the spatial vasculature model. He just can select it for usage. Basic scientists create the hypomodel. This model is linked to the molecular hypomodel.

Necessary developments by basic scientists and IT-people:

The spatial vasculature model needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

3.3.2.3.3 Phenomenological hypomodel

The end-user will not be able to interfere with the phenomenological model. He just can select it for usage. Basic scientists create the hypomodel.

The phenomenological (hypomodel) needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

3.3.2.4 Execution of the hypermodel as an integrated tool

After the selection of the hypermodel by the end-user this hypermodel will be executed without further intervention by the end-user. In case that the system is used in the *research mode*, all model parameters, which have assigned default values, will be possible to take new values by the user.

Necessary developments by basic scientists and IT-people:

The execution of the hypermodel needs to be fluently within a certain time accepted by clinicians. If the tool will be used in the future for decision support results should be available with a short period of time.

3.3.2.5 Visualization, reporting and storage of the results

The execution of the hypermodel will produce visualization results and a report via the CRAF. The visualization results as well as the report will be automatically stored in the CHIC in-silico-trial repository linked to the specific patient. The report contains information about who did run the hypermodel at what time with which data. The composition of the hypermodel by the different hypomodels is explained as well. All data used for the different hypomodels are displayed in tables related to the hypomodels. If the molecular hypomodel is used the deregulated pathways are shown. The final result of the hypermodel is the answer to the question, if the tumour will shrink during pre-operative chemotherapy. For that purpose the tumour volume will be displayed as a graph over the time period of pre-operative chemotherapy with a function of prediction uncertainty. The volume data are also displayed in a table. The tumour shrinkage or growth during pre-operative

chemotherapy is also shown in a 2d line, a 3d and a 4d graph/video (space and time). The end-user can always select between the report and the visualization tool.

Necessary developments by basic scientists and IT-people:

The visualization tool, the reporting document and a tool to automatically store the results in the CHIC repository need to be developed. This includes the following:

I. HYPERMODEL DESCRIPTION
Text: Description of the hypermodel
Image: Topology of the Hypermodel
Note: Both produced from the Editor

II. HYPERMODEL INPUT PARAMETERS OF THE SIMULATION
Text: Input Parameter names and values
Note: Produced from the editor

III. VISUALIZATION OF SIMULATION RESULTS

1. *Predicted change in tumor volume:*
   - ***Value**: Tumor Volume Reduction Percentage:*
   - ***2D line graph** : Time evolution of tumor volume*
2. *Predicted change in overall tumor shape:*
   - ***3D graphs: renderings** of the external shape of the initial and final simulated tumor. It would be nice if the user could "scroll" a bar to see different slices.*
   - ***3D graphs** with rotation: Animated view: same as above, but with an animated camera view that exposes concealed geometry.*
   - ***4D graphs**: Animated evolution: 3D rendering of the change in shape of the final tumor (rendered with the initial) over simulated time*

3.3.2.6 Validation of the result in a single patient and fine-tuning of the hypermodel

In case the prediction of the hypermodel needs to be validated the end-user will run the validation from the CHIC frontend (portal) or the CRAF. This will only be possible if a hypermodel for this patient has already been run and post-processed data of the imaging studies after pre-operative chemotherapy are available in the CHIC data repository. If this is not the case the end-user is not able to select the validation hypomodel from the CHIC portal. To run the hypermodel is explained above, to post-process imaging data after pre-operative chemotherapy is possible as described under 3.2.2.2. To discriminate imaging data from diagnosis from those after pre-operative chemotherapy the date the investigation of the imaging in the patient was done is used. If the data are available and also a result of the hypermodel the validation will automatically compare the predicted volume of the hypermodel with the real volume obtained after pre-operative chemotherapy and gained from post-processing of the imaging studies after pre-operative chemotherapy. If the difference of the predicted tumour volume is more than 10% of the real tumour volume a fine-tuning of the hypermodel needs to be done. It can be expected the more results of the hypermodel will be validated in different patients the better the prediction of the hypermodel will get. The validation will be reported to the end-user and also a visual comparison between the predicted and the real tumour volume will be displayed.

Necessary developments by basic scientists and IT-people:

The validation tool needs to be developed in CRAF and ready to use for clinicians. This tool is divided into the first part of just comparing the predicted result of the hypermodel with the real tumour volume after pre-operative chemotherapy. The results of this comparison need to be reported and visualized. This needs to be developed. The second part of the tool is the fine-tuning that needs to be developed by basic scientists.

3.4 Demonstrators for upcoming reviews

In the following paragraph four possible demonstrators are described. An outline of these demonstrators is given in figure 3.4 below. As the validation tool will only be needed in future demonstrations at the end of the project, no demonstration is foreseen now.

Altogether four clinical demonstrators dealing with the nephroblastoma hypermodel are possible. They should all be integrated in the CRAF:

1. Entering the CRAF and downloading clinical and miRNA data from ObTiMA into the CHIC data repository
2. Post-processing imaging data and uploading them to the CHIC data repository
3. Demonstrating the CHIC data repository with all its features
4. Selecting hypomodels or an hypermodel, executing the hypermodel and show the results of the hypermodel

Within these clinical demonstrators technical demonstrators can be implemented. One technical demonstrator might deal with the legal tools to demonstrate data safety and privacy. This might be able to do within the first demonstration of data upload into the CHIC repository and to show, if we need this the difference between the CHIC clinical and research data warehouse.

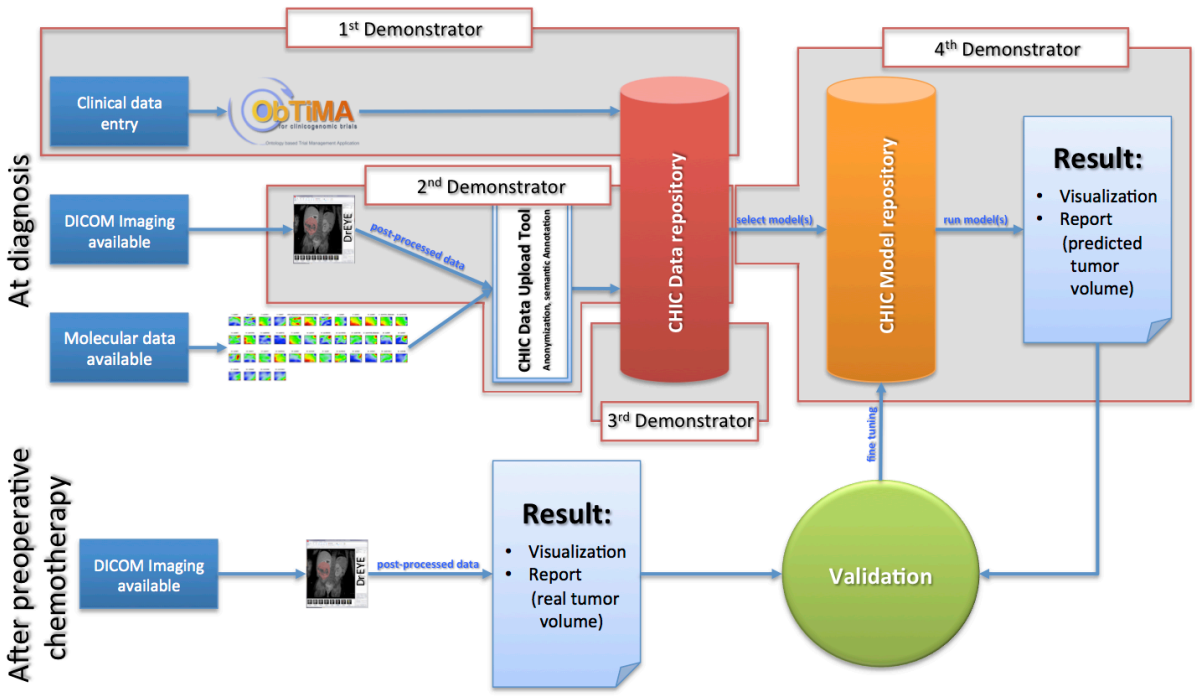


Fig. 3.4: Schematic overview of four different clinical demonstrators for the nephroblastoma hypermodel

The content of all clinical demonstrators is given in the sections above. The concrete demonstrators need to be defined between the clinicians and al other members of CHIC to get realistic demonstrators that do run smoothly. This selection is based on the availability of the different hypomodels and the possibilities of implementing such a closed workflow as written. For the nephroblastoma hypermodel all data that are needed are already available in the CHIC data repository.

3.5 Timeline for developing the hypermodel for clinical usage

An iterative process between all stakeholders of CHIC developed the timeline for the next 6 months to build all necessary components for the smooth run of the demonstrators. During this process a prioritization was established which components are most important to be finalized and which components will lack full functionality at that time. The following table outlines this timeline of tools to be developed. Red indicates the time for development and green gives the time when the tool is ready to be used. It also contains those features that will not be developed within the next 6 months.

Tools	August 2015	September 2015	October 2015	November 2015	December 2015	January 2016
Frontend / CHIC Portal						
Features to be integrated into DrEye						
CHIC data repository (clinical / research part)						
Legal tools						
Data upload tool and the CRAF						Prototype available
All hypomodels						
Visualization of result						
Report of results						
Validation of results						

4 Glioblastoma hypermodel

4.1 Short outline of the glioblastoma hypermodel

The clinical need for new therapies to treat Glioblastoma Multiforme (GBM) was described in D2.2. To summarize, current standard treatment consists of maximal safe neurosurgery, radiochemotherapy and adjuvant chemotherapy. Despite this multimodal treatment, median Progression Free Survival (PFS) is only 6.9 months and median Overall Survival (OS) only 14.6 months. Approximately 10% of patients are alive after 5 years of diagnosis. Many new treatment modalities are being explored both in a pre-clinical setting as in clinical trials. Immunotherapy, and in particular Dendritic Cell Vaccination (DC vaccination) is an emerging treatment strategy. In DC vaccination, it is aimed to activate the patient's own immune system against the tumor (more technical details are found in D2.2) to kill remaining tumor cells. By inducing immunological memory, this could theoretically lead to sustained long-term anti-tumoral protection.

At UZ/KU Leuven, DC vaccination trials have been running many years in the setting of relapsed GBM. In 2006, a trial was started in which DC vaccination was incorporated in the standard radiochemotherapy scheme. In this trial, safety and feasibility of this treatment schedule was shown. In 2010, a prospective randomized, double blinded trial was initiated in which patients with newly diagnosed GBM were randomised between DC vaccination and placebo injections. Six months after the start of the immunotherapy (i.e. after completing adjuvant temozolomide chemotherapy) there is a cross-over where placebo-treated patients can start real but late DC vaccination. The primary outcome of the clinical study was PFS at 6 months, and OS was a secundary outcome parameter. The inclusion of this trial was stopped in 2014 with 136 patients and results are expected in the near future.

During this trial multiscale data were collected for each patient: clinical data, MRI imaging data, biochemical and pathological data. These data include known clinically important parameters (e.g. extent of surgical resection) but also samples for more experimental data collection (e. g. immune monitoring) of which the exact contribution is currently unknown but part of the research question posed in the CHIC environment. All these elements of the GBM scenario will be available to compose the hypo- and hypermodels, as outlined in figure 4.4 and made available via a data management system.

We already know from the previous trials that DC vaccination is not useful for every GBM patient. Moreover, optimal timing of vaccination (i.e. during or after standard radiochemotherapy or early versus late vaccination) is currently unknown. Preliminary analyses of study patients made at our instution showed no difference in PFS or median OS between early and late vaccinated patients. However, a subgroup of patients reaches long term survival, defined as OS of more than 24 months after surgery. From a clinical point of view, it is of much interest to characterise this subgroup and be able to predict which patients are expected to reach long term survival. This is important for organisation and intensity of therapy and follow-up, counseling of patients and stratification of patients in clinical trials testing new therapies. Indeed, Glioblastoma Multiforme is probably a more heterogenous group of malignant brain cancer subtypes with different outcomes than classically described. Subtyping of patients, predicting outcome and stratification of patients is a first requirement for personalized medicine in which a more targeted therapy to a specific patient with a specific subtype of GBM can be administered. Through modelling within the CHIC environment, we want to explore if a patient with certain characteristics at diagnosis and its specific immune profile is predisposed to become a long-term survivor (or vice versa is predicted to have early relapse). This is our primary research question in CHIC. Moreover, we want to know whether immunotherapy in the form of DC vaccination is helpful to reach this long-term survivorship and if so, at if this vaccination should be given early or late.

From a clinical perspective the following steps describe the GBM scenario:

1. Steps outside of the CHIC project:
    a. At time of diagnosis: clinical, pathological (immunohistochemistry and tumor genetics), radiological (MRI) and biochemical (blood samples) data are collected.
    b. A patient with GBM is enrolled in a prospective, randomized clinical trial after surgery. Standard therapy starts and the patient is randomised between early and late DC vaccination (i.e. DC vaccination as “add-on” therapy during or after standard therapy, respectively).
    c. During the trial clinical, radiological, biochemical and therapy data (depending on the randomisation) are collected on a longitudinal base.
    d. All data are collected as source documents and are entered in a data management system with particular CRFs.
2. Steps within the CHIC project:
    a. Project compliant data management system:
        i. A system to store and manage trial data is needed. ObTiMA, a web-based GCP compliant data management system, was used to store the GBM trial data in newly created case report forms (CRFs).
        ii. The data need to be exported in a format, which is compliant with tools that are used in further steps.
    b. Pseudonymization and anonymization of data:
        i. These steps are necessary to protect patient data and to comply with the law when developing the models. In a clinical setting, when the platform is functional beyond the CHIC project, personal data will be needed.
        ii. The data are pseudonymized in the data management system and are exported likewise.
        iii. Anonymization is reached by a tool set up within the legal framework.
    c. Data upload to the CHIC data repository
        i. After the steps mentioned above, the different data types should be uploaded to the corresponding data repository by a suitable upload tool. Safety, security and semantic interoperability should be guaranteed.
    d. Post processing of data
        i. For the imaging hypomodel, post processing of imaging data will need to be performed.
        ii. First, categorical values for regional blood flow, perfusion, diffusion etc. need to be annotated by a qualified radiologist.
        iii. Additionally, computational segmentation of the images can to be done to differentiate tumor from necrotic region, oedema and normal brain tissue. Preoperative images can be segmented by an automated tool developed within the CHIC framework (BraTumIA). Segmentation during or after radiochemotherapy is more difficult due to the presence of ‘pseudoprogression’ in brain tumor regions after therapy. Tools for postoperative tumor auto-segmentation would be very helpful, but need to be developed and validated.
    e. Selection of hypomodels/hypermodel

i. The end-user is able to select a hypermodel or compose a hypermodel by connecting different hypomodels, depending on which data are available.

f. Execution of the hypermodel

i. After selecting/composing the hypermodel, it will be executed automatically for a previously selected patient.

g. Reporting and storage of the results

i. The result will be reported and stored automatically in the CHIC platform linked to the specific patient.

h. Validation of the results in single patients and fine-tuning of the hypermodel

i. During the randomized trial additional clinical, radiological, biomechanical and therapy data (depending on the randomisation) were collected on a longitudinal base. These additional data can be used to refine the result of the prediction.

ii. When the primary research question (Overall Survival at 24 months post-surgery) is known, these data can also be uploaded to the CHIC data repository to allow the validation and fine-tuning of the hypermodel.

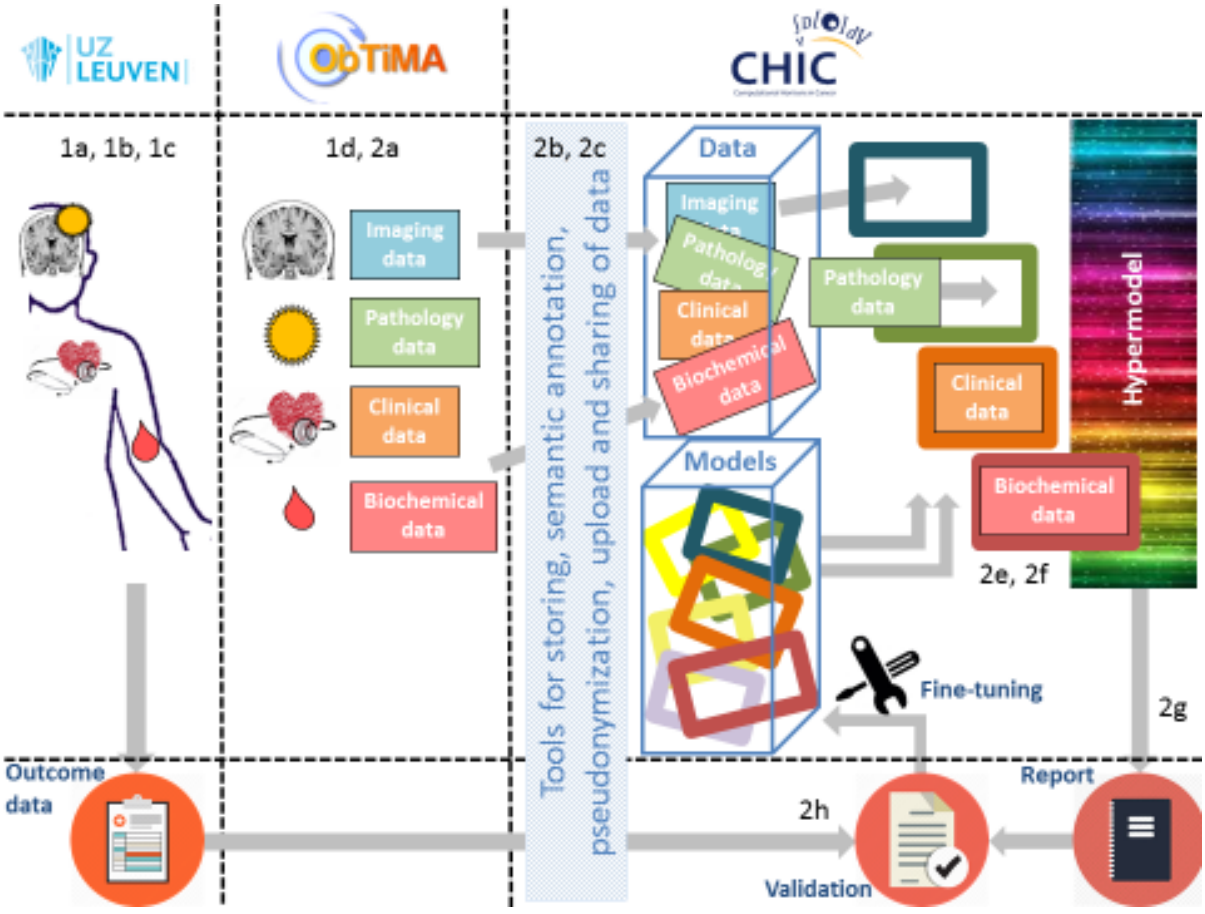


Fig. 4.1: Schematic overview of the steps in the glioblastoma scenario (with references to the points mentioned above)

4.2 Gross topology of the glioblastoma hypermodel

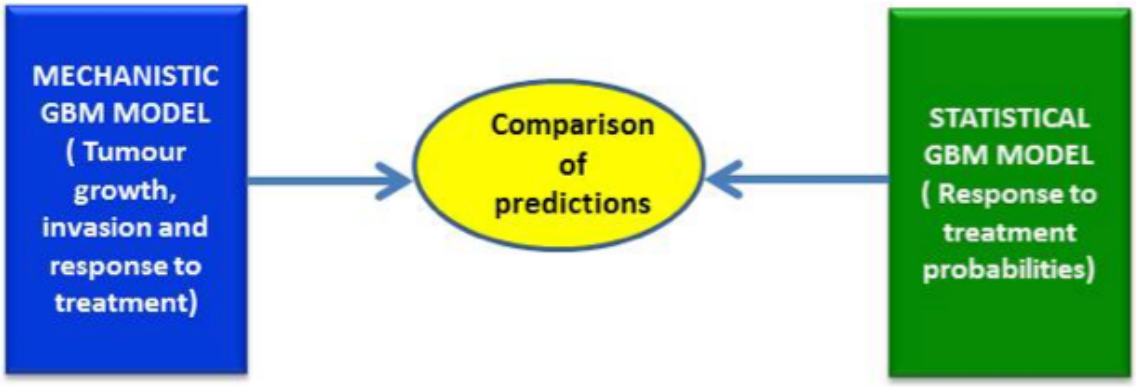


Fig. 4.2: High level gross topology of the CHIC glioblastoma hypermodel. Basic science view.

Up until now, mechanistic hypomodels simulating free growth and response to temozolomide and radiotherapy are available.

Taking account of the immune system cells system (dendritic cells, regulatory T cells, NK cells, cytotoxic T cells, T helper cells, myeloid derived suppressor cells etc.) in a mechanistic model consists in the following:

- First, at least a statistical correlation should be found on how the populations of these cells and their evolution are affected by patient, tumor, radiochemotherapy induced lymphopenia and vaccine characteristics.
- Second, at least a statistical correlation should be found on how the populations of these cells and their evolution affect tumor progression, i.e. to obtain some quantitative indications on how these cells affect the cancer cell kill ratio.

Up to now, only classic blood counts are available. Statistical analysis of these may reveal a shift in the immune system indicating qualitatively an immune response. Therefore, as yet, to include a mechanistic model regarding the evolution of immune system cells and their effect on tumor progression is feasible only in a phenomenological way.

This means that supposing that the patient has high probability of benefiting from dendritic cell vaccination and that typical cell kill rates due to chemo- and radiotherapy are adopted in the simulation, a “cell kill rate due to immunotherapy” parameter can be included in the hypomodels, taking account for the GBM cells killed by the immune system. Indications for this parameter (e.g. order of magnitude) could probably be obtained by patients in the lower arm of the study, i.e. patients who receive placebo vaccines during chemo- and radiotherapy and receive the actual vaccine after the completion of radio-chemotherapy treatment.

4.3 Steps of interactions of the clinician with the integrated glioblastoma hypermodel

An overview of all clinical relevant steps of the glioblastoma scenario is given above (figure 4.1). These steps describe main points of interaction of a clinical end-user with the platform and the hypermodel. Points of intersection with basic scientists, IT-people and modellers, that will make these hypermodels executable and usable for end-users, are given in the descriptions below. All tools developed by basic scientists, IT-people and modellers, as well as the ethical and legal framework at

these intersection points will guarantee a smooth run of the hypermodel execution, by presenting a closed workflow to the end-user with easy and understandable interactions. The first steps are data creation and collection; these raw data are needed for running the hypermodel and need to be uploaded to the CHIC data repository.

At each step the necessary developments by basic scientists and IT-people are mentioned. In addition the legal and ethical considerations are shown that need to be fulfilled. These necessary developments are important to guarantee a smooth execution of the hypermodel by end-users. They are described from a clinical perspective and written in cursive blue letters as separate sections.

4.3.1 Steps outside the CHIC platform

4.3.1.1 Patient enrolment and randomization

A double blinded, randomized trial for primary GBM patients was started at UZ Leuven in 2010. Patients were treated with standard treatment consisting of maximal safe neurosurgery, followed by 6 weeks of radiotherapy and concomitant temozolomide chemotherapy, followed by an additional 6 cycles of adjuvant temozolomide. Immunotherapy delivered as DC vaccination is randomised between early (during standard therapy) or late (after standard therapy) vaccination. In the early vaccination group, vaccination starts with 4 induction vaccines (dendritic cells loaded with tumor lysate) after radiochemotherapy and before adjuvant temozolomide. After induction vaccines, boost vaccines consisting of tumor lysate are administered between the cycles of temozolomide. In the late vaccination group, 4 induction vaccines followed by lysate vaccines are started after completing the standard therapy.

4.3.1.2 Data creation and collection

For included patients, the following data are collected at time of diagnosis and inclusion:

- Clinical data: baseline characteristics (age, sex, presenting symptoms, comorbidity and medication intake) and data after neurosurgery (extent of resection, clinical condition (Karnofsky Performance Scale and Mini-Mental State Examination) and RPA classification (Recursive Partitioning Analysis, the stratification parameter for randomisation)
- Imaging data: the pre- and postoperative MRI are annotated with categorical values (volume, contrast enhancement, perfusion, diffusion and oedema) by a qualified radiologist
- Pathology data: paraffin slices of the resected tumor are available for immunohistochemical analysis and genetic analysis
- Biochemical data: blood samples for standard cell counts and experimental research (serum and Peripheral Blood Mononuclear Cells (PBMCs)) are taken at time of leukapheresis as a reference point for further analysis (see biochemical data during treatment)

During treatment more data are collected:

- Standard therapy information: the course of radiotherapy, concomitant chemotherapy and adjuvant chemotherapy
- 'Product quality' parameters of the autologous elements of the immunotherapy: composition of the leukapheresis collection, cell counts and viability during cell culture, FACS analysis of the tumor-loaded dendritic cells and lysate concentration
- Imaging data: MRI images are taken at scheduled time points or at clinical need and should be annotated with categorical values (volume, contrast enhancement, perfusion, diffusion and oedema) by a qualified radiologist afterwards

- Biochemical data: Blood samples are taken at regular intervals and PBMCs are isolated: multicolour flow cytometry will be used to assess the different immune cell populations and their relevant activation/suppression markers (e.g. PD-1). Also serum samples are collected: relevant cytokines (e.g. VEGF) will be measured from these serum samples by Cytometric Bead Assay (CBA).
- Immune profile, i.e. data concerning populations of cells of the immune system would allow statistical analysis aiming at correlating those populations and/or their evolution during treatment with the outcome; overall survival (OS) or progression free survival (PFS).
- General follow up:
    - 'Complete blood count' values from the samples taken at regular time points during the trial (cfr. Biochemical data) and when available also during the temozolomide treatment
    - patient self-assessment on general abilities, self-reliance and quality of life at every vaccination time point
    - relevant medication intake during the course of the treatment, especially immune interfering medication
    - relevant adverse events at clinical examinations
- Outcome data (which drive the clinical question): Progression Free Survival status after 6 months of immunotherapy (after 6 cycles of adjuvant temozolomide chemotherapy), progression date, salvage therapy after progression, Overall Survival date

All data from the different participating reference hospitals are collected and stored as source documents at the local study centre, the UZ Leuven hospital. During the lifetime of the CHIC project the data were entered in the project's data management system in specific CRFs in order to have a project-compliant system. Access to these patient data (in the data management system) is restricted to the UZ Leuven hospital's trial clinicians.

Fig. 4.3: The course of data collection and storing: from Belgian reference hospitals to the trial centre at UZ Leuven and in the data management system.

4.3.1.3 Data Management System

All data mentioned above will be stored in ObTiMA (Ontology based Trial Management Application), the clinical trial management system provided by USAAR. A new set of CRFs were developed to store the glioblastoma data for the CHIC project. CRFs for pathology and biochemical data will be developed when the experimental work is finalized and the data are available. Access to the glioblastoma data is restricted to the UZ Leuven hospital's trial clinicians and assistants.

MRI images are stored as DICOM files in the PACS (Picture Archiving and Communication System*)* at UZ Leuven hospital.

Name	Version	Description
Adverse events	1.0019	Adverse events during trial
Baseline Patient Characteristics	1.0153	Patient information Course of disease
Chemotherapy	1.0009	Adjuvant chemotherapy
Immunotherapy	1.0163	Leukapheresis Dendritic cell quality Lysate vaccines
Medication	1.0059	Relevant medication during trial
Patient Blood Counts	1.0081	Blood values during trial
Patient Questionnaire	2.0001	Fertigkeitenskala Münster-Heidelberg Karnofski Performance Scale EORTC QLQ-C30 BCM20
Peri-operative radiology	1.0051	Pre-operative imaging Post-operative MRI
Radiochemotherapy	1.0040	Radiotherapy Concomitant chemotherapy
Trial radiology	1.0018	Imaging during the clinical trial
Vaccination	1.0033	Vaccination information

Fig. 4.4: Overview of the CRFs in ObTiMA for the GBM data in CHIC

4.3.2 Steps within the CHIC platform

If a clinician as an end-user will execute a hypermodel he needs primarily to enter the CHIC platform. For that purpose the IT people developed CRAF as a front-end. The clinician needs to be registered in CHIC and receives credentials to enter the CHIC platform. Via CRAF he/she can easily select the hypermodel he/she wants to run. After selection he/she is guided through all steps to successfully execute the hypermodel. He/she will be notified, if the execution finished and a result is available.

Legal and ethical considerations:

The legal and ethical framework needs to be in place including necessary signed contracts as given by the legal WP. Insofar as the validation process requires the clinician to upload and/or process patient data in personal form, the specific informed consent of the patient or his representative will be obtained. In addition the end-user will not interact with any tool guaranteeing data safety and data security that is developed by CHIC. These tools are all integrated in tools processing the data, like the upload tool.

Necessary developments by basic scientists and IT-people:

A CHIC portal as a frontend needs to be in place including a roles and rights management with a log-in application and the easy selection of hypermodels to be executed. After the selection of the hypermodel the end-user will be guided automatically through all steps of the executable workflow.

4.3.2.1 Data management system

A clinician has to be able to present his data as input for the models in a form that is compliant with the tools of the CHIC platform. The GCP-approved web-based data management system ObTiMA foresees in this compliance by an export function, which generates a CDISC ODM or CSV file. This export is another way to provide the data for upload, but less straightforward than the direct upload out of the system (see 4.2.2.3 Data upload). Providing the data via an export will only be done in the stadium of making the models.

ObTiMA also guarantees semantic interoperability of the data.

Necessary developments by basic scientists and IT-people:

The clinician needs secure access to only his own patient information in ObTiMA. If the indirect transfer (i.e. via export) of data into the CHIC platform is used, the clinician needs to be able to export the data out of ObTiMA to upload the generated file to the platform. Access to the glioblastoma data is restricted to the UZ Leuven hospital's trial clinicians.

Also, the sustainability of ObTiMA needs to be guaranteed.

4.3.2.2 Pseudonymization and anonymization

These steps are necessary to protect patient data and to comply with the law when developing the models.

A first pseudonym is added to the data in ObTiMA. When the data are uploaded as CDISC ODM or CSV file, they are first exported with this pseudonym as only identifier.

The MRI images are downloaded from the UZ Leuven PACS. The images themselves are anonymized by the download tool, but the package of images gets the same pseudonym as the according patient.

All data undergo de-identification by a tool provided by Custodix when uploaded to the CHIC platform.

In a clinical setting, when the platform is functional beyond the CHIC project, personal data will be needed: when a clinician runs the hypermodel the pseudonym will be replaced by the patient's name.

Legal and ethical considerations:

A tool for anonymizing the pseudonymized data and images needs to be developed. This tool needs to be customized according to the specific data that will be generated for each cancer type. This tool needs to run automatically when the data are uploaded to the CHIC data repository, without interaction of the user.

4.3.2.3 Data upload

For uploading data the data upload tool will ask to select a patient and the different data types to make an upload of data for this patient to the CHIC data repository possible. All data are already pseudonymized for the first time, but when uploaded the next step of anonymization takes place through the upload tool. When data are created and available in the CHIC clinical data repository, according to the available data many hypermodels can be selected.

By selecting the type of data that are available at diagnosis the user is forwarded and automatically logged in into ObTiMA and automatically forwarded to the download page, where the specific patient is automatically selected. The end-user can select on this page, which of the data he wants to download from ObTiMA into the CHIC data repository. After selection of this data he starts the direct download - upload process within ObTiMA by clicking a button on the download page in ObTiMA. After pressing this button he will be automatically logged-out in ObTiMA and ObTiMA will close. The end-user can then also select the imaging studies of the specific patient by the upload tool.

Necessary developments by basic scientists and IT-people:

Single sign on (SSO) between the CHIC portal and ObTiMA is needed at a later step of the project. In ObTiMA the patient selected in CRAF needs to be automatically found and the end-user will be directed to the page for download in ObTiMA. A button for download and upload to the CHIC data repository needs to be available within this page in ObTiMA. The end-user will be signed off after clicking this button, ObTiMA will then automatically close and the end-user is redirected to the upload tool, to select the imaging data of the patient. Also not-segmented images should be able to be uploaded without going through BraTumIA first.

Legal and ethical considerations:

Data upload is only possible by partners who have an ObTiMA account with access to the specific data that need to be uploaded.

4.3.2.4 Post-processing of data

Post-processing of data is only needed for imaging data, this can be done by BraTumIA in the CHIC platform. BraTumIA is validated to auto-segment pre-operative brain images and calculates the volumes of contrast and non-contrast enhancing tumor, necrosis and peri-tumoral oedema for these images.

During and after radiochemotherapy, assessing the residual or recurrent tumor volume is difficult because of the occurrence of 'pseudoprogression' in a subset of treated brain tumors. This pseudoprogression shows contrast capitation and hence mimics true progression. Software to auto-segment these MRIs and differentiate between true progression and pseudoprogression would be very helpful from a clinical point of view.

Pathological, genetic and biochemical data will be analysed in the respective hypomodels but do not require post-processing.

Necessary developments by basic scientists and IT-people:

A tool for auto-segmentation of the (longitudinal) brain tumor images should be available for clinicians in and out of the CHIC platform. When in the upload tool, the images of a specific patient can be selected and will automatically be uploaded into BraTumIA. After auto-segmentation the data and images are uploaded to the CHIC data repository.

Legal and ethical considerations:

The imaging data and segmentation data have to be transferred to the CHIC repository in a secure way.

4.3.2.5 Selection of hypomodels/hypermodel

After uploading of all requested data of a specific patient the end-user is guided to connect hypomodels or directly a hypermodel. For that purpose different hypomodels are shown to the end-user that are related to the GBM scenario. These are the clinical, pathological, biochemical and imaging hypomodels. The end-user can select independently each alone or together with others or all hypomodels. The hypomodels are connected in the Editor, where he/she can freely design a new hypermodel for the specific case after the data of the specific patient are automatically requested by the hypomodels from the CHIC data repository. As soon as the so composed hypermodel is created the end-user can execute the hypermodel.

Based on the data available at diagnosis the following hypomodels can be selected:

- Clinical hypomodel ← baseline characteristics, surgery data, RPA
- Imaging hypomodel ← pre-operative and postoperative imaging data
- Pathology hypomodel ← immunohistochemical and genetical data
- Biochemical hypomodel ← blood count, serum for CBA and PBMCs for FACS at leukapheresis

For the research question in the GBM scenario it would be useful if the prediction can be reassessed based on new data that become available during the course of treatment. The end-user has to run through the same procedure as described in the previous sections.

If the concerned data have a predictive value, the following hypomodels can be added to the hypomodels selected at diagnosis to compose a more extensive hypermodel:

- Clinical hypomodel ← patient self-assessment, relevant medication and relevant adverse events during treatment
- Standard therapy hypomodel ← radiotherapy, concomitant and adjuvant chemotherapy
- Imaging hypomodel ← imaging data at scheduled time points and clinical need

The following hypomodels are under discussion if feasible by data provision:

- Immunotherapy hypomodel ← leukapheresis collection, cell culture data, FACS DCs, lysate
- Biochemical hypomodel ← blood count, serum for CBA and PBMCs for FACS at scheduled time points

The validation tool will automatically compare the results given by the hypermodel for GBM and the real data for fine-tuning the hypermodel. This tool can only be selected, if results from the hypermodel are available and the real outcome data are uploaded to the CHIC data repository for the specific patient (see bottom of figure 4.1).

On the other hand it is possible to use directly a composed hypermodel for the selected patient. This hypermodel is always composed of all hypomodels.

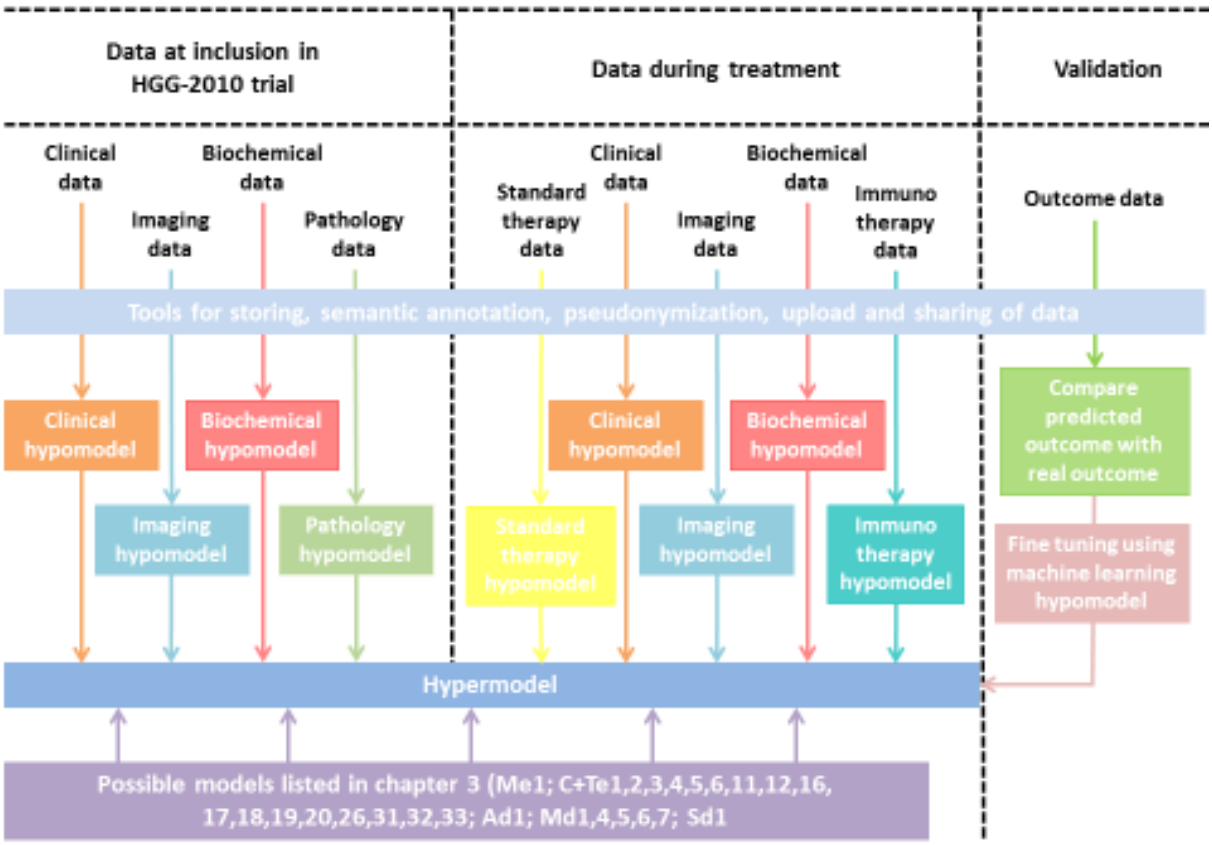


Fig 4.4: Schematic view of the glioblastoma scenario described by the linkage of hypomodels composing the glioblastoma hypermodel (Adopted and adjusted from D2.2)

Necessary developments by basic scientists and IT-people:

The hypomodels need to be finalised from the basic scientific viewpoint. By selecting different hypomodels their linkage to compose a functioning hypermodel needs to be done semi-automatically. The data of the specific patient can be retrieved automatically from the CHIC data repository. If this is not possible after a second pseudonymization the CHIC data repository needs to be divided into a clinical data repository and a research data repository. This means that the upload of specific data of a patient by the data upload tool or from ObTiMA should not perform a second pseudonymization on the data. To get these data available in the research data repository a tool needs to be build that will download the data from the CHIC clinical data repository to the CHIC research repository (see figure 4.5). In addition, access rights to the CHIC clinical data repository need to be restricted to only the hospital's trial clinicians that are allowed to access the glioblastoma patient's data and run the hypermodel in the specific clinical setting, even if the results of the hypermodel is still research. Presently it is not envisaged that for the glioblastoma scenario personal data will be used; all data can be used de-identified/anonymous.

The data that have a predictive value in the GBM scenario should be determined out of all data that can become available during treatment. Hypomodels based on these predictive data need to be composed and an end-user should be able to re-run the hypermodel with inclusion of these hypomodels with predictive value during the course of the treatment to reassess the prediction that was made based on the data that were available at diagnosis.

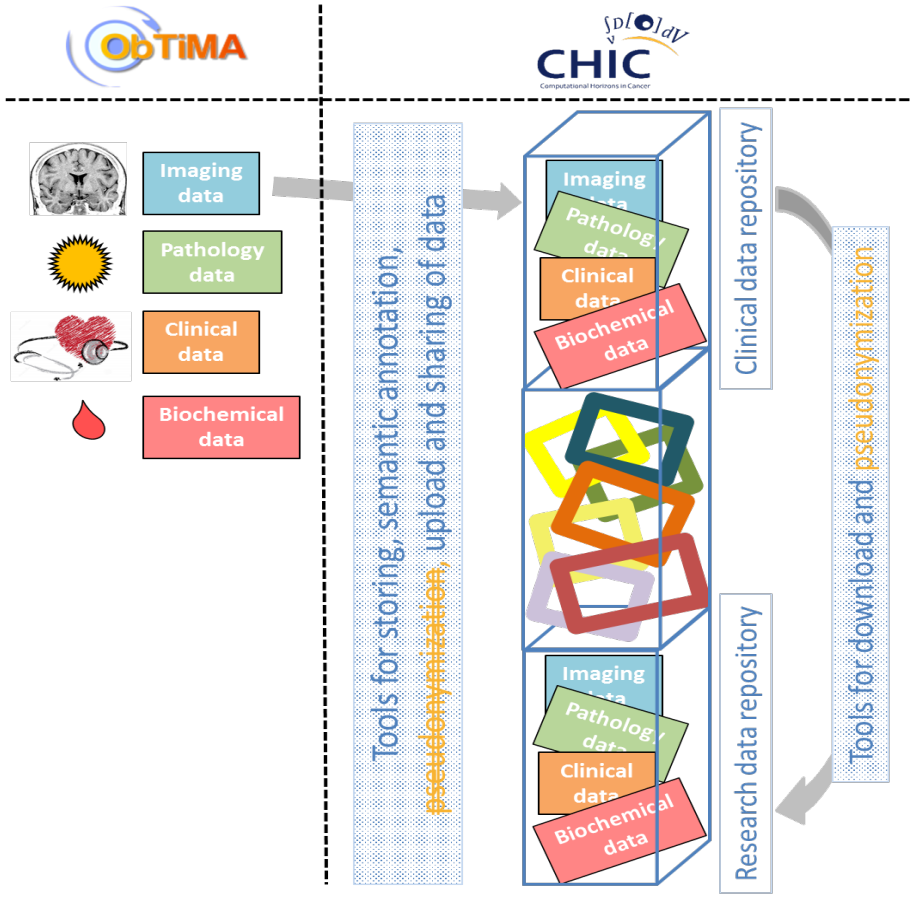


Fig. 4.5: Detail from figure 4.1 adopted to the situation when it is not possible to automatically retrieve the data of a specific patient from the CHIC data repository. Then, the CHIC data repository needs to be divided into a clinical data repository and a research data repository.

If the validation service will be used the workflow has to check if results of the hypermodel are available and GBM outcome data are uploaded to the CHIC data repository. Otherwise the validation tool cannot be selected. In that case the end-user will be asked to primarily run the hypermodel for this patient and if necessary to provide the outcome data.

Legal and ethical considerations:

If separate data repositories are needed this implies legal considerations to guarantee restricted access, data safety and privacy.

4.3.2.6 Execution of the hypermodel as an integrated tool

After the selection of the hypermodel by the end-user this hypermodel will be executed without further intervention by the end-user.

Necessary developments by basic scientists and IT-people:

The execution of the hypermodel needs to be fluently within a certain time accepted by clinicians. If the tool will be used in the future for decision support results should be available with a short period of time.

4.3.2.7 Reporting and storage of the results

The execution of the hypermodel will produce a report, which will be automatically stored in the CHIC in-silico-trial repository linked to the specific dataset/patient. The report contains information about who did run the hypermodel at what time with which data. The composition of the hypermodel by the different hypomodels is explained as well. All data used for the different hypomodels are displayed in tables related to the hypomodels.

The final result of the hypermodel is the answer to the question which patients will become long term survivors (defined as OS > 24 months), which patients are predicted to suffer early relapse and death (OS <12 months), whether immunotherapy in the form of DC vaccination is helpful to reach long term survivorship and if so, at if this vaccination should be given early or late.

Necessary developments by basic scientists and IT-people:

The reporting document and a tool to automatically store the results in the CHIC repository need to be developed.

Legal and ethical considerations:

Legal considerations to guarantee restricted access, patients' data safety and -privacy of the report and the storage of the results are needed.

4.3.2.8 Validation of the result in a single patient and fine-tuning of the hypermodel

In case the prediction of the hypermodel needs to be validated the end-user can select from the CHIC front-end (portal or CRAF) the validation tool. This will only be possible if a hypermodel for this patient has already been run and outcome data (survival, immune profiles) are available in the CHIC data repository. If this is not the case the end-user is not able to select the validation tool from the CHIC portal or CRAF. If both elements are available the validation tool will automatically compare the predicted outcome of the hypermodel with the real outcome. If the prediction is not correct, a fine-tuning of the model needs to be done. Basic scientists using also machine-learning programs will carry out this task. It can be expected the more results of the hypermodel will be validated in different patients the better the prediction of the hypermodel will get. The validation will be reported to the end-user and also a comparison between the predicted and the real outcome will be displayed.

Necessary developments by basic scientists and IT-people:

When the outcome data are not available, or when the hypermodel has not been executed before for that patient, it should not be able to select the validation service.

The validation tool needs to be developed and ready to use for clinicians. This tool is divided into the first part of just comparing the predicted result of the hypermodel with the real overall survival. The results of this comparison need to be reported. This needs to be developed. The second part of the tool is the fine-tuning that needs to be developed by basic scientists.

Legal and ethical considerations:

Legal considerations to guarantee restricted access, patients' data safety and -privacy of the validation of the results are needed.

For the glioblastoma scenario this validation can only be done on retrospective data because there is no trial running to collect prospective data from. So presently it is not envisaged that the glioblastoma hypermodel will be validated using personal data, only de-identified/anonymous data.

4.4 Demonstrators for upcoming reviews

For the GBM scenario several demonstrators are possible for an upcoming review. The demonstrators with the most clinical relevance are mentioned here, basic science and IT background can be integrated per demonstrator. The validation tool can only be demonstrated in a later stage of the project. The concrete demonstrators need to be defined between the clinicians and all other members of CHIC to get realistic demonstrators that run smoothly.

1. Starting from data that are available in the data management system; entering CRAF or the CHIC portal and downloading the data available at diagnosis from ObTiMA into the CHIC data repository

   Demonstrating the CHIC portal and data repository with its features

2. Selecting separate hypomodels or the complete hypermodel, executing the hypermodel and showing the results of the hypermodel

   Selection of hypomodels depends on which data are available at time of diagnosis, if possible both options should be shown: selection of hypomodels and execution of the resulting hypermodel when all data are available at diagnosis vs. when not all data are available

3. More data become available during the treatment; entering the CHIC portal again and downloading the new available data from ObTiMA into the CHIC data repository

   Selecting additional hypomodels, executing the hypermodel and showing the new results of the hypermodel together with the previously obtained results

4. *The outcome data are available: entering the CHIC portal again and downloading the outcome data from ObTiMA into the CHIC data repository*

   Selecting the validation tool, executing of the validation and automatic comparison between the predicted outcome and the real outcome

Within these clinical demonstrators technical demonstrators can be implemented. One technical demonstrator might deal with the legal tools to demonstrate data safety and privacy. This might be able to do within the demonstration of data upload into the CHIC repository and to show the difference between the CHIC clinical and research data warehouse.

For the glioblastoma hypermodel all data (except the pathology and biochemical data) for the hypomodels are already/soon to be available in the CHIC data repository. The datasets for validation still need to be entered in the data management system, but are available at UZ Leuven.

4.5 Timeline for developing the hypermodel for clinical usage

An iterative process between all stakeholders of CHIC developed the timeline for the next 6 months to build all necessary components for the smooth run of the demonstrators. During this process a prioritization was established which components are most important to be finalized and which

components will lack full functionality at that time. The following table outlines this timeline of tools to be developed. Red indicates the time for development and green gives the time when the tool is ready to be used. It also contains those features that will not be developed within the next 6 months.

Tools	August 2015	September 2015	October 2015	November 2015	December 2015	January 2016	January 2017
Front-end / CRAF						Prototype available	
CHIC data repository							
Connection between ObTiMA and CHIC data repository							
Upload tool and CRAF						Prototype available	
All hypomodels							
Report of results							
Validation of results							
Fine-tuning of the hypermodel after validation							

5 Non-small cell lung cancer hypermodel

5.1 Short outline of the non-small cell lung cancer hypermodel

The deliverable describes the non-small cell lung cancer (NSCLC) scenario as outlined in D2.2. Patients with NSCLC which are at a surgically treatable stage (e.g. stage I up to stage IIIa) are usually treated by lung lobe resection and lymphadenectomy. In most cases clinically and by routine imaging tools (thorax X-ray, CT) the tumour load seems removed by this procedure. Despite these results at least 40% of the NSCLC patients, even in stage I of the disease, show local or metastatic recurrence of the disease within the next years. For these patients it is of utmost importance to know which clinical, pathological and/ or molecular findings indicate the risk of recurrence and which of these patients will benefit from surgical treatment alone. The hypermodel therefore shall answer the question: Will a given surgically treated NSCLC patient show clinical and/ or pathological i.e. molecular parameters indicating the risk of recurrence? The hypermodel will be validated in two data sets of patients with distinctive micromorphological features: 1) Patients with primary pulmonary adenocarcinoma of the acinary adenocarcinoma type, corresponding to a tumour grading of G2 and 2) Patients with a primary pulmonary adenocarcinoma of the solid type, corresponding to a tumour grading of G3 (TNM classification). After validation of the hypermodel the tool can be used prospectively to advice clinical follow-up treatment of primarily surgically treated NSCLC patients. Beyond the hypermodel can be integrated in the follow-up scenario of patients with high stage of the disease (stage IIIb and IV) to check if the parameters detected by the hypermodel will be of clinical and/ or predictive or prognostic relevance in this group of patients with a usually extremely bad prognosis (mean overall survival time less than 12 months after pathological diagnosis of the disease).

From a clinical perspective the following steps describe the NSCLC scenario:

1. Steps outside of the CHIC project:
    a. A specific patient with NSCLC is enrolled in a prospective clinical trial
    b. Clinical, imaging and molecular data are collected at the time of surgical resection
    c. Clinical, imaging and histopathological/ molecular data (if a biopsy procedure is performed) are collected at the time of recurrence

2. Steps within the CHIC project:
    a. Data upload to the CHIC data repository
        i. A tool is needed for upload of the different data types, for guaranteeing data safety and security, to handle semantic interoperability
    b. Data post-processing
        i. Tumour segmentation using DrEye needs to be done and the segmented data needs to be uploaded
    c. Selection of hypomodels or a hypermodel
        ii. The clinician is able to select a hypermodel composed of different hypomodels.
    d. Execution of the hypermodel
        iii. After selection of the hypermodel this will be automatically executed

e. Visualization, reporting and storage of the results

   iv. Results of the hypermodel will be visualized, reported and stored in the CHIC platform linked to the specific patient

f. Validation of the results in single patients and fine-tuning of the hypermodel

   v. At the time of recurrence new data are uploaded to the CHIC data repository to allow the validation and fine tuning of the hypermodel

A detailed description of all these steps is given in the following sections of this chapter.

5.2 The Master Topology of the CHIC Lung Cancer Multimodeller Hypermodel

The CHIC lung cancer multimodeller hypermodel consists of two modelling pylons (Fig. 5.1). On the left hand side a multimodeller hypermodel provides a detailed multiscale simulator of the complex phenomenon of *in vivo* tumour growth and response to treatment ("first pylon"). On the right hand side a set of simplistic hypomodels provide a gross description of the same phenomenon ("second pylon").

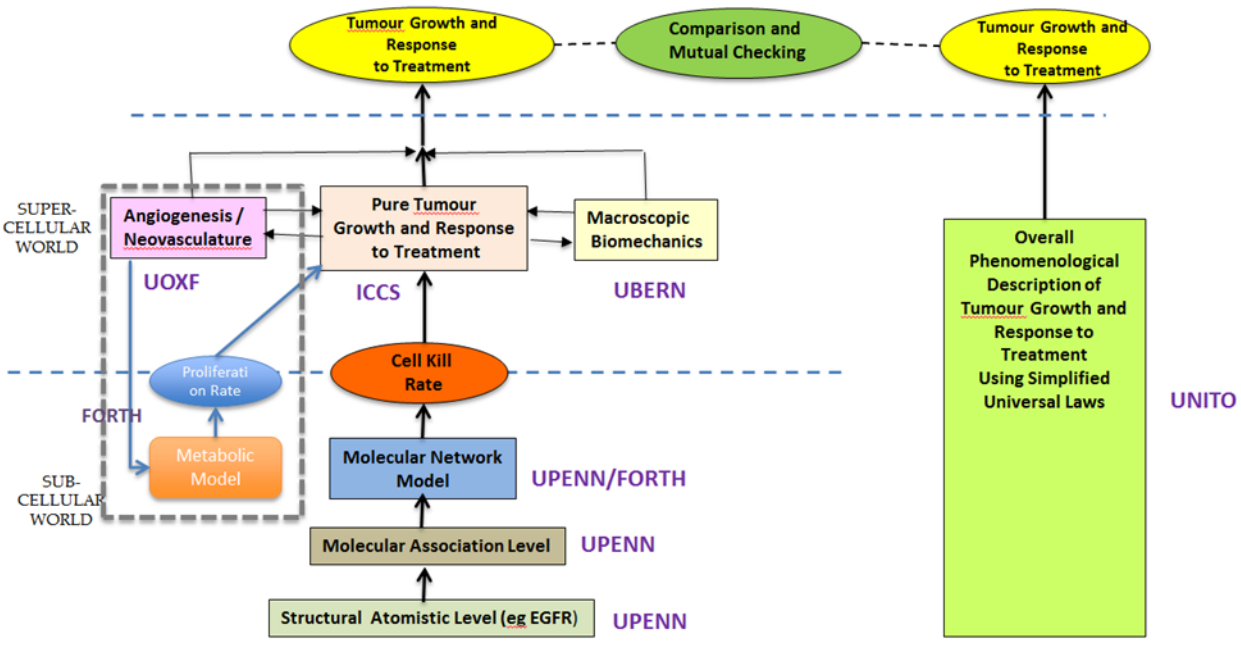


Fig. 5.1: High level topology of the CHIC multimodeller hypermodel of lung cancer. Basic science view.

ICCS has undertaken the development of the pure tumour growth hypomodel as well as the integration of the hypomodels developed by other partners in order to end up with the integrated multiscale hypermodel concerning angiogenetic tumour growth and response to treatment. UOXF has undertaken the development of the tumour related angiogenesis hypomodel. UBERN has undertaken the development of the macroscopic biomechanics hypomodel. UPENN has been

charged with the development of a composite hypomodel dealing with the atomistic and the molecular levels. In particular this hypomodel comprises a structural atomistic hypo-hypomodel, a molecular association hypo-hypomodel and a molecular network hypo-hypomodel. A refinement of the molecular network hypo-hypomodel focusing on metabolism has been undertaken by FORTH. The crucial parameter of cell kill probability that is meaningful for both the subcellular and the cellular levels of biocomplexity has been used as the linking point of the subcellular with the cellular and supercellular world. In fact atomistic and molecular modelling lead to the patient individualization of this central hypermodelling parameter. In parallel with this complex and detailed multiscale hypermodel, UNITO has produced a number of gross hypomodels based on simplified universal growth laws. Their purpose is to provide an initial very gross but still useful check of the overall behaviour of the multimodeller hypermodel. Conversely, the detailed multiscale multimodeller hypermodel is to serve as a possible corrector of the UNITO gross set of hypomodels.

5.3 Steps of interactions of the clinician with the integrated non-small cell lung cancer hypermodel

All clinical relevant steps of the NSCLC scenario are given above. These steps are describing main points of interaction of a clinical end-user with the hypermodel. Points of intersection with basic scientists, IT-people and modellers are given in the descriptions below that will make these hypermodels executable and usable for end-users. All tools developed by basic scientists, IT-people and modellers, as well as the ethical and legal framework at these intersection points will guarantee a smooth run of the hypermodel execution by presenting a closed workflow to the end-user with easy and understandable interactions. The first step of data creation is already within CHIC. It describes the data that are needed for running the hypermodel and that need to be uploaded to the CHIC data repository.

At each step the necessary developments by basic scientists and IT-people are mentioned. In addition the legal and ethical considerations are shown that need to be fulfilled. These necessary developments are important to guarantee a smooth execution of the hypermodel by end-users. They are described from a clinical perspective and written in cursive blue letters as separate sections.

5.3.1 Steps outside the CHIC platform

5.3.1.1 Data creation and collection

A new patient with a NSCLC will be recruited for participation in diagnostic branch of the Institute of Pathology. For the NSCLC hypermodel the following data are needed:

1. Clinical data
2. Imaging data
3. Histopathological data
4. Molecular data
5. External data (e.g. drug database, literature, KEGG database, etc.)

Clinical data will be stored in ObTiMA[5] (Ontology based Trial Management Application) as the clinical trial management system (CTMS). In ObTiMA a first pseudonymization will take place. ObTiMA will also store the molecular data of the tumour tissue at diagnosis. Imaging data are stored as DICOM files in the local PACS (Picture Archiving and Communication System) of the treating hospitals. They

[5] http://obtima.org

are transferred with the first pseudonym to the CHIC study at USAAR and stored on a local DICOM server. If MRI is clinically indicated and performed MRI data of the different modalities are used. All data used for the NSCLC hypermodel are defined by the different hypomodels to guarantee their scientific correct development by basic scientists (e.g. the KEEG database for the pathway hypomodel; etc.).

5.3.2 Steps within the CHIC platform

If a clinician as an end-user will execute a hypermodel he needs primarily to enter the CHIC platform. For that purpose the IT people developed CRAF as a front-end. The clinician needs to be registered in CHIC and receives credentials to enter the CHIC platform. Here he/she can easily select the hypermodel he/she wants to run. After selection he/she is guided through all steps to successfully execute the hypermodel.

Legal and ethical considerations:

The legal and ethical framework needs to be in place including necessary signed contracts as given by the legal WP. Where the validation process requires the clinician to upload and/or process patient data in personal form, the specific informed consent of the patient or his representative will be obtained. In addition the end-user will not interact with any tool guaranteeing data safety and data security. These tools are all integrated in tools processing the data, like the upload tool and the CRAF.

Necessary developments by basic scientists and IT-people:

CRAF or the CHIC portal as a frontend needs to be in place including a roles and rights management where a user can easily select models and execute them. After the selection of the hypermodel the end-user will be guided automatically through all steps of the executable workflow.

5.3.2.1 Data upload

The first step after selection of the hypermodel is to run the data upload tool. For that purpose the data upload tool will ask to select a patient and the different data types to make an upload of data for this patient to the CHIC data repository possible. In the NSCLC the clinical data and the molecular data, i. e. data on so-called driver mutations as well as miRNA data from tumour tissue are stored in ObTiMA. The imaging data are stored on a local DICOM server. All data are already pseudonymized for the first time. ObTiMA also guarantees semantic interoperability of the data.

Concrete steps are:

By selecting clinical and molecular data the user is forwarded to ObTiMA, where the specific patient can be selected. The end-user can select on this page, which of the data he/she wants to download from ObTiMA into the CHIC data repository. After selection of this data he/she starts the download - upload process within ObTiMA by clicking a button on the download page in ObTiMA. If other molecular data besides the driver mutation data are needed they have to be uploaded via the upload tool. The end-user then selects the imaging studies of the specific patient by the upload tool. After selection DrEye will open and the imaging data of this patient are uploaded into DrEye. The end-user needs to render the tumour in all modalities before the imaging data and the tumour segmentation can be uploaded to the CHIC data repository.

Necessary developments by basic scientists and IT-people:

Security integration between CHIC and ObTiMA is needed. Web services from ObTiMA to CRAF are needed for single sign on (SSO) and connection between CHIC and ObTiMA. In ObTiMA the patient

selected in the portal needs to be automatically found and the end-user will be directed to the page for download in ObTiMA. A button for download and upload to the CHIC data repository needs to be available within this page in ObTiMA. The end-user will be signed off after clicking this button, ObTiMA will then automatically close and the end-user is redirected to the upload tool, to select the imaging data of the patient.

5.3.2.2 Post-processing of data

Regarding the NSCLC hypermodel imaging data needs to be post-processed to gain data of the tumour volume, the tumour composition and the surrounding tissue. This post-processing of the data is done within DrEye. The imaging data are automatically opened in DrEye as described in 5.3.2.1. After tumour segmentation is finished the end-user will store the imaging studies, the tumour segmentation data and further statistical data of the tumour composition, done by DrEye, on the local DICOM server. At the same time DrEye will close and the end-user is redirected to the CHIC upload tool, where selecting the post-processed data the upload of them will start.

As soon as a completely automatic segmentation tool for the NSCLC's is available this tool will be integrated. This will change the workflow in a way that DrEye will be integrated after the automatic segmentation process. The end-user will select the imaging studies and the tool for automatic segmentation of the tumour. This tool will then call DrEye after finishing the segmentation process and upload the images with the segmentation via CRAF for reviewing by the clinician. After the end of this review process the workflow continues as described above.

No other data do need post-processing. The interpretation of the driver mutation data and miRNA data will be done in the molecular hypomodel.

Necessary developments by basic scientists and IT-people:

The clinician selects imaging data by using the end user application CRAF, and then CRAF calls DrEye for execution. After the end-user has finished the segmentation process and stored these data on the local DICOM server by using DrEye, DrEye will be closed and the imaging data as well as the segmentation data and further statistical data provided by DrEye will be uploaded to the CHIC data repository by the CRAF and the data upload tool.

5.3.2.3 Selection of hypomodels/hypermodel

After uploading of all data of a specific patient the end-user is guided to select hypomodels or directly a hypermodel. For that purpose different hypomodels are shown to the end-user that are related to the NSCLC scenario. The end-user can select independently each alone or together with others or all selection hypomodels. The clinician is guided to the Hypermodelling Editor where he/she can freely design a new hypermodel, by connecting different hypomodels. Then the data of the specific patient are automatically requested by the hypermodel from the CHIC data repository. As soon as the so composed hypermodel is created the end-user can execute the hypermodel. The validation tool will compare the results given by the hypermodel for NSCLC and the real data for fine-tuning the hypermodel. It is possible to use directly a composed hypermodel for a selected patient. This hypermodel is always composed of all hypomodels as described in the following sections.

Necessary developments by basic scientists and IT-people:

All hypomodels as described above need to be finalised. By selecting different hypomodels their linkage to compose a functioning hypermodel needs to be done via the Hypermodelling Editor. The

data of the specific patient can be retrieved automatically from the CHIC data repository. If this is not possible after a second pseudonymization the CHIC data repository needs to be divided into a clinical data repository and a research data repository. This means that the upload of specific data of a patient by the data upload tool or from ObTiMA should not perform a second pseudonymization on the data. To get these data available in the research data repository a tool needs to be build that will download the data from the CHIC clinical data repository to the CHIC research repository. In addition access rights to the CHIC clinical data repository needs to be restricted to clinicians that are allowed to run the hypermodel in the clinical setting even if the results of the hypermodel is still research.

If validation will be done the workflow has to check if results of the hypermodel are available and post-processed data of imaging studies after pre-operative chemotherapy are uploaded to the CHIC data repository. Otherwise the validation tool cannot be selected. In that case the end-user will be asked to primarily run the hypermodel for this patient and if necessary to post-process the imaging DICOM files after preoperative chemotherapy. This will be done accordingly to the described process for the imaging studies at diagnosis.

Legal and ethical considerations:

If a clinical data repository is needed this implies legal considerations to guarantee data safety and privacy.

5.3.2.3.1 Molecular Network Model

The end-user will not be able to interfere with the molecular network model. He just can select it for usage. Basic scientists create the hypomodel.

Necessary developments by basic scientists and IT-people:

The molecular network model needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

5.3.2.3.2 Imaging hypomodel

The end-user will not be able to interfere with the imaging hypomodel. He just can select it for usage. Basic scientists create the hypomodel. This hypomodel provides the cell kill rate.

Necessary developments by basic scientists and IT-people:

The imaging hypomodel needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

5.3.2.3.3 Metabolic hypomodel

The end-user will not be able to interfere with the metabolic hypomodel. He just can select it for usage. Basic scientists create the hypomodel. There are different possibilities to use the miRNA data within the metabolic hypomodel. The first one is to characterize tumour proliferation or tumour necrosis or apoptosis by the expression levels of specific miRNAs that are related to proliferation, necrosis or apoptosis. For that purpose data from literature are needed that show such correlations between miRNA expression and proliferation or apoptosis and necrosis. The second possibility is to use the expression levels of the miRNAs to find deregulated molecular pathways that interfere with

proliferation, necrosis or apoptosis. For that reason data from KEGG are needed and will be used in this hypomodel. The hypomodel will provide the proliferation rate.

Necessary developments by basic scientists and IT-people:

The metabolic hypomodel needs to be finalised for usage. Necessary external data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources.

5.3.2.3.4 Supercellular hypomodels

The supercellular hypomodels are dealing with angiogenesis, the pure tumor growth and response to treatment and the macroscopic biomechanics. The end-user will not be able to interfere with these hypomodels. He/she just can select them for usage. Basic scientists create the hypomodels. For the usage of the pure tumor growth and response to treatment hypomodel results of the metabolic, the molecular network, the angiogenesis and the biomechanical hypomodels are needed or data from literature have to be used.

Necessary developments by basic scientists and IT-people:

The supercellular hypomodels need to be finalised for usage. Necessary data need to be uploaded to the CHIC data repository by basic scientists or directly used from the external resources. The linkage between the different hypomodels needs to be established.

5.3.2.4 Execution of the hypermodel as an integrated tool

After the selection of the hypermodel by the end-user this hypermodel will be executed without further intervention by the end-user.

Necessary developments by basic scientists and IT-people:

The execution of the hypermodel needs to be fluently within a certain time accepted by clinicians. If the tool will be used in the future for decision support results should be available with a short period of time.

5.3.2.5 Visualization, reporting and storage of the results

The execution of the hypermodel will produce visualization results and a report. The visualization results as well as the report will be automatically stored in the CHIC in-silico-trial repository linked to the specific patient. The report contains information about who did run the hypermodel at what time with which data. The composition of the hypermodel by the different hypomodels is explained as well. All data used for the different hypomodels are displayed in tables related to the hypomodels. If the molecular hypomodel is used the deregulated pathways are shown. If the drug-associated results are available from the histopathological and molecular data the most relevant drugs are listed and linked to the molecular pathways they interact with. The final result of the hypermodel is the answer to the question, if the tumour will recur postoperatively. For that purpose the predictive tumour parameters will be displayed with a function of prediction uncertainty. The end-user can always select between the report and the visualization tool.

Necessary developments by basic scientists and IT-people:

The visualization tool, the reporting document and a tool to automatically store the results in the CHIC repository need to be developed.

5.3.2.6 Validation of the result in a single patient and fine-tuning of the hypermodel

In case the prediction of the hypermodel needs to be validated the end-user can select from the CRAF the validation tool. This will only be possible if a hypermodel for this patient has already been run and post-processed data of the imaging studies after pre-operative chemotherapy are available in the CHIC data repository. If this is not the case the end-user is not able to select the validation hypomodel from the CHIC portal. To run the hypermodel is explained above. To discriminate imaging data from diagnosis from those at the time of recurrence the date the investigation of the imaging in the patient was done is used. If the data are available and also a result of the hypermodel the validation tool will automatically compare the predicted tumour dimensions. If the difference of the predicted tumour volume is more than 10% of the real tumour volume a fine-tuning of the hypermodel needs to be done. In the future basic scientists will enhance the validation by using also machine-learning programs. It can be expected the more results of the hypermodel will be validated in different patients the better the prediction of the hypermodel will get. The validation will be reported to the end-user and also a visual comparison between the predicted and the real tumour volume will be displayed.

Necessary developments by basic scientists and IT-people:

The validation tool need to be developed and ready to use for clinicians. This tool is divided into the first part of just comparing the predicted result of the hypermodel with the real tumour volume at the time of surgery. The results of this comparison need to be reported and visualized. This needs to be developed. The second part of the tool is the fine-tuning that needs to be developed by basic scientists.

5.4 Demonstrators for upcoming reviews

In the following paragraph four possible demonstrators are described. An outline of these demonstrators is given in figure 3.3 above. As the validation tool will only be needed in future demonstrations at the end of the project, no demonstration is foreseen now.

Altogether four clinical demonstrators dealing with the NSCLC hypermodel are possible. They should all be integrated in CRAF:

1. Entering CRAF and downloading clinical, histopathological, molecular and miRNA data from ObTiMA into the CHIC data repository
2. Post-processing imaging data and uploading them to the CHIC data repository
3. Demonstrating the CHIC data repository with all its features
4. Selecting hypomodels or an hypermodel, executing the hypermodel and show the results of the hypermodel

Within these clinical demonstrators technical demonstrators can be implemented. One technical demonstrator might deal with the legal tools to demonstrate data safety and privacy. This might be able to do within the first demonstration of data upload into the CHIC repository and to show the difference between the CHIC clinical and research data warehouse.

The concrete demonstrators need to be defined between the clinicians and al other members of CHIC to get realistic demonstrators that do run smoothly. This selection is based on the availability of the different hypomodels and the possibilities of implementing such a closed workflow as written. For the NSCLC hypermodel most data that are needed are already available in the CHIC data repository.

5.5 Timeline for developing the hypermodel for clinical usage

An iterative process between all stakeholders of CHIC developed the timeline for the next 6 months to build all necessary components for the smooth run of the demonstrators. During this process a prioritization was established which components are most important to be finalized and which components will lack full functionality at that time. The following table outlines this timeline of tools to be developed. Red indicates the time for development and green gives the time when the tool is ready to be used. It also contains those features that will not be developed within the next 6 months.

Tools	August 2015	September 2015	October 2015	November 2015	December 2015	January 2016
Frontend / CRAF						Prototype available
Clinical Data Repository						
Features to be integrated into DrEye						
Report of results						
Data Upload tool and CRAF						Prototype available
All hypomodels						

6 Prostate cancer hypermodel

6.1 Short outline of the (recurrent) prostate cancer hypermodel

Dosing the Prostate Specific Antigen (PSA) in serum can easily monitor prostate cancer, one of the most diffused tumours in males.

Whenever the PSA level overcome a safety threshold and following multifocal biopsy, patients commonly undergo either radical prostatectomy (RP) or radical radiotherapy (RRT).

As a matter of fact, within 5 years, about 1/3 of them suffer for a biochemical recurrence, i.e. PSA values start growing again.

Recurrence probability has been intensively studied. It has been related to the main clinical parameters by nomograms (Partin et al 2001 [2], Briganti et al 2012 [3]) to clinical and pathological characteristics of the primary tumour (e.g. Gleason score, pathological staging, etc.) and of the surgical/ radiotherapeutical procedure (e.g. surgical margins, resected nodes…).

Here the story becomes somehow different for patients who underwent RP and RRT. Since following RP PSA can be only produced by recurrent prostate cancer (either localized or by nodes and/or metastatic sites) this is a clear warning bell. Moreover, for these patients it is conceivable that PSA production is directly correlated with tumor growth, and a careful monitoring of PSA is the basis for a dynamic approach (PSA- hypomodel).

Contrary to surgery, RTT does not abolish the production of PSA by the prostate gland. A very much complex situation occurs. Following RT the PSA reduction may be very slow and the lowest value, called ‘nadir’, may be detected after two years. Moreover, transient increases of PSA, called ‘bounces’, may be monitored especially after brachitherapy.

For these patients it is therefore preferred a ‘nomogram’-approach, which relates the probability of recurrence to the specific pre-treatment values of a large number of clinical parameters (Candiolo nomogram hypomodel).

In both cases **it is therefore mandatory for modelling validation to account for all the above parameters. A large data collection is also needed to represent all the possible sub-cohorts defined by the previous parameters (EUREKA1/2 studies).**

Presently EUREKA1 database contains data from 3538 patients who underwent RP and EUREKA2 database contains data from 3757 patients who underwent radical RT between January 1999 and December 2012. Data collection is still in progress according to the amendment of the ethical committee.

In particular, the PSA-hypomodel can compare the post-RP PSA history of each patient with the corresponding EUREKA1 sub-cohort. The model has been challenged by a MANOVA statistical model accounting for all the available clinical parameters. The statistical relevance of the post surgery PSA values is far larger than the other clinical parameters. This also happens with the values of the mathematical variables estimated by Gompertzian and West PSA growth simulation.

Moreover, only three PSA values evaluated within 18 months showed to be sufficient to predict early (< 24 months) or late (> 48 months) recurrence.

The Candiolo nomogram hypomodel has been validated on the EUREKA2 subcohort.O The post-RT patients showed a very good reliable stratification in 5 class risks of recurrence.

Based on this rationale, patients at high/very high risk of early recurrence, as soon as the PSA level increases after radical surgery/RT, are commonly urged to undergo a adjuvant hormonal therapy (Androgen Deprivation Therapy).

Clinical data collected by Eureka1 and Eureka 2 studies, together with emerging literature, show that:

1) ADT temporarily reduces the PSA level only during the therapy,
2) ADT promotes the development of a hormone-resistant clone, prejudicing any further late therapy.

The use of ADT should therefore be very prudential only when tumour development is very fast as actual adjuvant (second line) therapy. In the meanwhile accurate tumour imaging (US, CT, MRI) has to be produced, to precisely quantify the local and global tumour extension.

The ADT-hypomodel aims at predicting the effectiveness of adjuvant ADT. It has been already implemented for RP patients and predictions are available for patients whose PSA values are compared with the corresponding EUREKA1 sub-cohort (i.e. similar clinical and histological parameters and ADT duration, timing and dosage). Work is in progress for developing the ADT-hypomodel also for RRT patients.

Due to the poor correlation between post RRT PSA and tumor regrowth, a specific study of the pre and post treatment tumor volume is mandatory. A specific data collection of DICOM images has begun at FPO-IRCCS Cancer Center of Candiolo (EUREKA-2). The study is focusing on patients treated with IMRT-IGRT technique, and includes for each patient DVH data, multi-parametric-MRI data and Choline-PET-CT data.

The IMAGE-HYPOMODEL can manage imaging results with DREYE'S (or similar) tools to evaluate tumor extension and proliferation characteristics. It will evaluate tumor proliferation according to the Gompetzian and West growth models..

Finally, also based on the primary therapy previously selected, personalized Radio/chemo-hormone therapies have to be provided to control tumour proliferation (third line therapy).

SALVAGE RT-HYPOMODEL can produce a personalized schedule for RT accounting for estimated tumor proliferation.

SALVAGE CH-HYPOMODEL can produce a personalized schedule for combined therapy accounting for estimated tumor proliferation.

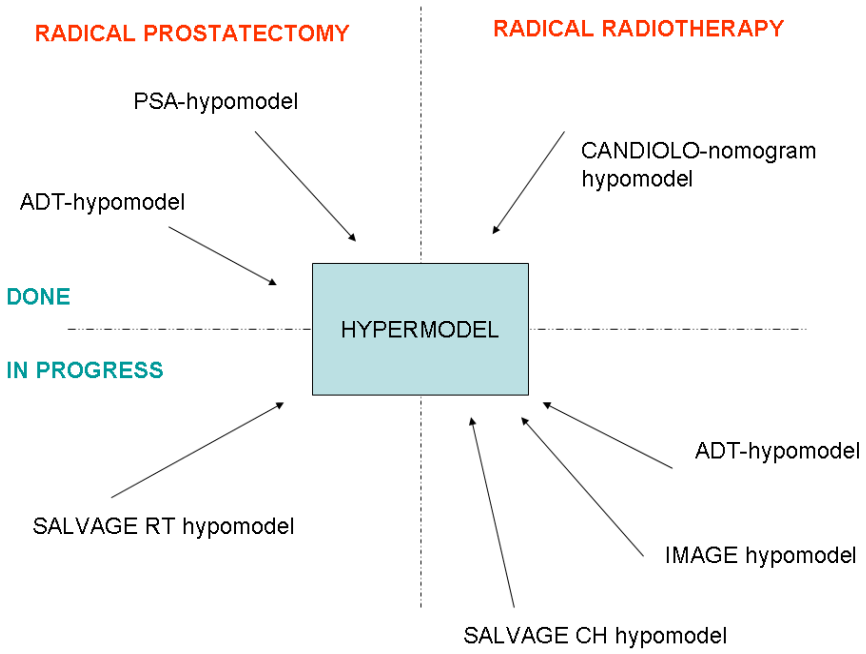


Fig. 6.1: Schematic view of the recurrent prostate cancer scenario described by the linkage of the hypomodels composing the advanced prostate cancer hypermodel

6.2 Steps of interactions of the clinician with the prostate cancer hypermodel

From a clinical perspective the following steps describe the recurrent prostate tumour scenario:

6.2.1 Steps outside the CHIC platform

Clinicians are expected to perform examinations to get clinical and imaging data in their own working context.

a. Pre-operative or radical-RT clinical and imaging data are available

b. Clinical and imaging data can be collected at all times during second and third line therapies

6.2.2 Steps within the CHIC platform

Clinicians can enter the CHIC platform by a dedicated portal, after a proper registration procedure.

At the moment the procedure can be performed only for those who joined the Eureka1/2 studies based on the amendment of the Ethical Committee authorization. In the near future according to the CHIC data provider agreement signed by UNITO Eurekas database will be more widely available.

Via CRAF or the portal he/she can easily select the hypermodel he/she wants to run. After selection he/she is guided through all steps to successfully execute the hypermodel. As soon as the execution is finished the user will be notified and results are given.

6.2.2.1 Enrolment in Eureka studies

A specific patient who underwent radical prostatectomy/RT is enrolled in a prospective clinical trial. To strictly fit our national rules, the patient should be already listed in the Eureka studies, which have been approved by the ethical committee also for prospective follow up of enrolled patients.

6.2.2.2 Data upload

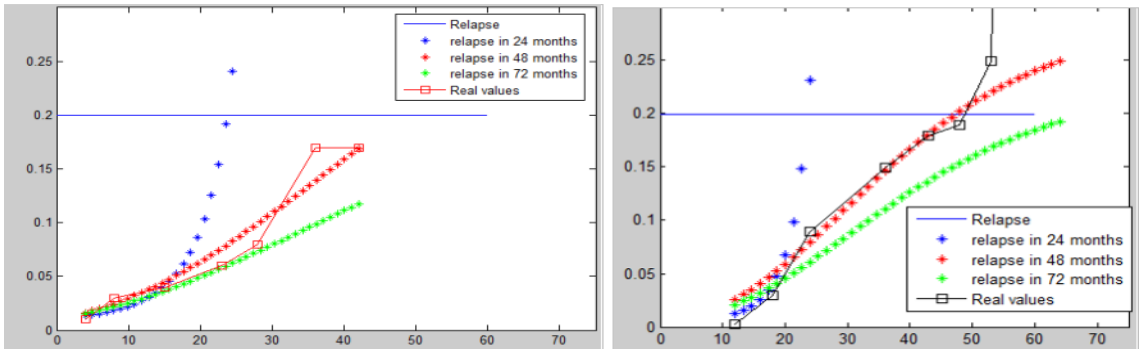
Actually a csv file is needed. It will be possible to create a user-friendly interface to guide the clinician to create the csv file using his/her information.

- Clinicians presently involved in the research (Eureka1/2 team) have a direct access to patients database. In perspective interested clinicians will define with CHIC a protected procedure to consult data from the CHIC database according to the CHIC data provider agreement signed by UNITO
- Data from Eureka1 and Eureka2 database are selected corresponding to the same clinical parameters pertaining to the patient (patient sub-cohort)
- Clinical data of the patient are recorded. A tool is needed for uploading different data types, for guaranteeing data safety and security, to handle semantic interoperability

6.2.2.3 Post-processing of data: according to the requirements of the hypermodel tools

- For RP patients: the first 3 or 4 PSA values will be used to forecast the relapse at 24, 48 and 72 months using Gompertz and/or West models (PSA-HYPOMODEL); the next PSA values will be compared with the three curves, giving an idea of the tumor velocity and the time to relapse (see Figure 3.3).

Figure 6.2: simulations of the behavior of the tumor in case of relapse at 24, 48 and 72 months using only the first 4 PSA values. For the formulas see [4], [5] and [6].



- For RRT patients: inserting the clinical parameter values required by the CANDIOLO NOMOGRAM hypomodel the patient may be inserted in the proper class of risk for recurrence. When pre and post treatment images are available, the IMAGE hypomodel can assess the tumor volume
- In case of adjuvant ADT therapy, using the West growth law, an estimation of the strength of proliferation of the tumor is done (ADT-HYPOMODEL), eventually comparing the current patient with the similar patients into the DB
- SALVAGE treatments (RT and CT) are simulated using known algorithms.

6.2.2.4 Selection of hypomodels/hypermodel

The clinician could chose if he/she wants to run only the PSA-HYPOMODEL, the CANDIOLO NOMOGRAM HYPOMODEL, the IMAGE-HYPOMODEL, the ADT-HYPOMODELS or all together. The SALVAGE RT and CT-HYPOMODEL may require the others hypomodels as pre-requisite.

FOR RP PATIENTS: PSA-hypomodel: Recurrence timing prediction

a. At least 3 PSA successive values are collected after suspecting recurrence
b. The proliferative tumour strength 'a' is computed
c. According to the a values estimated in the patient's sub-cohort, a early or late recurrence of tumour is predicted

- In case of slow progression, other PSA values are collected at least every 3 months and the procedure related to PSA-hypomodel is repeated until a quick progression is foreseen.

FOR RRT PATIENTS: CANDIOLO NOMOGRAM HYPOMODEL is used for class risk classification

- In case of quick progression or high/very high recurrence risk:
  a. PSA data are collected every month

b. The ADT-hypomodel is run using the proper patient' sub-cohort data

c. ADT is eventually prescribed with proper dose and timing

d. Accurate tumour imaging is performed, tumour segmentation using DrEye is done and the segmented data uploaded

e. Tumour progression is estimated based on imaging data (IMAGE- hypomodel)

- Selection of third line therapy based on 'SALVAGE therapy hypomodels'

a. Provided the tumour is at least partly localized in the pelvis and the patient underwent surgery as first line therapy a personalized, optimized RT procedure is tailored by SALVAGE-RT -Hypomodel

b. Provided the tumour is spread outside the pelvis a personalized, optimized, multi-drug (docetaxel + ADT) is tailored by SALVAGE-CH-Hypomodel.

c. In both cases the eventuality of a hormone-resistant clone has developed following ADT is considered, and therapy parameters will account for it.

6.2.2.5 Execution of the hypermodel by means of a prototype of prostate-oncosimulator (URO-ANGEL?)

After connection of hypomodels or after selection of a hypermodel, the hypermodel will be automatically executed for the selected patients.

Results could be inserted in a new part of the Eureka database for further validation.

The Hypomodels can run in sequence.

6.2.2.6 Visualization, reporting and storage of the results

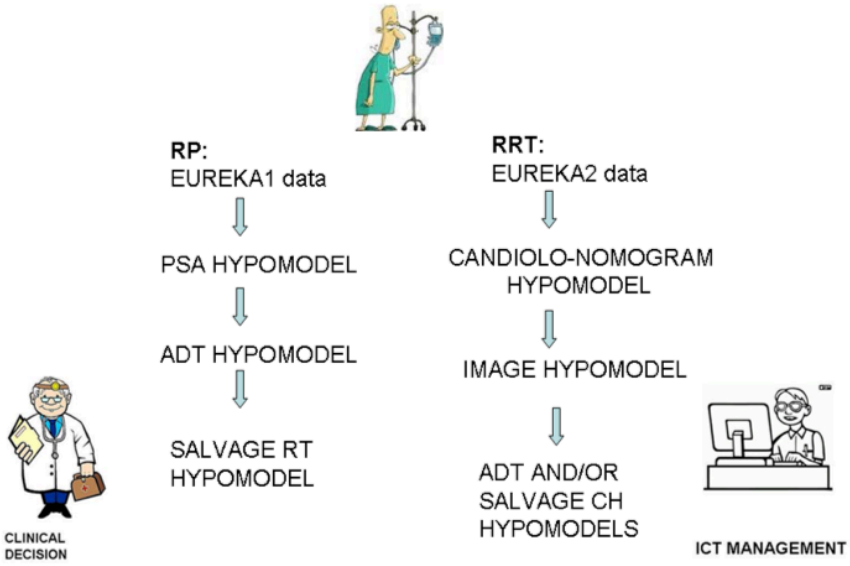
The outputs will be some jpg/png/pdf files with the plots of the simulations and Excel files with the estimated parameters; a metric to chose the more probable time period of the future relapse could be implemented.

Results of the hypermodel will be visualized, reported and stored in the CHIC platform linked to the specific patient.

6.2.2.7 Validation of the result in a single patient and fine-tuning of the hypermodel

We have more than 100 patients with a relapse, more than 5 PSA values and no adjuvant therapy before the relapse. We will be able to test and fine-tuning the model using them. After the end of the third line therapy, new data are uploaded to the CHIC data repository to allow the validation and fine-tuning of the hypermodel.

A detailed description of all these steps is given in the following sections of this chapter.



6.3 Demonstrators for upcoming reviews

All the hypomodels needed for running the (recurrent) prostate cancer hypermodel are nearly almost ready for a first demonstration. They should all be integrated in the CHIC frontend (portal):

1. Entering the CHIC portal and downloading clinical data from a standard patient into the CHIC data repository
2. Evaluating the probability of early – late tumour recurrence (PSA-HYPO ready as MATLAB file) for RP patient or the class of risk for recurrence (CANDIOLO NOMOGRAM –HYPO) for RRT patient
3. Evaluating the effectiveness of adjuvant ADT (ADT-HYPO almost ready as MATLAB file)
4. Evaluating tumour volume from imaging features for RRT patient: we should be more acquainted with Dr Eye's
5. Tailoring salvage therapies: tools are not ready at the moment, but the same routines already shared for lung cancer can be implemented in short time

6.4 Timeline for developing the hypermodel for clinical usage

An iterative process between all stakeholders of CHIC developed the timeline for the next 6 months and beyond to build all necessary components for the smooth run of the demonstrators. During this process a prioritization was established which components are most important to be finalized and which components will lack full functionality at that time. The following table outlines this timeline of tools to be developed. Red indicates the time for development and green gives the time when the tool is ready to be used. It also contains those features that will not be developed within the next 6 months.

Tools	August 2015	September 2015	October 2015	November 2015	December 2015	January 2016	January 2017
Frontend / CRAF (in cooperation with ICT people)						Prototype avaliable	
PSA-Hypomodel							
CANDIOLO NOMOGRAM-Hypomodel							
ADT-Hypomodel							
IMAGE-Hypomodel (in cooperation with ICT people)							
SALVAGE THERAPIES-Hypomodel							

6.5 Common and re-usable developments

6.5.1 Re-usable developments for hypermodels for other cancer types and beyond cancer

The basic idea of the hypermodel could be applied on all the other tumor types.

In particular, the hypomodels IMAGE-HYPO and SALVAGE RT AND ADT-CT HYPO can be adapted to the other three types of tumors within the CHIC project. Although the parameters about the cellular metabolism will change, the skeleton of the hypermodel will remain the same.

6.5.1.1 Beyond cancer

6.5.1.1.1 Vaccination strategies to prevent Prostate Cancer

A very promising therapy is based on the activation of the Dendritic Cells (DC), which are the most potent antigen-presenting cells and are responsible for the induction of specific antitumor immune response. [7]. Since UNITO has a specific expertise and international patents on nanotechnological carriers for the delivery of gases, drugs, molecular and genetic materials [8] an experimental

investigation comparing adjuvant ADT with DC vaccination on mice is foreseen in order to validate the mathematical model.

6.5.1.1.2 Personalized medicine approach: rehabilitation of surgery collateral effects

Following Radical Prostatectomy a temporary impairment of the urogenital functions is normally suffered. Incontinence problems normally disappear within a few months, but it may continue to occur. It worsens the patient's quality of life, unless specific therapeutic strategy is performed.

Rehabilitation of the pelvic pavement is one of the best approaches. The basic idea is to integrate two different approaches: a user-friendly visual interface and a more rigorous control by the clinician using quantitative parameters.

About the first approach, we want to create a set of 3D animations using free tools like Unity3D and Blender, then to use them as 'serious games' usable by the patients. These SW allow planning of personalized ' at home' training sessions, with or without the active surveillance of the clinician.

About the second approach, the improvements of the patient should be checked at each step with technological tools, like motion capture or ultrasound analysis.

hypomodels	**Radical prostectomized Patients (RP)**	**Radical Radiotreated (RRT) patients**
MOLECULAR LEVEL (circulating bio marker PSA)	**X** PSA-HYPO	
CELLULAR LEVEL	**X** ADT-HYPO	**X** ADT-HYPO IMAGE-HYPO
ORGAN-TISSUE LEVEL	**X** SALVAGE RT-HYPO	**X** ('Candiolo' nomogram-HYPO) SALVAGE ADT- CT HYPO
WHOLE BODY LEVEL- PERSONALIZED	**X** rehabilitation	

7 Conclusions

The clinical relevance of the different hypermodels for nephroblastoma, glioblastoma non-small cell lung cancer, and prostate cancer is given in this deliverable. The upcoming demonstrators are explained in detail and demonstrate the benefit of the hypermodels for clinical decision support. Legal and ethical issues are addressed to guarantee data safety and security. All necessary implementations from the IT perspective are described and a timeline for all the hypermodels are given.

This document also serves as a test of principal for hypermodels in other cancer types. This will guarantee the reusability of the developments in CHIC for other cancer specific hypermodels and will be a major output of the project. If the evaluation of the described hypermodels by enduser shows that they can be used in clinical care the CHIC platform will need to sustain for further developments. By the usage of machine learning methods the accurate prediction of the hypermodels will improve over time. One can compare the output of the hypermodels of today with the weather forecast of 30 years ago. Therefore the expectation from clinicians about immediately usage of hypermodels in clinical care needs to take this comparison into consideration. In addition to use the result of a hypermodel for clinical decision support they need to be validated and it needs to be shown that decision base prediction of a hypermodel is better than a decision based on the clinical standards of today. For that purpose clinical trials are needed to show this benefit. These clinical trials have to start as soon as the hypermodels are validated.

8 References

[1] http://obtima.org

[2] Partin AW, Mangold LA, Lamm DM, Walsh PC, Epstein JI, Pearson JD., Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. Urology. 2001 Dec;58(6):843-8.

[3] Briganti A, Capitanio U, Abdollah F, Gallina A, Suardi N, Bianchi M, Tutolo M, Salonia A, Freschi M, Rigatti P, Montorsi F. , Assessing the risk of lymph node invasion in patients with intermediate risk prostate cancer treated with extended pelvic lymph node dissection. A novel prediction tool. Prostate. 2012 Apr;72(5):499-506.

[4] Nahum, The radiobiology of hypofractionation, Cli.Onc. 27(2015) 260-269

[5] Castorina, Deisboeck, Gabriele, Guiot, Growth Laws in Cancer: Implications for Radiotherapy, Rad. Res. 168 (2007), 349-356

[6] Jones, Secomb, Dewhirst, El-Kareh: The additive damage model: a mathematical model for cellular responses to drug combinations, J Theor Biol. (2014);357:10-20

[7] Youlin K et al, 4-1BB protects dendritic cells from prostate cancer-induced apoptosis. Pathol Oncol. Res. (2013) 19:177-181

[8] Cavalli R, Guiot C, Prato M, Troia A: A Nanostructure for the vehiculation of gas and/or active ingedients and/or contrast agents and uses thereof. Patent: Int publ n. WO 2015/028901 A1, 05.03.2015

Appendix 1 – Abbreviations and acronyms

ADT	Androgen Deprivation Therapy
CRAF	Clinically Relevant Application Framework
CT	Computer tomography
CTMS	Clinical Trial Management System
DICOM	Digital Imaging and Communications in Medicine
DWI	Diffusion Weighted Imaging
KEEG	Kyoto Encyclopaedia of Genes and Genomes
MRI	Magnetic Resonance Imaging
ObTiMA	Ontology based Trial Management Application
PACS	Picture Archiving and Communication System
PSA	Prostate specific Antigen
RP	Radical prostatectomy
RRT	Radical radiotherapy
RT	Radiotherapy
US	Ultrasound
VPH-HF	Virtual Physiological Human Hypermodel Framework