Clinically Oriented Translational Cancer Multilevel Modelling

Deliverable D2.3

Requirements and specification for the *ContraCancrum* integrated platform and clinical validation studies

Project Number:	FP7--IST-*223979*
Deliverable id:	D2.3
Deliverable Name:	Requirements and specification for the *ContraCancrum* integrated technological platform
Submission Date:	08/06/2009

Information Society Technologies

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	CONTRACANCRUM
Project Full Name:	Clinically Oriented Translational Cancer Multilevel Modelling
Document id:	D2.3
Document name:	Requirements and specification for the *ContraCancrum* integrated platform and clinical validation studies
Document type	INT
Version:	1
Submission date:	08/06/2009
Editor: Organisation: Email:	Gordon Clapworthy BED gordon.clapworthy@beds.ac.uk

Document type PU = public, INT = internal, RE = restricted

ABSTRACT:

Moving from the characterisation of the intended users produced in Deliverable D2.1, this deliverable identifies the anticipated use of the system from the perspective of the individual workpackages and uses this to identify requirements for the final ContraCancrum system.

KEYWORD LIST: User, specification, requirement

MODIFICATION CONTROL			
Version	Date	Status	Author
1.0	15/5/2009	Draft 1	All WP leaders
2.0	21/5/2009	Draft 2	All partners
3.0	31/5/2009	Penultimate	All partners
4.0	4/6/2009	Final	All partners

List of Contributors

- all partners contributed widely in the creation of this deliverable, initally via the workpackage leaders

Contents

Publishable Executive Summary

The ContraCancrum, i.e. the Clinically Oriented Translational Cancer Multilevel Modelling, project aims at developing a composite multilevel platform for simulating malignant tumour development and tumour and normal tissue response to therapeutic modalities and treatment schedules.

The project aims at having an impact primarily in (a) a better understanding of the natural phenomenon of cancer at different levels of biocomplexity and, most importantly, (b) a disease treatment optimisation procedure in the patient's individualised context by simulating the response to various therapeutic regimens. The predictions of the simulators to be developed will rely on the imaging, histopathological, molecular and clinical data of the patient. Fundamental biological mechanisms involved in tumour development and tumour and normal tissue treatment response such as metabolism, cell cycle, tissue mechanics, cell survival following treatment etc. will be modelled. Stem cells will be addressed in the context of both tumour and normal tissue behaviour. From a mathematical point of view, the simulators will exploit several discrete and continuous mathematics methods such as cellular automata, the generic Monte Carlo technique, finite elements, differential equations, novel dedicated algorithms etc. A study of the analogies of tumour growth with embryological development is expected to provide insights into both mechanisms.

ContraCancrum will deploy two important clinical studies for validating the models, one on lung cancer and one on gliomas. The crucial validation work will be based on comparing the multi-level therapy simulation predictions with the actual medical data (including medical images), acquired before and after therapy.

ContraCancrum aims to pave the way for translating clinically validated multilevel cancer models into clinical practice.

This deliverable builds on Deliverable D2.1, which identified and characterised the anticipated user population of the ContraCancrum integrated technological platform and the uses to which it will be put to produce a broad specification for the system. It views the system from the perspective of the individual workpackages and identifies needs that they anticipate having in the course of the project.

This deliverable should be read alongside Deliverable D3.1, which deals with the technological aspects of the system that will be created to address the users' requirements.

1. Introduction

This document represents work performed within WP2. Extending the outcomes of D2.1, which concentrated on the various user types who will employ the final ContraCancrum system, the main body provides a summary of the requirements identified by the individual ContraCancrum WPs. More detailed technological aspects are covered in Deliverable D3.1, *Report on Architecture Design of ContraCancrum Environment.*

The remainder of this report is as follows. Section 2 defines the objectives of this deliverable and Section 3 introduces the ContraCancrum workpackages. Section 4 is the major part of the report and this defines the requirements at a workpackage level. The outcomes are summarised in Section 5.

2. Objectives of the Deliverable

This deliverable identifies the individual needs of the ContraCancrum workpackages. When allied with the data map to be explored in WP3, it will form the base upon the ContraCancrum integrated platform and clinical validation studies will be built.

A very important innovation of the ContraCancrum project is to provide an integrated multi-level, scenario-based approach for cancer modelling, and this is also introduced in the deliverable.

3. The ContraCancrum Workpackages

The ContraCancrum WPs are as follows.

WP	Work package title	WP Leader
1	Project management	FORTH
2	User needs analysis and specifications	BED
3	Technological environment	UCL
4	Simulation at the cellular and higher levels of biocomplexity	ICCS
5	Simulation at the molecular level of biocomplexity	UCL
6	Simulating cancer biomechanics	UBERN
7	Tumour imaging and visualisation	PFL-H
8	Integration of the simulation system	ICCS
9	Clinical guidance, validation and translation	USAAR
10	Dissemination and Exploitation	PHL-H

Of these, WPs 1 & 10 relate to project management and project outcomes, and WPs 2 & 3 relate to the design/development of the actual system to be deployed, so are not relevant to the content of this deliverable.

WP8 relates to the integration of the ContraCancrum system, and an indication of how ContraCancrum components can be combined is provided in Section 4.

WPs 4, 5, 6, 7 & 9 relate to users and the way in which the system is to be employed, and these form the focus of the discussions in Section 5.

4. System Integration

The concept of the integrated multi-level, scenario-based approach for cancer modelling has been thoroughly discussed in the context of WP9 and the first scenario–based integration design of the modules of the project is illustrated in Fig. 1, which relates to the case of lung cancer.

This scenario-based integration involves the following modules:

- biochemical-level simulations *in silico* to define the best EGFR inhibitor drug for the individual lung cancer patient in WP5
- molecular-level simulations to define statistical models (e.g. cell kill probabilities), also in WP5
- biomechanical simulations in WP6
- medical image analysis modules in WP7
- discrete and continuous tissue-level models of cancer growth and response to therapy that gradually integrate all the above, in WP 4 and WP8

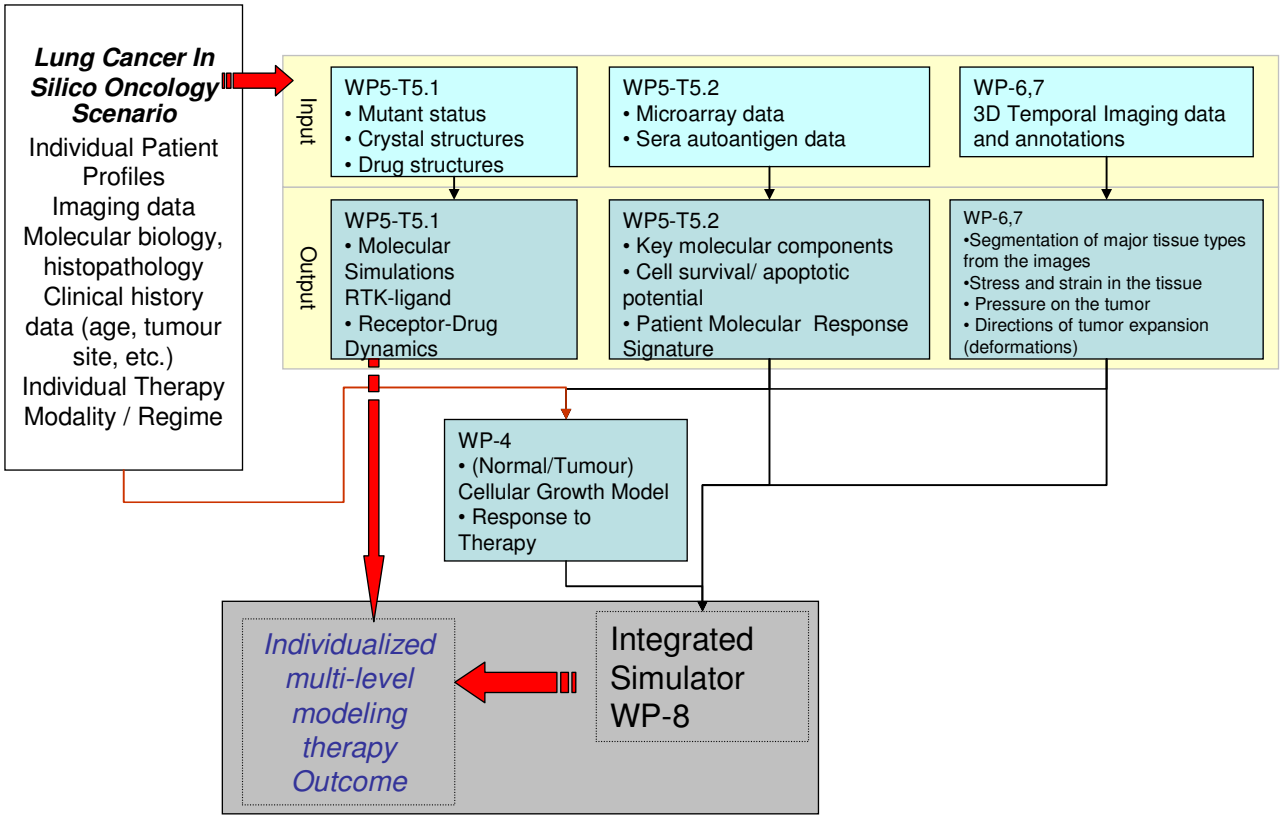

Figure 1. A scenario-based integration strategy of ContraCancrum components in order to provide individualised, modelling-assisted decision making for therapy.

5. Workpackage Requirements

The requirements associated with each of the separate workpackages are highly influenced by the types of user involved in it and the work that will be undertaken. We shall, therefore, consider the needs of each WP individually, as much of what is described is specific to the WP involved.

5.1 WP4: Simulation at the cellular and higher levels of biocomplexity

Objective

The objective of WP4 is to develop a set of multilevel simulation models of tumour growth as well as tumour- (and to a lesser extent, treatment-) affected normal tissue, and of the response to radiotherapy and/or chemotherapy for the cases of gliomas and lung cancer in the patient's individualised context. For the models to be translatable into clinical practice, thorough clinical adaptation, optimisation and validation will be performed in tight interaction with WP9.

General outline

The cell and higher biocomplexity level simulators of the project will be primarily based on the "top-down" simulation approach described above which will be applied for the first time to the simulation of treatment response of lung cancer. This approach will also be refined and adapted for the case of gliomas. Special emphasis will be put on stem and limited mitotic potential tumour cells. In this way, multilevel information will be extracted and used at a realistic level of detail by always taking into account that the whole endeavour is to be confined within the clinical setting.

In addition, a "bottom-up" approach will be adopted in simulating non-imageable tumour growth, especially the process of glioblastoma local invasion. It should be noted that, in parallel with the tumour-response models, rather simple experimentally and clinically based toxicology models (potentially in combination with discrete simulation of the replenishment of normal tissue stem, transit and differentiated cells) will provide safety limits beyond which any candidate treatment scheme would be clinically unacceptable, regardless of the predicted outcome of its tumour control. The multilevel clinical testing, optimisation and validation processes to be performed mainly within the framework of workpackages WP8 and WP9 will constitute *per se* another novelty of the project in the context of gliomas and lung cancer. A detailed technical description of the cellular and higher biocomplexity level models including information flows is provided in Deliverable D4.1.

Interaction with selected other workpackages

Interaction with WP5

The available molecular data will be exploited, these include the status, amplification and expressions of critical genes, which have been shown to drastically affect the response of the tumour under consideration to the treatment addressed. Estimates (even semi-qualitative) of their effect on the cell kill ratio per cell due to the treatment considered will also be provided, based on pertinent literature. The idea is to use the cell-kill-ratio value provided by the Food and Drug Administration (FDA) or general literature as a cell-kill-ratio reference value (e.g. for a given area under curve or radiation dose) and then to perturb it based on the effect of the molecular and/or histological profile of the tumour in order to achieve higher patient individualisation of the simulation.

Interaction with WP6

Using imaging data, a three-dimensional biomechanical model of the brain and lung will be built. Both models will aim at assessing the changes in mechanical environment around the tumour during growth as well as the deformation of the surrounding tissues. First, the finite element method will be used to calculate only the mechanical environment around the tumour while diffusion of the cancerous cells in healthy tissue will be simulated using the continuous and the discrete approach described in Deliverable D4.1. Second, both cell

diffusion and mechanical environment will be calculated using the finite element method. This approach will link the cellular approach described to the macroscopic level.

Interaction with WP9

The integrated simulator outcome (WP8 including WP4) will be compared with the actual (primarily) imaging data provided by WP9 following treatment and an optimisation loop of the parameters estimation (and if necessary of the model algorithm itself) will be initiated.

Simulation scenario

The user of the simulator (either a researcher or a clinician) submits a simulation job through a dedicated user interface developed within the framework of WP3 in collaboration with the rest of the workpackages. The necessary multiscale input data (processed imaging data, processed molecular data, processed histological data, treatment and clinical data) are retrieved from the ContraCancrum database. The spatiotemporal discretisation characteristics of the simulation are defined (size of the geometrical cell of the mesh superimposed on the region of interest, time step). These discretisation data will be used by both the biological and the biomechanical modules. The simulation job is executed on either a grid or a cluster platform. The simulation output, in the form of text and raw files, is stored in the ContraCancrum database. Subsequently, selected output data are retrieved and visualised (e.g. standard graphs, 2D sections, 3D renderings, 4D animations) and compared with the corresponding actual clinical data. Standard statistical processing functionalities will be included in the system in order to facilitate clinical adaptation, optimisation and validation of the models.

Concerning the user interface, although full functionality will be available to both researchers and clinicians, the features directly relevant to clinical practice will be clearly denoted so as to facilitate and accelerate the handling of the system by clinicians without necessitating any extensive technical and basic scientific knowledge.

5.2 WP5: Simulation at the molecular level of biocomplexity

Objectives

(a) Patient-specific chemotherapy drug targeting

Modelling of the interactions of drugs with epidermal growth factor receptor (EGFR) and a number of other pertinent molecular entities in the chemotherapeutic treatment of gliomas and lung cancer: drug-receptor binding studies in glioma chemotherapy and in lung cancer.

(b) Molecular interdependence networks for cell survival probabilities

Design of a statistical model for cancer-type specific prediction of cell survival probabilities in individual patients based on their gene expression profiles. The model will be based on cell-line survival data in response to radio- and chemo-therapy. Validation of the model will carried out in clinical glioma and lung cancer data. Modelled cell survival probability will be integrated into cellular tier of the Oncosimulator through the perturbation of a cell kill parameter.

General Description

The molecular level simulator will study wild-type and mutant EGFRs with a family of their inhibitors through molecular dynamics techniques. A highly automated molecular simulation/free energy calculation workflow tool, the binding affinity calculator (BAC), will be used to study the inhibitor-EGFR interactions. Binding affinities will be calculated, which will help us to select subgroups of patients who are most likely to respond to interventions using

specific drugs. An accurate ranking of drug binding affinities should be achieved at clinically relevant time scales.

For ContraCancrum, WP5 will make use of a middleware tool, the Application Hosting Environment (AHE), to utilise grid resources in a quick, transparent manner. We have already integrated AHE with BAC, so that simulations and analyses can be launched on a wide variety of remote grid resources.

WP5 will need to run several simulations of inhibitor-EGFR systems concurrently on a computational grid and to maintain a suitable quality of service, high bandwidth networks to move terabytes of data around DEISA and other resources should be made available.

The final system should demonstrate the ability to:

- address the required (clinic) data for evaluating and optimising the molecular level simulator;
- gain HPC access, including DEISA;
- launch simulations/analyses on remote supercomputers;
- transfer data among a variety of resources including supercomputers, local machines and storage systems;
- make advance reservations and co-reservations on HPC machines.

The "*Glioma model for prediction of cell survival probabilities in response to radio- and chemotherapy"* will take a pharmacogenomic approach, thereby enabling the simulation of tumour growth based on individualised patient molecular profiles. The model will require data on molecular and radiological cell survival responses, and the corresponding basal gene expression profiles data. By combining the information on therapy response with gene expression, it is then possible to evaluate the impact of different molecular profiles of sensitivity and resistance to be considered *in-silico*.

At present this will utilise expression data, derived from microarray experiments, on cell lines and clinical patient tumours. These data are currently drawn from published datasets (NCI-60, REMBRANDT and TCGA databases); additional data is to be made available for validation through prospective studies generated by the University Hospital of Saarland as part of the clinical validation study (See WP9).

5.3 WP6: Simulating cancer biomechanics

Objectives

WP6 will develop macroscopic finite element models of the brain and lung to calculate the stress/strain in these tissues. Deformations of both the tumour and neighbouring normal tissues will be simulated, based on the tissue biomechanical properties and detailed anatomic atlases. Optimisation and validation will be performed in close interaction with WP8 and WP9.

General Description

The objective of the macroscopic simulator is to consider the biomechanical environment of the tumour in the model as well as the impact of tumour expansion on the surrounding tissues. Cellular information about tumour evolution is obtained from the molecular and cellular levels simulator (WP4).

One should keep in mind that the Oncosimulator is primarily a clinical tool and therefore it should be developed and presented as such. For the clinician, calculations at the micro and macroscopic level in the Oncosimulator should be a "blackbox" where inputs are images of

the patient and output are predictions of the treatment outcome. For this reason, simulations should be performed as fast as possible and results should be made clear for the clinician. This will allow end-user to easily test different clinical scenarios and possible treatments. Moreover, new data and new treatments can then be added and processed.

For researchers, this tool provides a unique platform for the study of tumour but also complex surrounding tissues such as lung or brain in the framework of biomechanics. Therefore, it should facilitate exchange and development of new ideas.

Workflow of the simulation

When a simulation is started for a specific patient, the biomechanical component will connect to the image database and retrieve patient CT/MRI data and the segmentation of the tissues performed by the image analysis work-package (WP7). The most important features of the tissue are delineated in the segmentation – tumour, skull, white matter, grey matter, ventricles, etc.

At best, for clinical use, segmentation should be performed automatically. Moreover, at this step, availability of all the required input data needs to be checked to ensure correct initialisation before the simulation commences.

From the input images, a finite element mesh will be automatically generated for the specific patient. From this, a continuum finite element model will be generated to simulate the tumour, its growth and the mechanical perturbations induced on the surrounding healthy tissues. The mesh and boundary conditions constitute the input model for the biomechanical simulation and need to be stored for future verification. Calculation will be performed by dedicated software on remote supercomputers to ensure fast computation. The model should account for sources and sinks of matter linked to the tumour growth. Input on local tumour growth and change in volume are obtained by coupling the biomechanical model to the cellular simulator.

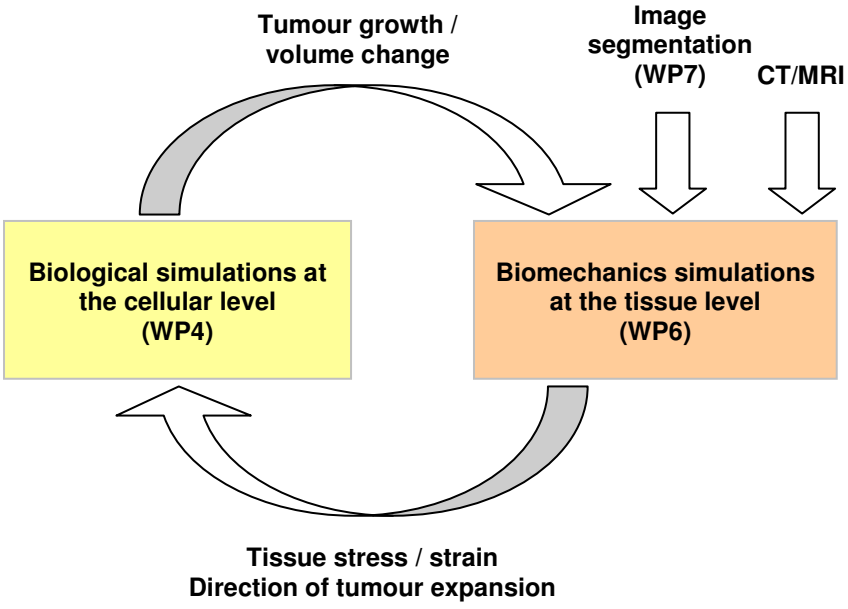

Figure 2. WP6 workflow and links with other WPs

One possible scenario for this coupling is to consider that each geometrical cell of the initial cubic mesh has unlimited biological cell capacity. This implies that, although the initial cellular mesh will be used throughout the simulation, its geometrical elements are able to

deform, expand or shrink so that their actual geometry may change throughout. This implies that very large deformations of the initial geometrical elements can take place according to the FE analysis.

In this scenario, at each scan the biological module calculates the new numbers of cells contained within each cellular mesh element. The new number of biological cells is provided to the biomechanical module which proceeds to the deformation of the corresponding element. Information regarding the update of the metabolic field is sent back to the biological module.

At each step, the result of both the cellular and biomechanical simulation are stored and made accessible to the researcher. Considering the huge amount of data generated by the simulation, this could be a challenging task and care should be taken to facilitate the access as much as possible. Moreover, any modification of the parameters of the simulator and the addition of new features (changes in the biomechanical model used, for example) should also be possible. This is extremely important since the combination of the cellular and biomechanics simulator can work at different levels and be realised with different scenarios.

Postprocessing

In terms of postprocessing, again, requirements are different for clinicians and researchers. Clinicians need only be able to access the final result of the simulation (combining macro and microscopic simulations) and obtain clinically relevant results. The preprocessing (handling of patient data, image processing, choice of treatment) and postprocessing should be gathered in the same clinical application.

For researchers, a web-based finite element viewer to study the different steps of the simulation could be an interesting possibility. Otherwise, easy access to the simulation data base for local postprocessing is mandatory.

Specific Requirements

The final system should provide:

- software-based access to the project DICOM databases to enable users to retrieve relevant data concerning the patient such as CT/MRI/DTI images and corresponding anatomical segmentation files and models.
- remote access and visualisation of the simulation results as well as simple download of these data to enable local analysis (or optionally, web-based visualisation).

The facilities provided should enable users to:

- launch automatic mesh generation software
- launch finite element simulations on remote computers automatically
- transfer results of analyses to the patient's database
- establish sequential coupling between simulation processes performed at the cellular and biomechanical levels, in particular by allowing the exchange of simulation parameters and messages to trigger the multi-level simulations.

5.4 WP7: Tumour imaging and visualisation

Objectives

In terms of medical image analysis tools, WP7 will develop the necessary image registration and data fusion components for *in silico* modelling of tumours. Imaging data before, during and after treatment should be adequately fused (if several imaging modalities are to be used

concurrently e.g. CT and MRI), segmented, reconstructed in 3D and registered so that the treatment outcome can be reliably visualised and quantified. A set of software modules will be developed to implement these processes.

WP7 will enable the visualisation of the imaging data as an input to the simulation model and the visualisation of the simulation predictions. It will also develop tools to compare the actual treatment result to the simulation predictions.

Whenever possible, WP7 will use image processing and visualisation methods that are already available either from open source or from Philips' software platform.

General description

WP7 provides image processing and visualisation related services to other ContraCancrum partners. In the ContraCancrum data processing chain, WP7 partners and their tools take clinical image data provided by the clinical partners as input. These images are processed and the results are stored in the ContraCancrum database, to be verified by the clinical partners, and to be used by partners in WP4 and WP6.

WP7 also provides visualisation tools for partners in WP4 and the clinical partners to interact with the image processing results.

From discussions within WP7 and with partners from WP4 and WP6, it was decided that the original clinical image data should be stored on a standard DICOM server and should be accessed by WP7 partners via the standard DICOM protocol. If the image processing result is an image itself, this should also be stored on a DICOM server. Other image-related data, such as registration information and surface meshes from image segmentation should be stored in VTK format.

Scenarios

Example scenarios for image processing were discussed during the WP6+7 workshop in November 2008 in Hamburg (PET/CT follow-up images with registration and segmentation). The description below is in chronological order and concerns image-related data only. It is intended to give a detailed description of the standard way in which such data from a patient will flow within the system.

1. A clinical partner (CP1) decides to create a new patient case in the database because complete patient data is available, e.g. from file; the data includes PET/CT images of the lung for different points in time.
2. CP1 creates a pseudonym, Ps, for the patient P.
3. CP1 creates a patient entry on the ContraCancrum's DICOM server for P
4. CP1 copies the image data for patient P under the pseudonym Ps from the hospital's PACS system to the ContraCancrum DICOM server.
5. CP1 uploads a case description for Ps to the database.
6. CP1 submits the image data for processing.
7. A notification is sent to the responsible image processing partner (IP1). This partner is responsible for registration of multi-modal PET/CT images of the lung.
8. IP1 receives the notification.
9. IP1 downloads the complete patient case of Ps (images and non-image data) to a local database.

10. IP1 uses the tools available at his/her organisation to register each CT image of Ps to the corresponding PET image.
11. IP1 uploads the resulting registration information to the database.
12. IP1 submits this data for further processing.
13. A notification is sent to the responsible image processing partner (IP2). This partner is responsible for registration of CT images of the same patient at different points in time.
14. IP2 receives the notification.
15. IP2 downloads the complete patient case of Ps (images and non-image data) to a local database. If IP2 works at the same partner as IP1, he/she may want to use the data that has been downloaded by IP1 already.
16. IP2 uses the tools available at his organisation to register the time series of CT image of Ps.
17. IP2 uploads the resulting registration information to the database.
18. IP2 submits this data for further clinical approval.
19. A notification is sent to the responsible clinical partner, CP2, who is responsible for approving the registration results.
20. CP2 receives the notification.
21. CP2 uses a verification and visualisation tool to download the registration information. The tool downloads the original images from the local DICOM server and presents them overlaid to CP2.
22. CP2 checks the registration result and eventually approves it with the help of the verification and visualisation tool.
23. The status of the registration in the database is set to "approved" and a notification is sent to the responsible image processing partner, IP3, who is responsible for segmentation of lung tumours in PET/CT images.
24. IP3 receives the notification.
25. IP3 downloads the complete patient case of Ps (images and non-image data) to a local database; if IP3 works at the same partner as IP2, he/she may want to use the data that has been downloaded by IP1 already.
26. IP3 uses the tools available at his organisation to segment the tumour and the normal tissue; registers for each of Ps' PET/CT images.
27. IP3 uploads the resulting segmentation information to the database.
28. IP3 submits this data for further clinical approval.
29. A notification is sent to the responsible clinical partner, CP3, who is responsible for approving the segmentation results.
30. CP3 receives the notification.
31. CP3 uses a verification and visualisation tool to download the registration and segmentation information; the tool downloads the original images from the local DICOM server and presents them overlaid to CP3.
32. CP3 checks the segmentation result and eventually approves it with the help of the verification and visualisation tool.

33. The status of the whole patient case Ps in the database is set to "image processing approved". Now the image processing tasks are completed and notification is sent to the relevant partners in WP4 and WP6.

Further scenarios/use cases to be considered may be:

- adding new image data to an existing patient
- removing image data from an existing patient
- removing a patient
- requesting a new approval
- requesting additional image processing.

This list is illustrative and is not intended to be complete.

Specific Requirements

Pseudonymisation

All clinical image data should be pseudonymised by the clinicians before being stored on the DICOM server. The pseudonymisation should be reversible by the clinical partner in order to be able to store new data for the same patient in the database.

Image processing results must be stored together with the pseudonym of the patient, either on the DICOM server or in the non-image database.

Storage needs

Registration results can be stored either as transformation matrices (rigid registration) or as point sets (non-rigid registration). In both cases, the storage needed is much less than for the original images. Segmentation results can be stored either as surface meshes or as label images. Label images typically have a bit depth smaller than the original images.

Therefore, on the DICOM server there should be storage allowed for twice the size of the original clinical images. In the database for the non-image data, storage for image-processing results should be allowed, maximally, for the size of the original clinical images.

Identification of processing steps

To ensure the quality of the processing, information on the processing steps must be stored in the database.

Each image processing result stored in the database should have stored with it the original data that was used to create the result.

Since image processing results depend upon the tools used, a record should be kept of the tools that were used to create the result and the person who did it.

Downloading data

It should be possible to download all data for a patient (identified by a pseudonym) to a local disk for further local processing. This would include original clinical images as well as processed images and non-image data. It should be possible to choose to download everything or to select only specific data.

Original clinical images should be accessible via the DICOM protocol to be copied to a local DICOM server. Furthermore, it should be possible to download them to a local directory.

Uploading data

It should not be necessary to upload the complete patient data at a single session; it should be possible to perform partial uploading.

Workflow management

Image processing is part of a processing chain, so a notification mechanism should be installed to inform responsible groups/users that new data is available. When uploading results from processing, such a notification should be sent automatically.

The image processing workflow may change with time, so a means by which the workflow and the notification receivers can be changed should be provided.

Data types

The DICOM server for clinical images and processed images should use the DICOM protocol for data transfer.

Whenever non-DICOM data must be stored, this should be done in the database for non-image data, which must be able to store any format (binary or text). The users of the database should be responsible for creating the necessary interfaces to read and process the data.

VTK is the preferred file format for surface meshes, rigid and non-rigid transformations.

User interface

The user interface for interacting with the data should be web-based; uploading and downloading data should be performed via a web browser.

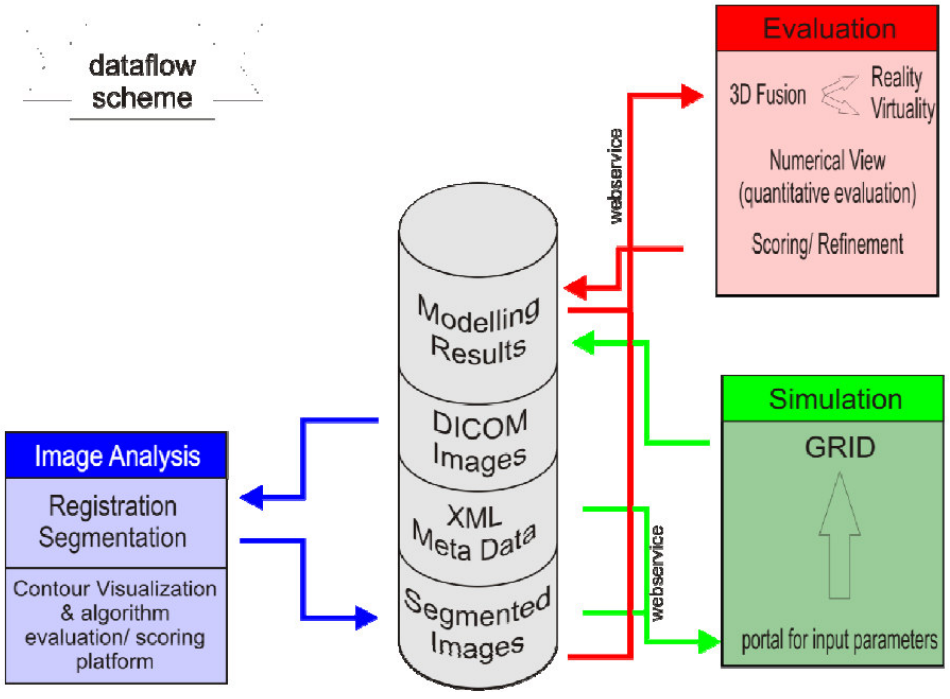

Figure 3. Data flow scheme associated with WP7

5.5 WP9: Clinical guidance, validation and translation

Objective

The objective of WP9 is to collect cross-platform clinical data, post-genomic data and image data from patients with gliomas and lung cancer. This data has to be integrated into the technological environment of ContraCancrum, so that *in silico* experiments and data analysis can be performed seamlessly and in clinically relevant timeframes. Furthermore, the results of the *in silico* experiments will be evaluated and their clinical relevance estimated.

Overview

The clinical aspects of ContraCancrum fall into three main areas:

1. bio-banking for ContraCancrum clinical studies
2. ContraCancrum clinical studies and *in silico* modelling scenarios
3. collecting and organising the ContraCancrum clinical data.

The purpose is to develop simulation models in glioma and lung cancer that can be used as a proof of principal for further exploitation in other cancer types. The emphasis is on individualised treatment of the patient and, with this in mind, the results of such modelling experiments must be available in a short timeframe after diagnosis, so all data that are necessary for running such experiments have to be available in a timely manner. Hence, the molecular biologists, radiologists and clinicians will have to produce reliable data very fast. The combined chemotherapy simulation model will be validated and optimised using pseudonymised data

Ultimately, the simulation models are expected to:

- support clinicians' decisions concerning various candidate cancer treatment schemes
- facilitate the optimisation of individualised treatment
- help to suggest new therapeutic strategies
- help to train or inform doctors, life scientists, researchers or interested patients by demonstrations of the likely tumour response to different therapeutic schemes.

Further details of the clinical aspects of ContraCancrum can be found in the combined deliverable D2.2/D9.1, *Definition of Clinical Scenarios in Cancer Modelling Studies and Validation* and *Protocols and Regulations for ContraCancrum Studies (including Bio-Banks)*.

Bio-banks

Bio-banks refer to organised collections of biological material (blood, tissues, DNA, RNA, proteins, etc.) and corresponding data. It is generally worthwhile sending normal tissue, blood and other fluids of the patient to the bio-bank as well as tissue under specific scrutiny.

Each bio-bank requires a database to support storage and access to the data; a suitable database is described in Deliverable D2.2/D9.1. Material donation has to be performed ethically, and informed consent must be obtained from the donor.

The bio-bank should have standard operating procedures (SOP) as part of its quality control mechanisms; this will also ensure that appropriate ethical and legal standards are applied throughout.

Simulation Models

From a clinical point of view, two preconditions are of the utmost importance, if one is to be able to trust predictions of *in silico* methods:

- every *in silico* method has to be part of a clinico-genomic trial

- every prediction of an *in silico* method has to be compared with reality.

This implies that data management, including anonymisation/ pseudonymisation of data, data security and tumour banking, are well established and that the trial is reviewed by an ethical committee and fulfils all other criteria to gain approval by regulatory authorities.

To make the simulation predictions as precise and realistic as possible, it is crucial to obtain as much information from each of the different categories of data – from the tumour (molecular biology, pathology, imaging), from the patient (clinical data) and from the possible treatment (pharmacokinetics of drugs that will be used, the treatment schema). Standards have to be defined regarding the clinical data that are needed, the imaging studies to be performed, the segmentation process to be employed for rendering the tumour and the molecular genetic data that is to be analysed. Minimum requirements for various types of data are provided in Section 4 of D2.2/D9.1. Only anonymised or pseudonymised data will be stored and used in ContraCancrum.

Data Management

The data to be used for gliomas and lung cancer are given in D2.2/D9.2, Sections 6 & 7, respectively. ObTiMA will be used as the data-management system; this was developed in the ACGT project and will be expanded to accommodate gliomas and lung cancer.

Workflows

Figures depicting the workflows for gliomas and lung cancer are presented below. These are reproduced from Sections 6 & 7 of D2.2/D9.1, which also provide notes on the respective validation processes. Further information is given in chapter 4.4 of this deliverable.

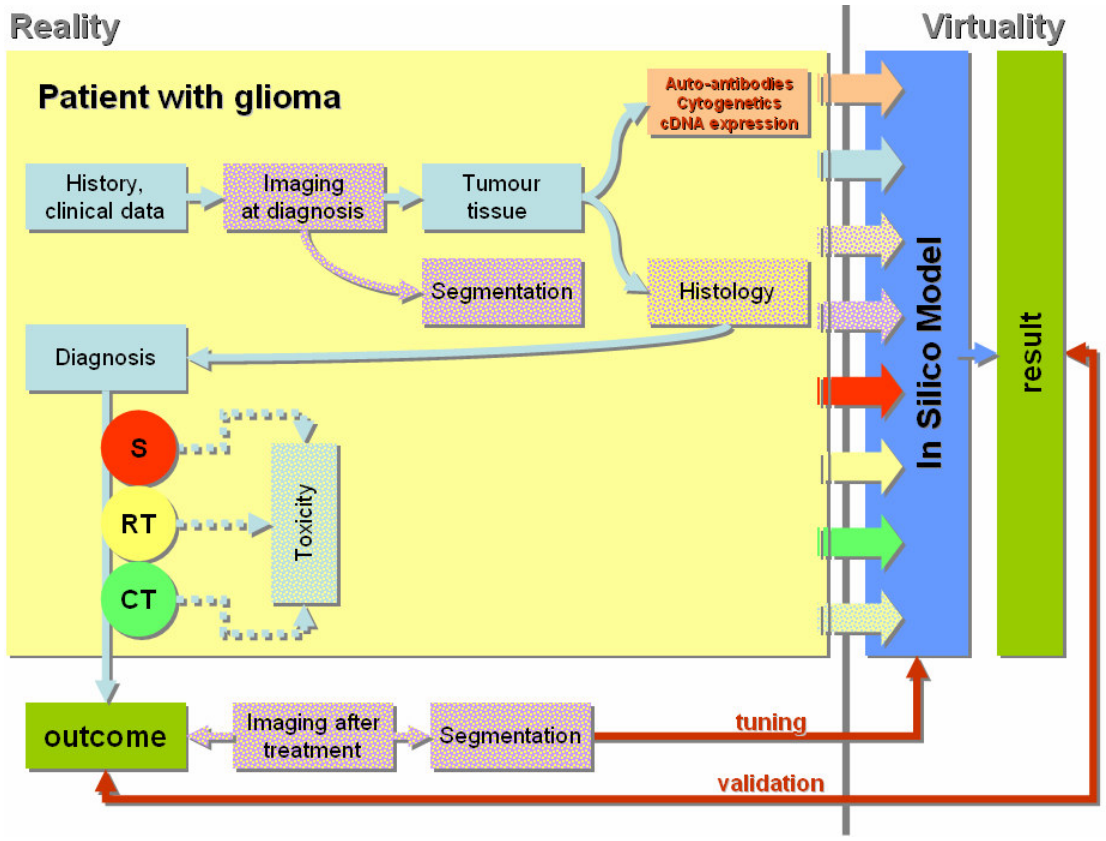

Figure 4. Workflow of the glioma scenario. S: surgery; RT: radiotherapy; CT: chemotherapy

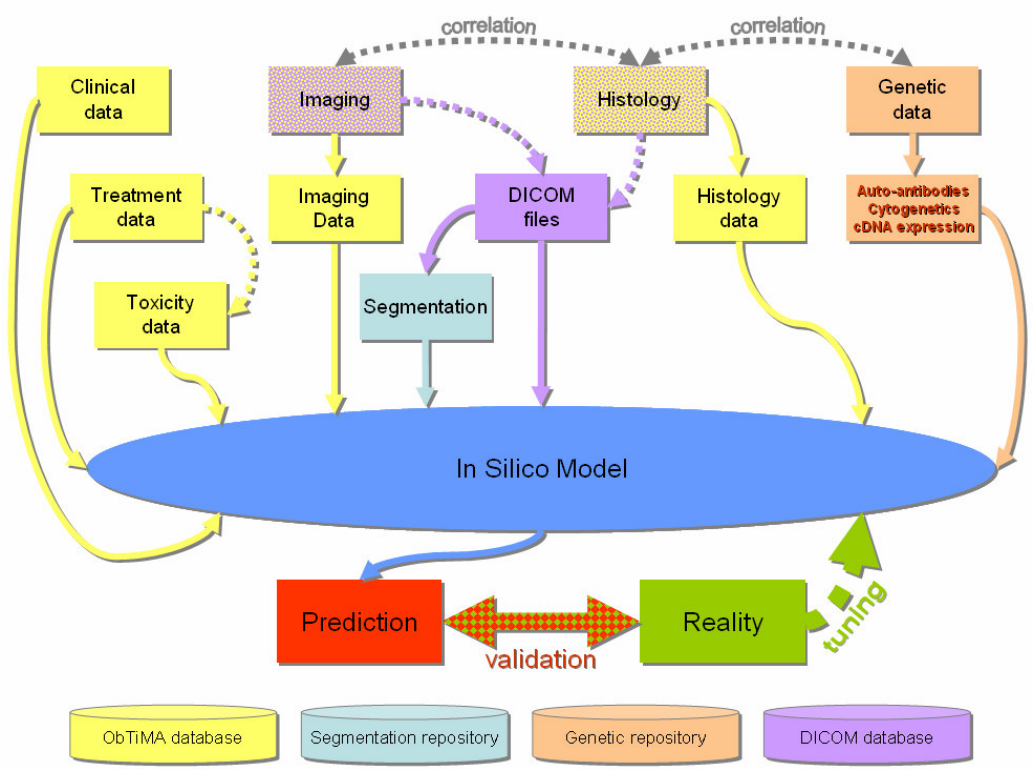

Figure 5. Workflow of data of the glioma scenario

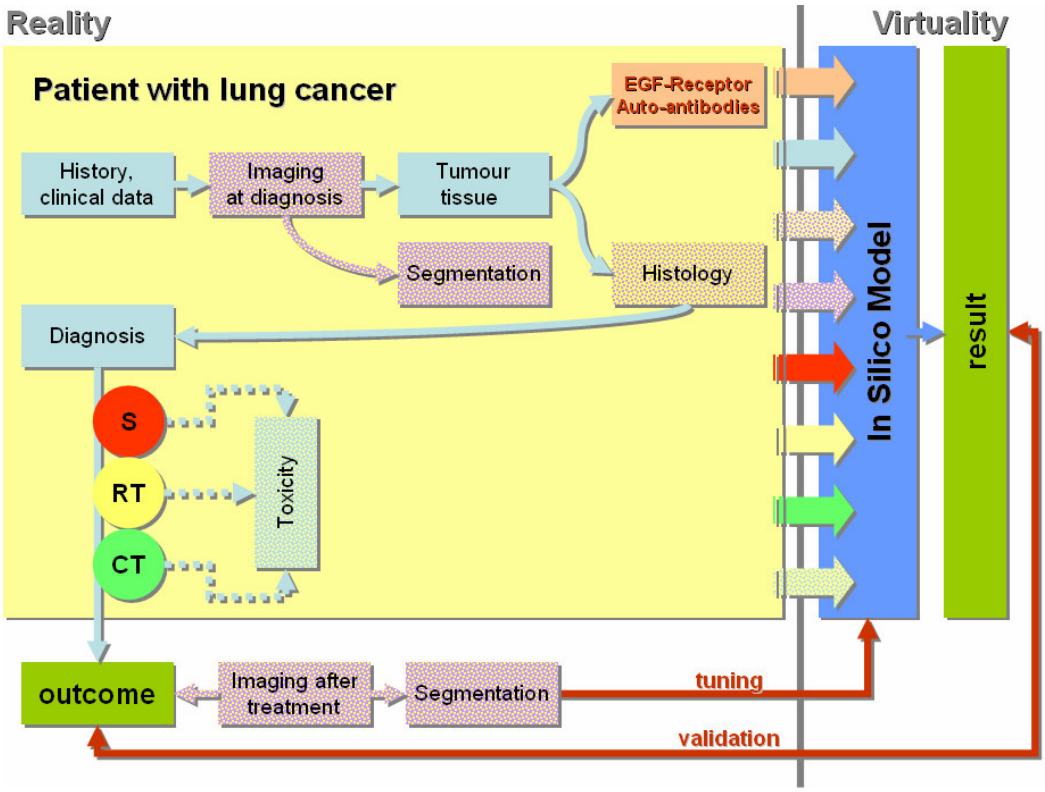

Figure 6. Workflow of the lung cancer scenario. S: surgery; RT: radiotherapy; CT: chemotherapy

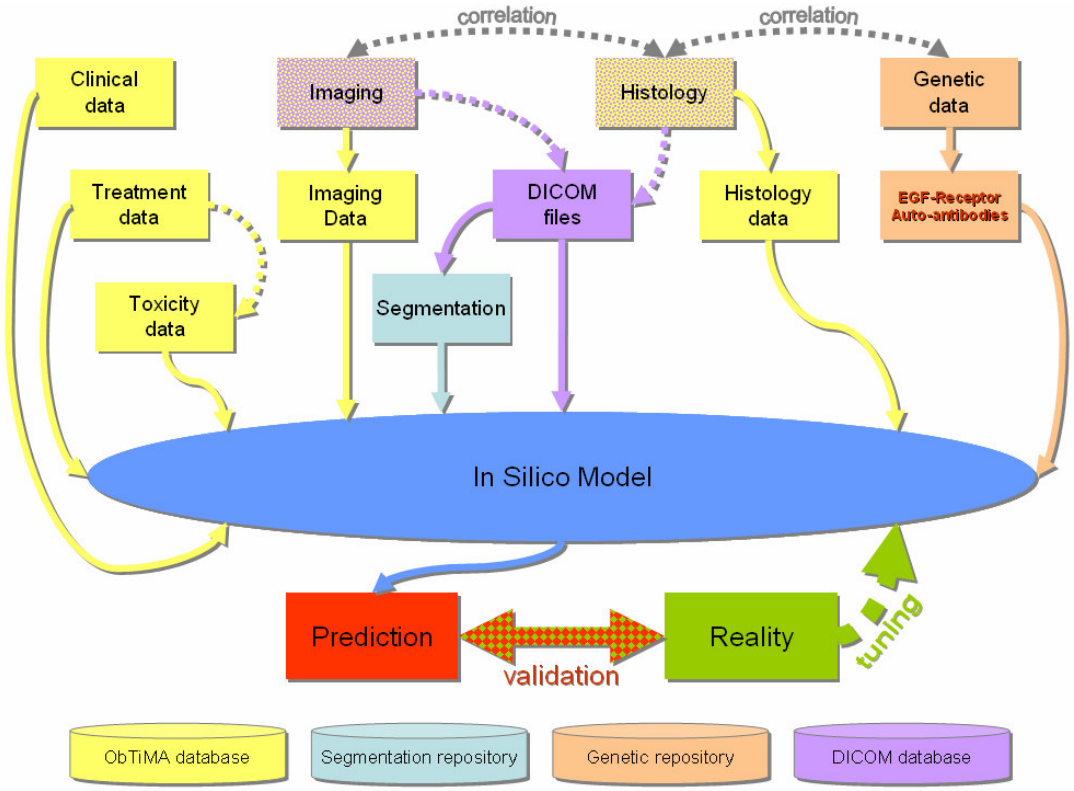

Figure 7. Workflow of data of the lung cancer scenario

The segmentation can be performed in a semi-automatic way by the clinician and in an automatic way by the image processing partner. Both approaches will be followed during the project. The results will be compared, and the best fitting segmentation, as chosen by the clinician, will be used in the *in silico* model. The segmentation tool to be used by the clinicians is DoctorEye, developed by FORTH.

Pseudomysed sequencing data of the EGFR (Epidermal Growth Factor Receptor) will be uploaded to a repository for molecular biological data and then processed before being entered into the *in silico* model. The same will be done for the Autoantigen data and other molecular biological data. The location of the repository for these data will be made known to project partners, when available.

Interface

The interface should be a Web-based portal that offers role and rights management and allows the sharing of data, including clinical data, by use of ObTiMA or some other patient-data management system, imaging data via a DICOM Server, and molecular biological data

The portal should provide the use/download of tools for data pre- and post-processing, such as a segmentation tool and a pseudonymisation/anonymisation tool; the latter will have to be used before any data can be uploaded. It should also give support with regard to the execution of workflows and allow the retrieval of results of models for a single patient. These results should be visualised and/or compared with the real data/imaging

6. Summary

This deliverable has looked at the system requirements from the perspectives of the individual workpackages. Given the diversity of activities across the WPs, the picture that has emerged is heterogeneous and is not open to simple categorisation.

This implies that the system developers will have to adopt a distributed communication model in order to maintain a close and continuous dialogue with all of the WPs on an individual basis rather than having a more centralised approach. This will ensure that the broad range of needs is fully met in terms of both the short-term evolution of the system throughout the project and the ultimate form that the system takes by the time the project ends.