Clinically Oriented Translational Cancer Multilevel Modelling

Deliverable D2.2

Definition of clinical scenarios in cancer modelling studies and validation

&

Deliverable D9.1

Protocols and regulations for ContraCancrum studies (including bio-banks)

Project Number:	FP7--IST-*223979*
Deliverable id:	D2.2 & D9.1
Deliverable Name:	Definition of clinical scenarios in cancer modelling studies and validation & Protocols and regulations for ContraCancrum studies (including bio-banks)
Submission Date:	30/04/2009

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	CONTRACANCRUM
Project Full Name:	Clinically Oriented Translational Cancer Multilevel Modelling
Document id:	D2.2 & D9.1
Document name:	Definition of clinical scenarios in cancer modelling studies and validation & Protocols and regulations for ContaCancrum studies (including bio-banks)
Document type	RE
Version:	2.0 final
Submission date:	30/04/2009
Editor: Organisation: Email:	Norbert Graf USAAR graf@uks.eu

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

ABSTRACT:

This deliverable provides definitions of clinical scenarios in cancer modelling studies. The data used in these studies as well as the workflows are given. The kind of validation of the models is described. This part belongs mainly to the deliverable D2.2. Protocols and regulations for ContraCancrum studies including bio-banking are explained. This part belongs mainly to deliverable D9.1. As D2.2 and D9.1 are linked both deliverables are joined for harmonisation purposes and to avoid writing repeated information. The whole document is divided into three main parts: bio-banking and regulations in ContraCancrum clinical studies, ContraCancrum clinical studies and in silico modelling scenarios, and ContraCancrum clinical data. As ContraCancrum focuses on glioma and lung cancer this deliverable deals with protocols and requirements for the integration of clinical, imaging and molecular data in these diseases to simulate tumour growth and response to treatment. The purpose is to develop simulation models in both diseases that can be used as a proof of principal for further exploiting them in other cancer types.

KEYWORD LIST: glioma, lung cancer, in silico oncology, cancer modelling, multiscale bio-modeling

MODIFICATION CONTROL			
Version	Date	Status	Author
0.1	20/03/2009	Draft	Norbert Graf
0.2 – 0.4	10/04/2009	Draft	Norbert Graf
0.5	14/04/2009	Draft	Norbert Graf
0.6	03/05/2009	Draft	Norbert Graf
0.7	06/05/2009	Draft	Norbert Graf
1.0	10/06/2009	Draft	Norbert Graf
2.0	10/07/2009	Final	Norbert Graf

List of Contributors

Major contributions:

Norbert Graf	USAAR, Germany
Eckart Meese	USAAR, Germany
Rainer Bohle	USAAR, Germany
Kostas Marias	FORTH, Crete
Amos Folarin	UCL, UK

All other partners did contribute as well

Contents

Publishable Executive Summary

The present document contains protocols and regulations for ContraCancrum studies. ContraCancrum focuses on glioma and lung cancer. Protocols and regulations are dealing with integration of clinical, imaging and molecular data in these diseases to simulate tumour growth and response to treatment. The purpose is to develop simulation models in both diseases that can be used as a proof of principal for further exploiting them in other cancer types. The document begins with a brief introduction of what will be one of the final goals of ContraCancrum shown by a concrete example. Subsequently both diseases (glioma and lung cancer) are explained from a clinical point of view. The data that will be provided from the clinical site are explained in detail. Bio-banking is discussed in detail. Scenarios and workflows for both diseases are given demonstrating how the simulation model will work and explaining what are the expectations from a clinical perspective. A validation process of the simulation model for both cancer types will be described using real patient data. Protocols for the scenarios in both diseases are written in summary and will be used in prospective clinical trials in future according to the regulations for clinical trials in Europe and in particular in Germany. All regulations needed for the simulation model from a clinical perspective are discussed. Both prospective clinical trials will start as single centre trials after approval at the University Hospital of the Saarland. Summaries of the trial protocols are included in this deliverable and are given in the appendix. Further appendixes provide more detailed information regarding several aspects of the work.

1. Introduction

This document describes the requirements of the *in silico modelling* to be developed within the ContraCancrum from a clinical perspective and provides information about data, workflows and regulations within two malignant diseases namely glioma and lung cancer. Cancer modelling is a complex and multiscale combination of sciences and technologies in order to simulate malignant tumour growth and tumour and normal tissue response to therapeutic modalities at all levels of biocomplexity.

The aim is to better understand cancer and related phenomena and to optimize therapeutic interventions by performing in silico (on the computer) experiments based on the individual data (clinical, imaging, histopathologic, molecular) of the patient.

The objective is to develop a technologically advanced and user friendly system able to spatiotemporally simulate within well defined reliability limits tumour growth and tumour and to a lesser extent normal tissue response to chemotherapy for glioma and lung cancer in the patient's individualized context. Pertinent clinical, imaging, histopathologic and molecular data in conjunction with ContraCancrum clinical trials will be exploited in order to validate the model both prospectively and retrospectively. For this purpose the definitions of the clinical scenarios in glioma and lung cancer modelling studies are provided. The data used in these studies as well as the workflows are given and regulations for ContraCancrum studies including bio-banking are explained.

2. Objectives of the Deliverable

This deliverable provides definitions and requirements of clinical scenarios in cancer modelling studies. The data used in studies for glioma and lung cancer as well as the workflows are given. The kind of validation of the models is described. Protocols and regulations for ContraCancrum studies including bio-banking are explained. The whole document is divided into three main parts giving the objectives of the deliverable:

1. bio-banking and regulations in ContraCancrum clinical studies
2. ContraCancrum clinical studies and in silico modeling scenarios
3. ContraCancrum clinical data.

As ContraCancrum focuses on glioma and lung cancer this deliverable deals with the integration of clinical, imaging and molecular data in these diseases to simulate tumour growth and response to treatment. The purpose is to develop simulation models in both diseases that can be used as a proof of principal for further exploiting them in other cancer types.

The question “What is the best treatment for a given tumour in an individual patient?” leads to a high level of individualization. To attain this goal it is of utmost importance that the results of such modeling experiments are available in a short timeframe after diagnosis. This implies that all data that are necessary for running such experiments have to be available in a timely manner. This is especially important for molecular biologists, radiologists and clinicians, who have to produce reliable data very fast. The experiment run by the ACGT Oncosimulator itself is not a time-consuming experiment.

The combined chemotherapy simulation model is validated and optimized using pseudonymized data from clinical trials as a proof of principal. The models to be developed have to be able to reproduce important aspects of the clinical reality and practice and generally produce reasonable predictions. Nevertheless, in order to enter routine clinical practice as a decision making tool, an exhaustive validation, adaptation, and optimization procedure has to take place. Furthermore, molecular methods of extraction of crucial histological constitution of tumours need to be tested and integrated into the models. Following completion of such testing procedures the simulation models are expected

- to support clinicians’ decisions concerning various candidate cancer treatment schemes
- facilitating individualized treatment optimization
- helping to suggest new therapeutic strategies
- helping to train or inform doctors, life scientists, researchers or interested patients by demonstrations of the likely tumour response to different therapeutic schemes

A successful performance of cancer modelling studies is seen as a strongly encouraging step towards the clinical translation of such integrated systems. These models can be adjusted to any other cancer type using data as described in this deliverable. It is always necessary to validate such systems by going through a learning loop until the prediction of the simulation is corresponding to the reality of the tumour response in an individual patient.

Bio-banking and regulations in ContraCancrum clinical Studies

3. Bio-banking

Human specimen like blood, tissues, DNA, RNA, proteins and other biological material are fundamental resources in medical research today. Bio-banks refer to organized collections of biological material and corresponding data. Research using biomaterial and data aims to discover the genetic basis of diseases and provide novel diagnostics, predictive biomarkers, and therapeutics that are safe and more efficacious by understanding for which patients those products are appropriate. Bio-banks have to be strict organized to provide high-quality bio-specimens to support these kinds of discoveries in basic research of diseases.

Bio-banks have to meet two goals. They have to foster **translational research** and **molecular profiling**. Translational research brings knowledge from the laboratory bench to the patient's bedside whereas molecular profiling needs patient's biomaterial for analyzing a variety of bio-molecules, including DNA, RNA, proteins, and metabolites. This approach closes a loop between scientific research and clinical medicine (fig. 3.1).

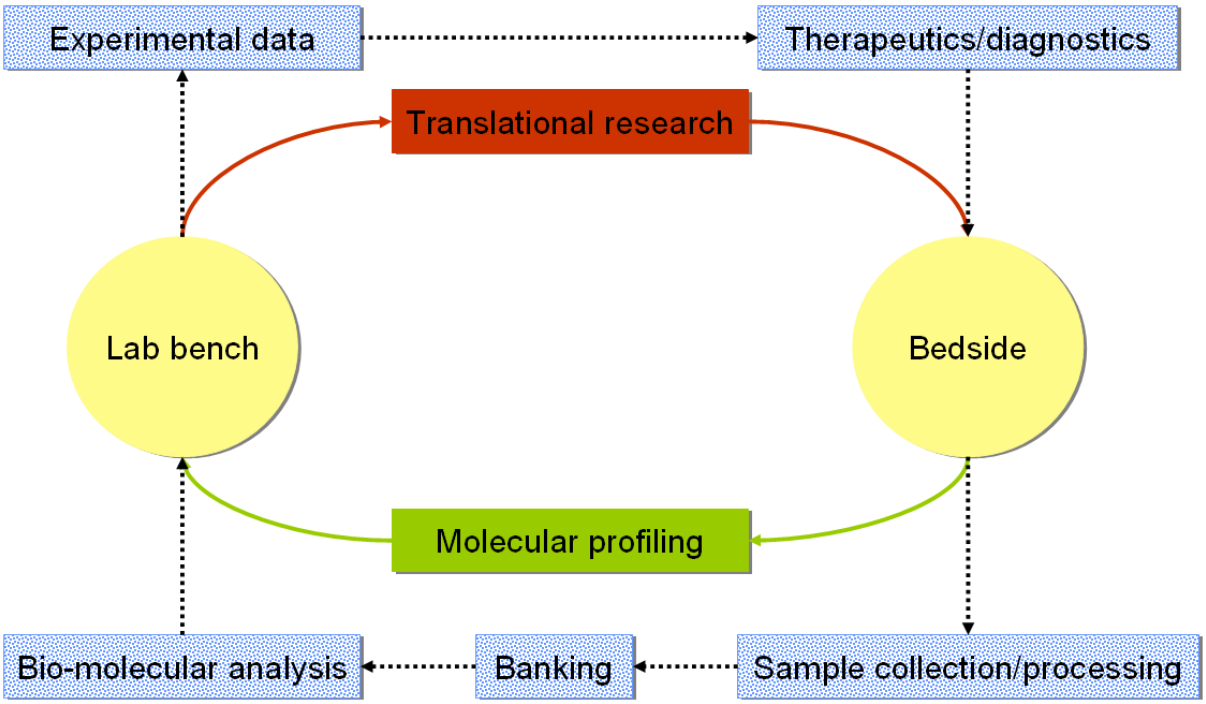

Fig. 3.1: Schematic view of the relationship between Laboratory and bedside (an adaptation of figure 1 from Bio-banks: Accelerating molecular medicine. Challenges facing the global bio-banking community[1])

Bio-banks must have the capacity to house large collections of well-defined and processed biological samples and corresponding data sets. This capability is a key for longitudinal clinical studies that are tracked over extended periods. Longitudinal data are critical for understanding the long-term benefits and/or side effects of drugs and various treatment options as well as the influence of various environmental factors. Bio-banks use different approaches to exploit tissue samples. Some collect material from a specific population whereas others study a specific disease. Most of them are project-driven and collect tissues from specific diseases or organs. Because of the associated ethical and

[1] BIOBANKS: ACCELERATING MOLECULAR MEDICINE. Challenges Facing the Global Biobanking Community. November 2004, IDC #4296. http://www-03.ibm.com/industries/global/files/Biobanks_Accelerating_Molecular_Medicine.pdf (last access: 13th April 2009)

regulatory issues surrounding the storage of an individual's genetic and medical information, a well-articulated, transparent and protected operation of a bio-bank is critical and of utmost importance to build public trust and to avoid abuse.

There are many open questions regarding bio-banking today. These questions can be divided into three categories:

- Scientific questions including logistics of bio-banks, e.g.:
  - *What are the best practise for sampling, collection, processing, annotation, storage and distribution of biomaterial (logistics)?*
  - *What are quality control mechanisms of bio-banks?*
  - *How can bio-banks work together and share samples?*

- Ethical and legal issues, e.g.:
  - *How is privacy of a patient protected?*
  - *Have samples to be donated by patients?*
  - *Who is the owner of a sample?*
  - *How has an informed consent look like?*
  - *What information has to be given to a patient after donation of a bio-specimen?*
  - *What are the ethical and legal issues in exchanging samples, especially across nations?*
  - *What are funding models of bio-banks?*
  - *How to deal with commercial issues?*

- Information technological issues, e.g.:
  - *What are necessary items in a basic dataset that allow to store the kind, processing and amount of samples, the research done with these samples, and the results coming up using these samples?*
  - *How to protect privacy from an IT point of view?*
  - *How to store results from different research laboratories using the same samples?*
  - *What Ontologies are needed for bio-banking databases?*

The following chapters will address these questions and give concrete answers for ContraCancrum in a way that the research projects in ContraCancrum regarding glioma and lung cancer are able to be done.

Logistics of bio-banking

Most important questions of bio-banks deal with logistic issues. In the following section the sampling, collection, processing, annotation, storage and distribution of bio-specimen is described in general for patients with cancer and concretized for the situation in ContraCancrum.

In every patient with a malignant disease diagnosis is done by analyzing tumour material. In solid tumours at least a biopsy of the tumour is performed and sent to a pathologist for histological diagnosis. Today in every cancer patient it should be standard to send fresh tumour tissue without fixation (formalin) directly to the pathologist, who will perform primarily the macroscopic staging of the tumour and then takes pieces out of the vital tumour for shock freezing, cell culture and fixation. Before a piece of the tumour is shock frozen immediately a small part of this piece will go to histological examination to provide histological diagnosis of the shock frozen specimen. This has to be done in every case and also for specimen going to cell culture or other investigations. Especially for heterogeneous tumours this is an important procedure and part of the quality control of bio-banking to be assured that the genetic analysis deals not with necrotic material or normal tissue.

It is also worthwhile to send normal tissue, blood and other fluids of the patient to the bio-bank. In some cancers, like familial inherited cancer syndromes or in patients who have a germline mutation material from family members might be useful for scientific analysis and to detect family members on risk for the specific cancer type. These biological specimens are also sent to the bio-bank.

To send specimen to the laboratory where bio-banking (bio-banking-lab) is done special boxes should be available, that guarantee that the sample will not melt during the transport to the laboratory. Fig. 3.2 shows such a ‘tumour box’ as it is used in Paediatric Oncology in Germany. After arriving in the bio-banking-lab the empty box will be send back to the local hospital so that for the next biological specimen a tumour box is available for transportation.

Tumour material from patients who are enrolled in clinical trials is send to the reference pathology of the trial to provide a uniform diagnosis and to do extra staining of the tumour according to the trial protocol.

In the laboratory where the bio-banking is done the tumour material will be processed and DNA, RNA and proteins will be extracted. The amount and quality of DNA, RNA and proteins will be checked and portioned in several vials for storing at – 80°C. Material for cell culture will be prepared according to the conditions for cell culture for the specific tumour type.

Fig. 3.2: Tumour box as used in Paediatric Oncology in Germany

There are standard operating procedures available for all steps of tumour material preparation and storing in the bio-banking-lab. This is part of the quality control of the laboratory.

To gain most of clinical data bio-banking should always be part of clinical trials. In single centre trials the bio-banking can be done in the local hospital. In multicentre trials the bio-banking should be centralized. Such centralization has to guarantee a highest quality of the bio-material stored. It will facilitate the track back of biological information in a trial if only one bio-bank is used.

In case of a single bio-bank it is necessary to establish a scientific board that guarantees the distribution of material to researchers according to scientific criteria. Every bio-bank must provide forms for researchers, where they can apply for receiving biological specimen (see chapter 'ethical and legal issues').

Each bio-bank must have a database that stores at least the kind, processing and amount of samples, the research done with these samples, and the results coming up using these samples. In ACGT such a database is defined in D2.2[2]. This database is explained in the chapter 'database for bio-banks' and will be used for ContraCancrum for the bio-banks for lung cancer and gliomas.

Figure 3.3 and 3.4 gives a schematic overview of bio-banks and bio-material.

[2] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. pp 54-61, 15th September 2007

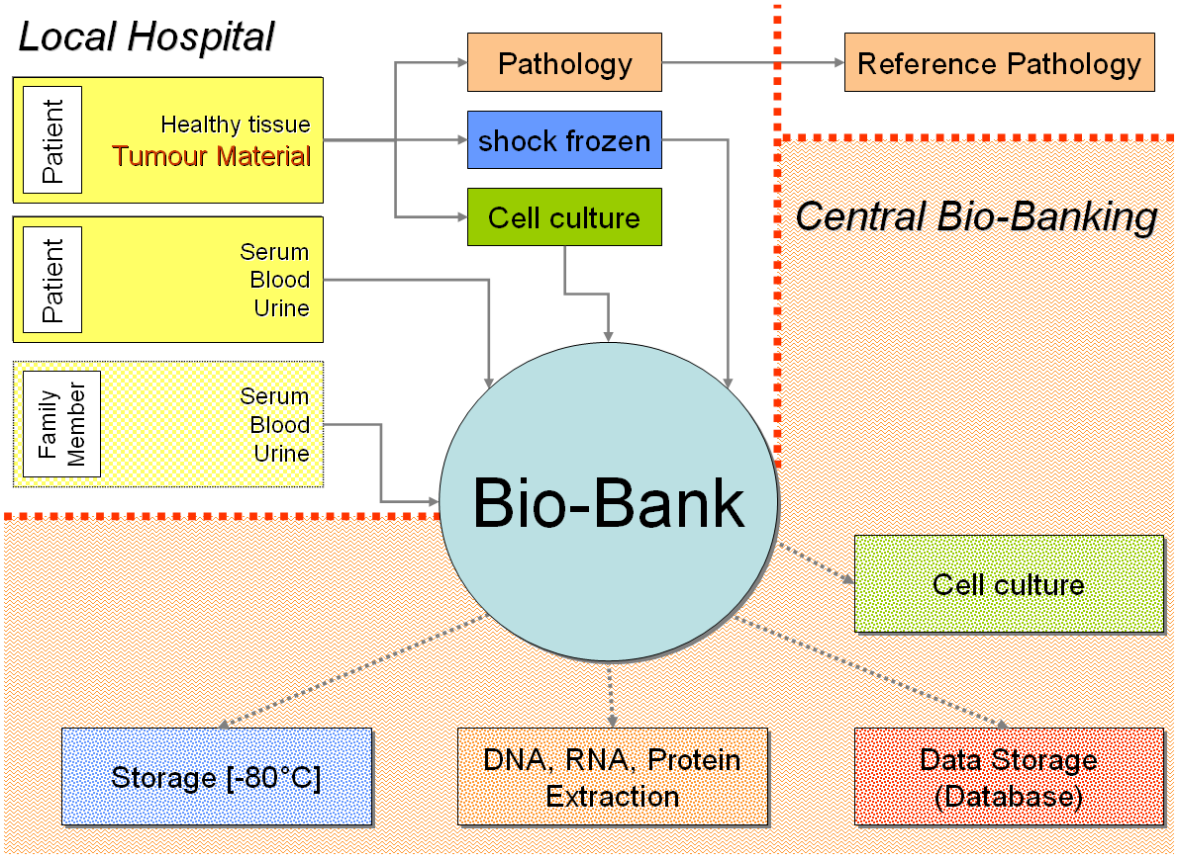

Fig. 3.3: Logistics of bio-banking: From local hospital to central bio-banking.

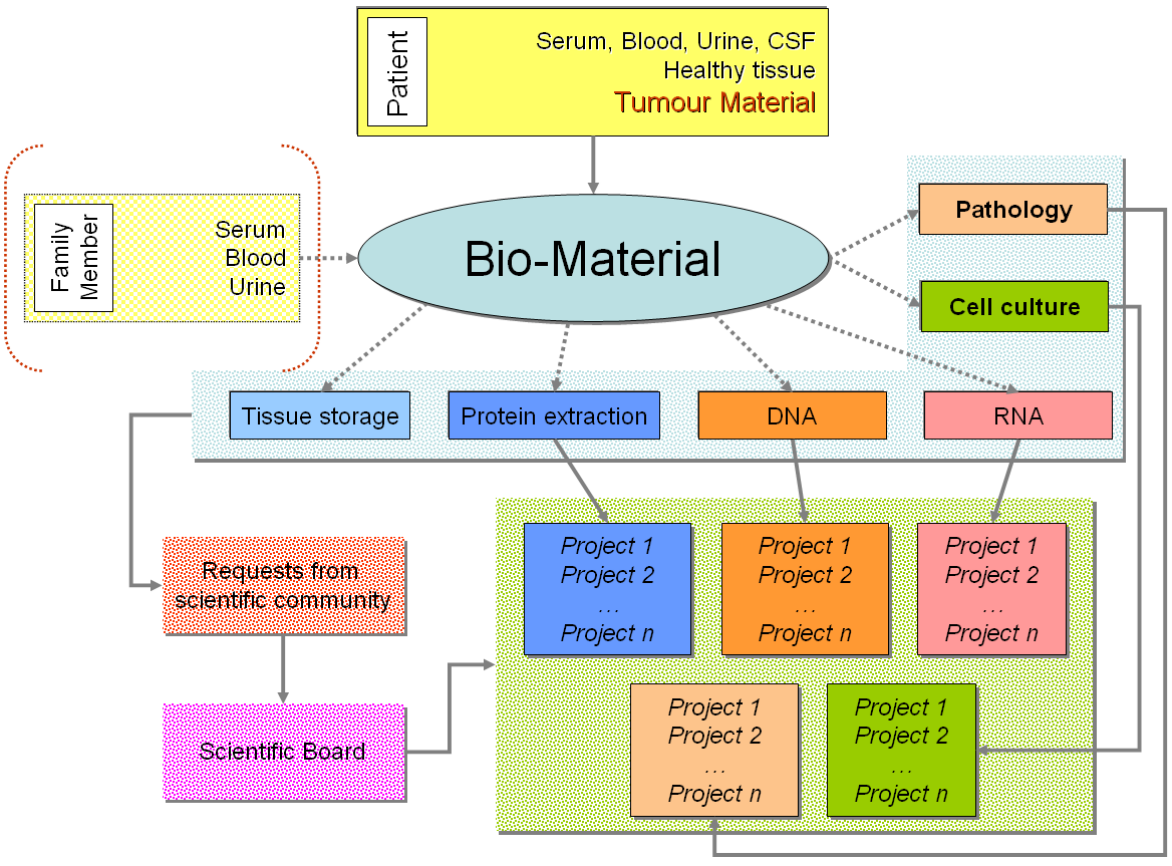

Fig. 3.4: Logistics of bio-banking: Research projects

Ethical and legal issues

Because of the complexity of the healthcare, quality, security, and privacy issues surrounding bio-banks, governance is a key issue. Recognizing this fact, many governments have established regulations concerning the operations of bio-banks at both local and national levels. These regulations deal mainly with the protection of the privacy of patients and funding issues and regulate how human biological material is to be collected, stored, and used for certain purposes. Bio-banks have to adhere to these regulations. The problem is the missing standardization across nations. For multicentre clinico-genomic trials involving different countries even on the European level makes it difficult to easy exchange data and biological specimen. The last mentioned problem will and can not be solved within ContraCancrum. As both clinical scenarios do need data and bio-material from one single centre (University Hospital of the Saarland) Standard Operation Procedures (SOP) will be established for the local banking of bio-material from lung cancer and glioma patients and corresponding clinical data. The bio-banking will be part of the prospective clinical scenarios and trials that are established for both tumour types. Regarding the exchange of data and bio-material the regulation regarding German law will be followed and ethical approval will be obtained by the local Ethics Committee (Ethikkommission der Ärztekammer des Saarlandes).

A comprehensive overview of the German law regarding bio-material is given in section 3.13.1 of the ACGT deliverable 12.2[3] and cited here:

"There is no federal German law concerning human tissues or bio-banks. The question of tissues and cadaver samples are regulated by common law. Nationwide rules can only be enforced by criminal law. According to the patrimony law ("Erbrecht"), the human body is not considered an object during life, so that it is not part of the heritage. As a consequence, a dead body has no owner in Germany. The legal condition of parts and fluids separated from the body has to be examined carefully: during life, the integrity of one's own body is a fundamental right of the person. Once parts are separated from the body, they are regulated by both property and personality right law (overlap thesis). Interestingly the property rights and the right to privacy can belong to two different persons, and they differ in their scope of protection: property of samples can be completely transferred, but the right to privacy cannot! In practice, if a patient gave no particular directive for the use of his tissues/body fluids, the physician (*e.g.* the pathologist) is free to trash them or to use them for quality control purposes in the laboratory or for research purposes if anonymization is warranted, if no immortalization is undertaken, and if the research project is not ethically debated. However, if the samples are not only used for research, but also for therapeutic purposes, the need for a specific informed consent given by the patient is increasing, in particular with large samples, or in the presence of personal characteristics (*e.g.* genetic material). For the researcher, the duty remains to protect the right to privacy of the donor/or the patient[4].

- If the banking of human genetic material includes the collection and storage of personal data associated with an identifiable person, the Data Protection Law has to be taken into account. The German Data Protection Law is very complex, since the applicable law depends on the status of the data collecting and storing institution (public, private, federal, state), it contains different permissions for collecting, storing, using, transferring for own or for other purposes and there are numerous exceptions in other laws. The Data Protection Laws are subsidiary towards these special provisions in certain areas.
- *Federal Data Protection Law (BDSG) 1990.* Non public DNA banks, which process personal data, have to take into account the Federal Data Protection Law if data are processed into or out of file commercially or for professional or commercial purposes. The processing of personal data and their use is only allowed if there is a legal permission or consent of the affected person. The transmission of personal data is permitted if it is necessary for scientific research. This permission does not apply for medical data.
- *Data Protection Law of the states.* The application field of the data protection laws of the states include all public authorities of the respective states. They have to be applied in public hospitals or the state or the municipalities. The University hospitals are either non-self-maintained public state institution or state companies with restricted independence. If these institutions bank identified or

[3] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. pp 54-61, 15th September 2007

[4] Reymond MA, Steinert R, Eder F, Lippert H. Ethical and regulatory issues arising from proteomic research and technology. Proteomics. 2003;3(8):1387-96

identifiable human genetic material, they have to take the state data protection acts into account. These laws are subsidiary towards special data regulation in law, specific for certain areas, for example the hospital laws.

- *Law of 20 June 1990 to regulate matters relating to gene technology.* The aims of this law are: (1) to protect life and health of human beings, animals and plants against possible threats of gene technology, and (2) to give a legal framework for research, development and support of scientific, technical and economic possibilities of gene technology.
- *The German Bundestag, Chancen und Risken der Gentechnologie Enquete-Commission, 1987.*
- *Resolution of the Conference on the Privacy of Information of Rheinland–Pfalz on the Subject of Gene Analysis and Informational Self-Determination (26–27 October 1989).*
- *The Federal and State Governments, Final Report of the Bund–Länder–Arbeitsgruppe Genomanalyse, 1990.*
- *Minister of Research and Technology, Arbeitskreis Genforschung Report, Frankfurt, 1991.*
- *German Parliament, Periodic Report of the Büro für Technikfolgen, 1992, 2000*[5]*.*

In 2004, the Biobank Working Group of the German Telematics Platform for Medical Research Networks (TMF) initiated a project to construct a generalized legal basis for the establishment and operation of bio-banks in Germany. Most bio-banks in Germany are currently operated by public institutions such as university clinics, institutes or departments. However, a substantial proportion of these bio-banks are considered "going private". This project involved the planning, writing and evaluation of an expert report that addresses in great detail the legal issues concerning property rights, medical professional regulations, general liability insurance, resource continuity and research secrecy[6]."

Protection of patients Privacy in ContraCancrum

All data that will be stored in ContraCancrum will be pseudonymized or anonymized to guarantee the privacy of personal data. Pseudonymization or anonymization is done at the University Hospital of the Saarland, before data are stored in the databases of ContraCancrum. This has to be done for all clinical data, molecular data, imaging data and all other data containing personal information. Personal data that have to be anonymized are:

- Name
- Birth date
- Place of residence
- And those given in 8.1.1 in this deliverable

Every patient has to give his/her informed consent for storing, analyzing and transferring his data and/or biomaterial. All institutions are listed who will have access to the data.

Donation of biomaterial samples

All patients are asked to donate their samples for research purposes. This is included in the informed consent form for taking part in the ContraCancrum scenarios and trials.

Owner of biomaterial and funding of ContraCancrum bio-banks

Owner of the biomaterial stored in the bio-banks for glioma and lung cancer will be the person responsible for the bio-bank. He has to guarantee the safe operation of the bio-bank according to the legal and ethical regulations mentioned above. He is also responsible for the funding of the bio-bank, so that the material can be safely stored at least as long as the clinical trials are running.

[5] Godard B, Schmidtke J, Cassiman JJ, Aymé S. Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits. A professional perspective. Eur J Hum Genet. 2003;11 Suppl 2:S88-122

[6] Simon J, Paslack R, Robienski J, Cooper DN, Goebel JW, Krawczak M. A legal framework for bio-banking: the German experience. Eur J Hum Genet. 2007;15(5):528-32

Informed consent

Patients taking part in the research project of ContraCancrum will receive an Information sheet explaining to them the research project (ContraCancrum) (appendix 3).

A template of the informed consent is given in Appendix 4 of this deliverable. It will be translated into German language, because patients taking part in the scenarios and trials of ContraCancrum are treated at the University Hospital of the Saarland. The informed consent addresses the following subjects:

- Voluntariness of taking part in the research project of ContraCancrum
- Pseudonymization / anonymization of personal data
- Use of clinical, imaging and genetic data
- Donation of biomaterial
- List of people/institutes/hospitals that will have access to his data

The single research projects will be explained to the patient but not written in the informed consent. The patient is asked to agree to research projects in general.

Commercial issues

There are no commercial issues within the research of biological material of patients enrolled in ContraCancrum scenarios and/or trials.

Standard Operating Procedure for Bio-Banks

Standard operating procedures (SOP) for bio-banks are part of the quality control mechanisms to operate bio-banks efficiently and to apply ethical and legal standards that are required. These SOPs can be divided into the following subcategories:

- Tissue or bio-material handling
- Processing of bio-material
- Storing of bio-material
- Application and distribution of bio-material to the scientific community
- Bio-banking database
- SOP for funding and maintaining the bio-bank
- SOP for guaranteeing data security and patient safety

These SOPs are under development for the ContraCancrum bio-banks.

Database for bio-banks

A database for bio-banking is described in D2.2 of the ACGT project[7]. The tables of this database are given here and can be seen as a minimal basic dataset for bio-banks. Further items can be added as needed.

For ContraCancrum trials the dataset for Bio-banking will be the same as for ACGT clinico-genomic trial. Such a dataset can be divided into different parts:

- Dataset for managing a bio-bank (logistics)
- Dataset for describing the stored biomaterial
- Dataset for describing the research that is done by whom
- Dataset dealing with the requirements for a bio-bank taking part in ContraCancrum

The description of the format of the data is given the column ‘type’ and explained in chapter 8. Reference to the TCGA data portal (http://cancergenome.nih.gov/dataportal/data/about/) is given.

[7] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. pp 54-61, 15th September 2007

Bio-bank logistics

Context	Category		Type	Format/ Coding/Options	TCGA Reference
Bio-bank	Name of the Bio-bank		A50		
	ID of the Bio-bank in ContraCancrum		I5		
	Localisation		I5	Institution-ID	
	Responsible person		I5	User-ID	
	commercial Bio-bank		I2	-1: not yet known 1: no 2: yes	
		If yes, costs for storage / vial and year	R8.2	€ for storing material	
	service provided		I2	-1: not yet known 1: no 2: yes	
		If yes, please specify	A254		
Opening Date	Date since running		D8		
Specification, what the Bio-bank is able to do (multiple entries possible)	Normal tissue	will be stored	I2	-1: not yet known 1: no 2: yes	
		If yes, specify	A25	Which tissue	
			I2	-1: not yet known 1: cell culture 2: DNA 3: RNA 4: Protein 5: Paraffin material 6: other	
		If other, specify	A25		
	Tumour material	will be stored	I2	-1: not yet known 1: no 2: yes	
		If yes, specify	A25	Which tissue	
			I2	-1: not yet known 1: cell culture 2: DNA 3: RNA 4: Protein 5: Paraffin material 6: other	
		If other specify	A25		

	other tissue	will be stored	I2	-1: not yet known 1: no 2: yes	
		If yes, specify	A25	Which tissue	
			I2	-1: not yet known 1: cell culture 2: DNA 3: RNA 4: Protein 5: Paraffin material 6: other	
		If other specify	A25		
Size of Bio-bank	Number of samples that can be stored		I7		
	Number of samples that are already stored		I7		
Information about the Bio-bank	provided		I2	-1: not yet known 1: no 2: yes	
		If yes, specify	M	Memo	
		Kind of Information	I2	-1: not yet known 1: only general Information 2: Templates for contracts	
		Shipping conditions	I2	-1: not yet known 1: no 2: yes	
		Homepage	A254	http://	
		Contact email	A100	Include @	
Bio-bank committee	Names	User-ID	I5	Multiple entries	

Stored biomaterial from a patient

Context	Category	Type	Format/ Coding/Options	TCGA Reference
Patient Pseudonym	Pseudonym	A25	Pseudonym	
Bio-bank-ID	Bio-bank-ID	I5		
Informed consent	Type	I2	-1: not yet known 1: donated material, every analysis possible 2: analysis restricted	

					3: new informed consent for each analysis needed	
Bio-material	collected	Date		D8	DDMMYYYY	
	stored	Date		D8	DDMMYYYY	
Material*		Material-ID		I15		
		type		I2	-1: not yet known 1: blood 2: plasma 3: serum 4: tumour tissue 5: normal tissue 6: bone marrow 7. cerebrospinal fluid 8: urine 9: other	
		If other, specify		A50		
storage of material before processing and definite storage		Done		I2	-1: not yet known 1: no 2: yes	
		If yes: Storage before processing		I2	-1: not yet known 1: room temperature 2: refrigerator 3: deep-frozen: - 20°C 4: deep-frozen: - 80°C 5: stabilisation agent added 6: directly processed 7: other	
		If stabilization agent added: please specify		A100		
		If other, specify		A50		
		Shipping/Transport necessary		I2	-1: not yet known 1: no 2: yes	
		If yes: Please specify	Kind of transport	I2	-1: not yet known 1: without cooling 2: with cooling 3: deeply frozen	
			Label for temperature	I2	-1: not yet known 1: no 2: yes	
			Condition of material after transport	I2	-1: not yet known 1: good 2: defrosted 3: deeply frozen during whole transport 4: cold chain interrupted	
			Duration [min]	I5		

Corresponding Histology	Done		I2	-1: not yet known 1: no 2: yes	
	If yes: Please specify	Histology	A50	*including normal tissue*	
		ICD-10	A20		
		Comment	M	Memo	
		Tumour	I2	-1: not yet known 1: inhomogeneous tumour 2: homogeneous tumour 3: only normal tissue	
		Tumour cells	I2	-1: not yet known 1: no 2: vital 3: only necrotic	
Time before processing	Minutes		I5		
Processing	Primary method	general	I2	-1: not yet known 1: nothing 2: cell isolation 3: cell sorting 4: cell isolation and sorting 5: other	
		If other, specify	A50		
		Method specification	M	Memo	
	Secondary method	general	I2	-1: not yet known 1: cell culture 2: controlled frozen for cell culture 3: DNA extraction 4: RNA extraction 5: protein extraction 6: membrane extraction 7: mitochondria extraction 9: other	
		If other, specify	A50		
		Method specification	M	Memo	
DNA	Quality		I2	-1: not yet known 1: 100 % pure 2: 90 % pure 3: < 90 % pure	

	Total amount isolated [µg]		R6.2		
	Number of vials		I2		
	Vial-ID**		A29	Multiple entries possible	PK dna_id
	Storage	Date	D8		
		Temperature			
		Place: Institution ID	I15		
RNA	Quality		I2	-1: not yet known 1: 100 % pure 2: 90 % pure 3: < 90 % pure	
	Total amount isolated [µg]		R6.2		
	Number of vials		I2		
	Vial-ID**		A29	Multiple entries possible	PK rna_id
	Storage	Date	D8		
		Temperature			
		Place: Institution ID	I15		
Protein	Quality		I2	-1: not yet known 1: 100 % pure 2: 90 % pure 3: < 90 % pure	
	Total amount isolated [µg]		R6.2		
	Number of vials		I2		
	Vial-ID**		A29	Multiple entries	
	Storage	Date	D8		
		Temperature			
		Place: Institution ID	I15		
other	Specify material		A10		
	Quality		I2	-1: not yet known 1: 100 % pure 2: 90 % pure 3: < 90 % pure	
	Total amount isolated [µg]		R6.2		
	Number of vials		I2		
	Vial-ID**		A29	Multiple entries	

	Storage	Date	D8		
		Temperature			
		Place: Institution ID	I15		
Paraffin	ID of the institution		I15		
	Histological No.		A25		
	Responsible physician User ID		A20		
	Paraffin Block available		I2	-1: not yet known 1: no 2: yes	
	Number of slides		I3		
	Slide-ID**		A29	Multiple entries	PK-slide_id

* For each material of a patient an extra dataset has to be provided
** Vial ID will be automatically generated: Material-ID + Kind of material + No. of the vial (for example: 123456-DNA-5: meaning this is the 5th vial containing DNA from material with the ID of 123456)

Research that is done by whom using material from the bio-bank

Context	Category	Type	Format/ Coding/Options	TCGA Reference
Bio-bank ID	Institution-ID	I5	Multiple entries possible	
Material ID	Material-ID	I15	Multiple entries possible	
Number of vials needed	Number	I5		
Vial ID	Vial-ID	A25	Multiple entries possible	
Institution doing the research	Institution-ID	I5		
Main Researcher	User-ID	I5		
Project	Name of project	A254		
	Acronym of project	A100		
	Main question	M	Memo	
	Method	M	Memo	
	Research protocol available	I2	-1: not yet known 1: no 2: yes	

	Research approved by EC/IRB	I2	-1: not yet known 1: no 2: yes	
		I5	ID	
	Research financed by	I5	ID	
Request for material	approved	I2	-1: not yet known 1: no 2: yes	
	Date of request	D8	DDMMYYYY	
	Date of approval	D8	DDMMYYYY	
	Contract signed	I2	-1: not yet known 1: no 2: yes	
	Contract signed – Date	D8	DDMMYYYY	
	Date material is shipped	D8	DDMMYYYY	
Result	published	I2	-1: not yet known 1: no 2: yes	
	Link to PubMed	A254	Multiple entries possible	

Requirements for Participation of a Bio-bank / Molecular Biological Laboratory

Context	Category	Type	Format/ Coding/Options	TCGA Reference
Bio-bank ID	Institution-ID	I5		
Bio-bank fulfils legal requirements to participate in ContraCancrum		I2	-1 not yet known 1: no 2: yes	
Bio-bank fulfils ethical requirements to participate in ContraCancrum		I2	-1 not yet known 1: no 2: yes	
Bio-bank has SOPs according to GLP criteria		I2	-1 not yet known 1: no 2: yes	
Bio-bank is a registered member of ContraCancrum		I2	-1 not yet known 1: no 2: yes	
Bio-bank has already participated in an ContraCancrum trial		I2	-1 not yet known 1: no 2: yes	

Responsible Person of the Bio-bank is registered in ContraCancrum		I2	-1 not yet known 1: no 2: yes	
	If, yes, User-ID	I5		
Date when contract with ContraCancrum was signed	Date	D8		
Person who signed contract for ContraCancrum	User-ID	I5		
Person who signed contract for the Bio-bank	User-ID	I5		

This database will be built as a module in ObTiMA. ObTiMA is a Ontology based Trial Management Application that is developed within ACGT[8]. To use ObTiMA the ACGT Master Ontology (MO)[9] will be expanded for glioma and lung cancer during the project. For molecular genetics the Gene Ontology[10] will be used.

Molecular biological and histological data

Context	Category	Type	Format/ Coding/Options	TCGA Reference
Material used	Institution-ID	I5	Multiple entries possible	
	Material-ID	I15	Multiple entries possible	
	Vial-ID / Slide-ID	A25	Multiple entries possible	
Pathology	Section location	R10.2		SECTIONLOCATION
	Number proliferation cells	R10.2		NUMBERPROLIFERATINGCELLS
	Percent tumour cells	R10.2		PERCENTTUMOURCELLS
	Percent normal cells	R10.2		PERCENTNORMALCELLS
	Percent necrosis	R10.2		PERCENTNECROSIS
	Percent stromal cells	R10.2		PERCENTSTROMALCELLS
	Percent lymphocyte infiltration	R10.2		PERCENTLYMPHOCYTEINFILTRATION

[8] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. 15th September 2007

[9] http://bioportal.bioontology.org/ontologies/38787 (last accessed 13th April 2009)

[10] http://www.geneontology.org/ (last accessed 13th April 2009)

Percent monocyte infiltration	R10.2		PERCENTMONOCYTEINFILTRATION
Percent granulocyte infiltration	R10.2		PERCENTGRANULOCYTEINFILTRATION
Percent neutrophile infiltration	R10.2		PERCENTNEUTROPHILINFILTRATION
Percent eosinophile infiltration	R10.2		PERCENTEOSINOPHILINFILTRATION
Endothelial proliferation	A50		ENDOTHELIALPROLIFERATION
Nuclear pleomorphism	A50		NUCLEARPLEOMORPHISM
Palisading necrosis	A50		PALISADINGNECROSIS
Cellularity	A50		CELLULARITY
Percent p53 staining	R10.2		
Percent ki67 staining	R10.2		

References regarding bio-banking

The following references are listed as an update of current source of knowledge regarding bio-banking issues. This list is not entitled to be complete.

Biobanks and Networks in general

1. Asslaber M, Abuja PM, Stark K, Eder J, Gottweis H, Trauner M, Samonigg H, Mischinger HJ, Schippinger W, Berghold A, Denk H, Zatloukal K. The Genome Austria Tissue Bank (GATiB). Pathobiology. 2007; 74(4):251-8.
2. Australasian Biospecimen Network. Biorepository Protocols. 2007. http://www.abrn.net/pdf/ABN_SOPs_Review_Mar07_final.pdf (last access: 13th April 2009)
3. Biobanks: Accelerating molecular medicine. Challenges Facing the Global Biobanking Community. November 2004, IDC #4296. http://www-03.ibm.com/industries/global/files/Biobanks_Accelerating_Molecular_Medicine.pdf (last access: 13th April 2009)
4. "Bio-bank Japan" toward the personalized medicine. Abstract for presentation at 11th International Congress of human genetics. http://www.ichg2006.com/abstract/497.htm.
5. Cooperative Human Tissue Network (CHTN). 2009. http://chtn.nci.nih.gov/ (last access: 13th April 2009)
6. Human Genome Epidemiology Network. 2009. http://www.cdc.gov/genomics/hugenet/default.htm (last access: 13th April 2009)
7. National Bio-specimen Network Blueprint. Andrew Friede, Ruth Grossman, Rachel Hunt, Rose Maria Li, and Susan Stern, eds. (Constella Group, Inc., Durham, NC, September 2003) http://prostatenbnpilot.nci.nih.gov/FINAL_NBN_Blueprint.pdf (last access: 13th April 2009)
8. Næss O, Søgaard AJ, Arnesen E, Beckstrøm AC, Bjertness E, Engeland A, Hjort PF, Holmen J, Magnus P, Njølstad I, Tell GS, Vatten L, Vollset SE, Aamodt G. Cohort Profile: Cohort of Norway (CONOR). Int J Epidemiol. 2008;37:481-485
9. Palmer LJ. UK Bio-bank: bank on it. Lancet. 2007;369:1980-2.
10. Pennsylvania Cancer Alliance. Bioinformatics Consortium. 2009. http://pcabc.upmc.edu/ (last access: 13th April 2009)
11. Quality of Life and Management of Living Resources. European Human Frozen Tumour Tissue Bank. TUBAFROST QLRI-CT-2002-01551 © Medlawconsult Deliverable 7.1, version 1.3 ETHICAL and LEGAL ASPECTS. http://www.tubafrost.org/research/moreinfo/deliverables/TUBAFROST%20Deliverable%207.1.pdf (last access: 13th April 2009)

12. Riegman PH, Dinjens WN, Oomen MH, Spatz A, Ratcliffe C, Knox K, Mager R et al. TuBaFrost 1: Uniting local frozen tumour banks into a European network: an overview. Eur J Cancer. 2006;42:2678-83.
13. Singapore Tissue Network. 2007. http://www.stn.org.sg/. (last access: 13th April 2009)
14. Swede H, Stone CL, Norwood AR. National population-based bio-banks for genetic research. Genet Med. 2007;9:141-9.
15. Swedish National Biobank Program, 2003 http://www.biobanks.se (last access: 13th April 2009)
16. The Shared Pathology Informatics Network (SPIN). 2009. www.cancerdiagnosis.nci.nih.gov/spin/#description (last access: 13th April 2009)
17. UK Bio-bank. http://www.ukbiobank.ac.uk/ (last access: 13th April 2009)

Infrastructure

18. ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. pp 54-61, 15th September 2007
19. Biobanking and Biomolecular Resources Research Infrastructure (BBMRI), 2009. http://bbmri.eu/ (last access: 13th April 2009)
20. Data storage and DNA banking for biomedical research: technical, social and ethical issues. Recommendations of the European Society of Human Genetics. Eur. J. Hum. Genetics 2003; 11(suppl 2): S2-S10.
21. Human Genome Organization. 2008. http://www.hugo-international.org/. (last access: 13th April 2009)
22. Public Population Project in genomics (P3G). 2009. http://www.p3gconsortium.org/ (last access: 13th April 2009)
23. Riegman PH, Dinjens WN, Oosterhuis JW. Bio-banking for interdisciplinary clinical research. Pathobiology. 2007;74:239-244.
24. Specialized Programs of Research Excellence (SPOREs). 2008. http://spores.nci.nih.gov/ (last access: 13th April 2009)

Databases and Guidelines

25. Anneca T, Bompiani A, Buttura R, Calori G. et al. Guidelines for clinical protocols in genetic research. 2006. http://www.fsk.it/eventi/guidelines.pdf. (last access: 13th April 2009)
26. Auray-Blais C, Patenaude J. A bio-bank management model applicable to biomedical research. BMC Med Ethics. 2006;7:E4.
27. GENATLAS, http://genatlas.medecine.univ-paris5.fr/ (last access: 13th April 2009)
28. Gottweis H, Zatloukal K. Bio-bank governance: trends and perspectives. Pathobiology. 2007;74:206-11.
29. Isabelle M., I. Teodorovic, M.M. Morente, D. Jamine, A. Passioukov, S. Lejeune, P. Therasse, W.N.M. Dinjens, J.W. Oosterhuis, K.H. Lam, M.H.A. Oomen, A. Spatz, C. Ratcliffe, K. Knox, R. Mager, D. Kerr, F. Pezzella, M. van de Vijver, H. van Boven, S. Alonso, D. Kerjaschki, J. Pammer, J.A. Lopez-Guerrero, A. Llombart Bosch, A. Carbone, A. Gloghini, E.-B. van Veen, B. van Damme, P.H.J. Riegman, "TuBaFrost 5: Multifunctional central database application for a European tumour bank", Eur J Cancer 2006;42:3103–9
30. ISBER. Best practices for repositories: Collection, Storage, and Retrieval of Biological Materials for Research. 2008.
31. Laurie G. Genetic databases: assessing the benefits and the impact on human and patient rights--a WHO report. Eur J Health Law. 2004;11:87-92. http://www.era.lib.ed.ac.uk/bitstream/1842/2447/1/WHOreportGeneticDatabases.pdf (last access: 13th April 2009)
32. Lickerman, Erik, Bry L., Herring N., Pandelidis Y., Afsheen, Nighat, Sharif A., "Grid-Enabled Bio-Sample Management Application for Data-Intensive Biomarker Analysis", Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)
33. Manolakou K, Vidalis T. Report on the collection and use of genetic data. http://www.bioethics.gr/media/pdf/reports/report_genetic_data_eng.pdf (last access: 13th April 2009)
34. Olund, Lindqvist, and Litton, "BIMS: An information management system for biobanking in the 21st century", IBM SYSTEMS JOURNAL, VOL 46, NO 1, 2007

35. Rai AJ, Gelfand CA, Haywood BC, Warunek DJ, Yi J, Schuchard MD et al. HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. Proteomics. 2005;5:3262-77.
36. Richard Casey, "How Federated Databases Benefit Bioinformatics Research", 2008 http://www.b-eye-network.com/view/2164 (last access: 13th April 2009)
37. Stenson et al (2003), The Human Gene Mutation Database, Hum Mutat (2003) 21:577-581
38. The Cancer Genome ATLAS. 2007. http://cancergenome.nih.gov/ (last access: 13th April 2009)
39. The caBIG project 2009. https://cabig.nci.nih.gov/ (last access: 13th April 2009)
40. UniProt. 2009. http://www.ebi.ac.uk/uniprot/ (last access: 13th April 2009)
41. WHO. IARC Working Group reports. Volume 2. Common Minimum Technical Standards and Protocols for Biological Resource Centres dedicated to Cancer Research / editors, E. Caboux, A. Plymoth, P. Hainaut. ISBN 9789283224426. © International Agency for Research on Cancer, 2007. http://www.iarc.fr/en/content/download/4910/43565/standardsBRC-1.pdf (last access: 13th April 2009)

Legal and ethical issues

42. Annas GJ. HIPAA regulations - a new era of medical-record privacy? N Engl J Med. 2003;348:1486-90.
43. Cambon-Thomsen A, Rial-Sebbag E, Knoppers BM. Trends in ethical and legal frameworks for the use of human bio-banks. Eur Respir J. 2007; 30:373-82.
44. CIOMS International Ethical Guidelines for Biomedical Research involving human subjects. January 2002. http://www.cioms.ch/frame_guidelines_nov_2002.htm (last access: 13th April 2009)
45. Creation and governance of human genetic research databases. http://213.253.134.43/oecd/pdfs/browseit/9306091E.PDF
46. Ethical and policy issues in research involving human participants. Volume I. Report and recommendations of the National Bioethics Advisory Commission. Bethesda, Maryland. August 2001. http://bioethics.georgetown.edu/nbac/human/overvol1.pdf (last access: 13th April 2009)
47. Genomics and world health. Report of the Advisory Committee on Health Research. WHO 2002. ISBN 92 4 154554 2 (NLM Classification: QZ 50) http://whqlibdoc.who.int/hq/2002/a74580.pdf (last access: 13th April 2009)
48. Godard B, Schmidtke J, Cassiman JJ, Aymé S. Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits. A professional perspective. Eur J Hum Genet. 2003;11 Suppl 2:S88-122.
49. International Bioethics Committee. Preliminary draft report on consent by the Working Group of International Bioethics Committee (IBC) on Consent. Paris 20-22 November 2006. http://portal.unesco.org/shs/en/files/10353/11631788121draft_report_consent.pdf/draft_report_consent.pdf (last access: 13th April 2009)
50. Lähteenmaki R. Estonian parliament considers genome law. Nat Biotechnol. 2000; 18:1135.
51. Reymond MA, Steinert R, Eder F, Lippert H. Ethical and regulatory issues arising from proteomic research and technology. Proteomics. 2003;38:1387-96.
52. Simon J, Paslack R, Robienski J, Cooper DN, Goebel JW, Krawczak M. A legal framework for bio-banking: the German experience. Eur J Hum Genet. 2007;15:528-32.
53. Tansey J. and M. Burgess. The foundations, applications and ethical dimensions of biobanks. Electronic Working Papers Series. W. Maurice Young Centre for Applied Ethics, University of British Columbia at www.ethics.ubc.ca. 2004 http://www.ethics.ubc.ca/workingpapers/deg/deg005.pdf (last access: 13th April 2009)
54. The stored tissue issue. Biomedical research, ethics and law in the era of genomic medicine. Weir R and Olick RS. Oxford University Press 2004.
55. Tri-Council Policy Statement. Ethical Conduct for Research Involving Humans. August 1998. http://www.umanitoba.ca/research/media/TCPS_gov_canada_statement.pdf (last access: 13th April 2009)
56. UNESCO. Universal Declaration on Bioethics and Human Rights http://portal.unesco.org/en/ev.php-URL_ID=31058&URL_DO=DO_TOPIC&URL_SECTION=201.html (last access: 13th April 2009)

4. Regulations needed for the simulation models from a clinical perspective

The goal of cancer research is to find better ways to treat patients with cancer. During the last decades basic research and clinical trials have gathered a lot of new insights in the molecular biology of cancer providing new drugs and treatment approaches. Nevertheless the outcome for most cancers is still dismal demanding new and better treatments for patients. With the help of *In Silico Oncology* it is expected that cancer growth and response to different treatments can be simulated. Such *in silico* experiments might help clinicians in future to find the best way of treating an individual patient by simulating different treatments in the computer before starting the treatment in reality.

The 'In Silico Oncology Group' (ISOG), National Technical University of Athens, has adopted an essentially "top-down" modeling approach and developed a number of hybrid discrete Monte Carlo / cellular automata and continuous differential equation simulation models of tumour growth and response to therapeutic modalities by fully exploiting the insight gained by molecular biology and other disciplines as well as the individual patient's data. The aim is a better understanding of biological mechanisms concerning cancer and related therapeutic interventions and, in the long term, a contribution to the design of patient individualized therapies.

From a clinical point of view **two preconditions** are of utmost importance, if one can trust predictions of *in silico* methods:

1. every *in silico* method has to be part of a clinico-genomic trial
2. every prediction of an *in silico* method has to be compared with the reality

In the process of developing *in silico* methods it is necessary to define the needed data in a first step, including 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). To make the simulation predictions as precise and realistic as possible it is crucial to get as much information from each of the different categories. The amount of data will be restricted by the availability of tumour material, imaging data and clinical data. Therefore *In Silico Oncology* has always to be integrated into or part of a clinico-genomic trial, where data management including data security and anonymisation or pseudonymisation of data as well as tumour banking is well established. In addition the trial is reviewed by an ethical committee and fulfils all other GCP criteria to get approval by regulatory authorities.

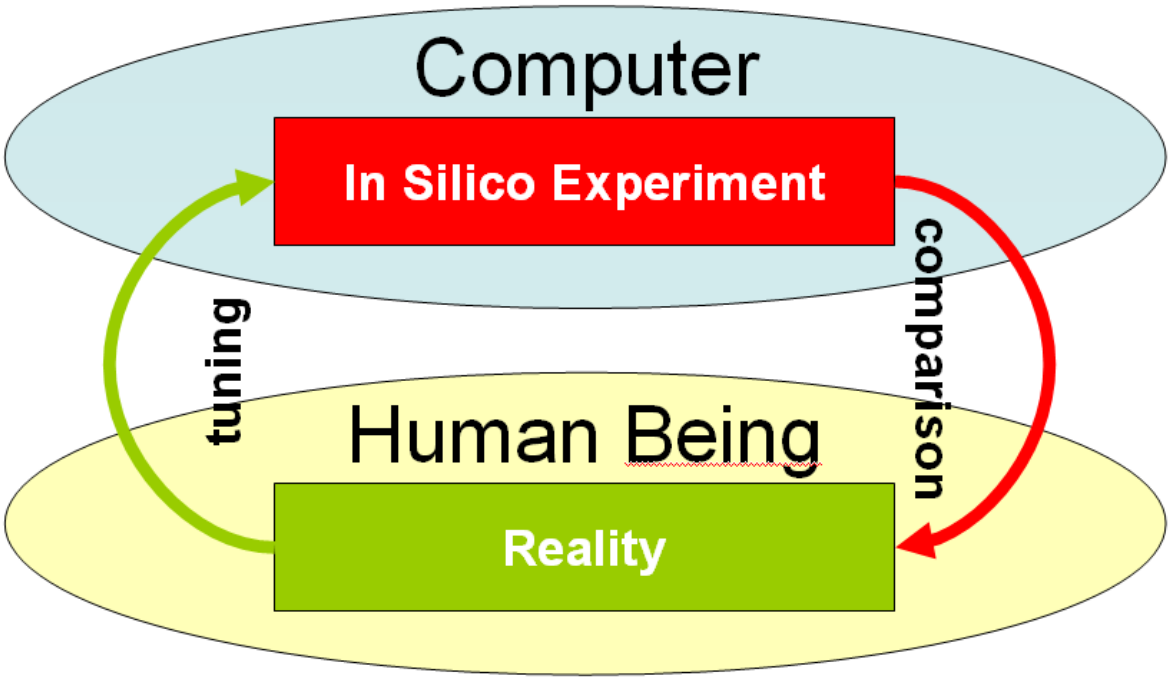

Fig. 4.1: Comparison between 'In Silico' experiments and the reality.

To make the simulation predictions as precise and realistic as possible, it is crucial to acquire as much information from each of the different categories as practicable. However, the amount of data will be restricted by the availability of tumour material, imaging data and clinical data. For that reason standards have to be defined regarding the clinical data that are needed, the imaging studies that are done, the segmentation process for rendering the tumour, and the molecular genetic data that are analyzed.

As written in D2.1 for the use the ContraCancrum system in clinics, it is necessary that every clinical scenario in ContraCancrum fulfils the following criteria:

- The question to be answered by the scenario is of clinical relevance.
- The data that are needed for the scenario can be easily provided by clinicians.
- All legal, ethical and data security requirements are fulfilled.
- The tools provided with the scenario are easy to use by clinicians.
- All scenarios must be validated in clinical trials before use in routine practice.

Figure 4.2 shows the use of ContraCancrum in the clinical setting.

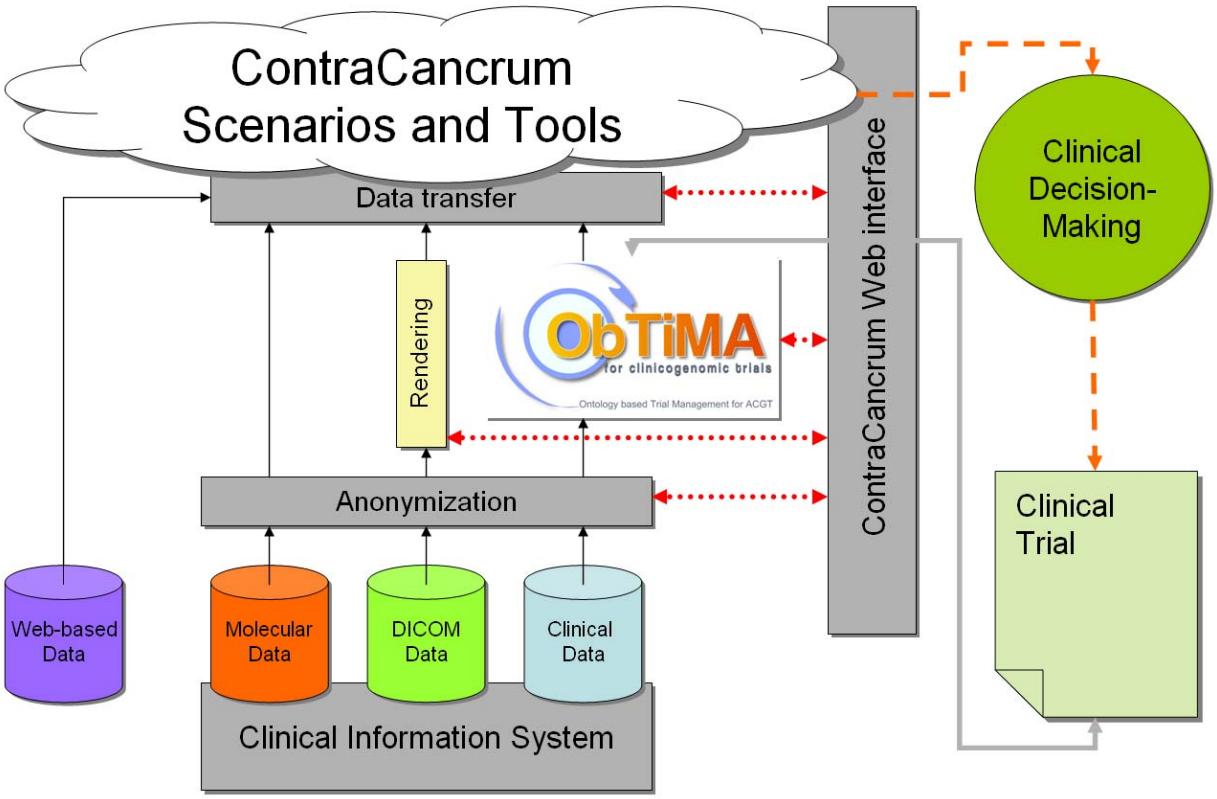

Figure 4.2: The use of ContraCancrum in the clinical setting

Clinical data

Clinical data that will be used in ContraCancrum are defined in deliverable D4.1 chapter 5.1 and in chapter 8 of this deliverable. They have to be prospectively collected. The data management system that will be used is ObTiMA, developed within ACGT. ObTiMA can be used with and without an ontology. To use the Master Ontology (MO) of ACGT it has to be updated for glioma and lung cancer during the first year of the project.

The following categories of clinical data are needed and defined by the scenarios/trials of ContraCancrum:

- History, symptoms
- Diagnostics
- Diagnosis

- Treatment
- SAE and SUSAR
- Outcome

All items of each of the categories are defined in chapter 8.

Imaging data

Imaging studies have to be done at the time of diagnosis and the end of the treatment at least. It is desirable to have more studies done during the treatment phase. An absolute need is before and after every treatment element (surgery, irradiation and chemotherapy).

Since for **glioma** MRI imaging is not always performed at the University Hospital of the Saarland, the following requirements are defined for ContraCancrum (in blue are minimal requirements):

- T2 axial — 3-4 mm slice thickness
- T1 axial — 3-4 mm slice thickness before contrast enhancement
- MPRAGE — (*Magnetization Prepared Rapid Gradient Echo*) 3D, 1 mm slice thickness after contrast enhancement reformation in all planes possible
- FLAIR — (*Fluid Attenuation Inversion Recovery*) 4 mm slice thickness after contrast enhancement
- DWI (ADC) — (*Diffusion weighted Imaging, (Apparent Diffusion Coefficient*)) 4 mm slice thickness, axial
- DTI — (*Diffusion Tensor Imaging*)
- MR-Spectroscopy — (Chemical Shift Imaging)
- SWI — (Susceptibility-weighted Imaging)

For **lung cancer** the following minimal requirements are required for ContraCancrum:

- T2 — Single shot T2 weighted
- CT — without and with contrast enhancement
- PET-CT

Data will further be characterized in D9.2.

The standard examination should consist of a T2-weighted SE dual echo sequence preferably in the axial plane. The short echo T2-sequence may be substituted by a FLAIR-sequence. The slice thickness should not exceed 7 mm and the slice factor should not exceed 20%.

A T1-weighted sequence, preferably in the axial plane, should be obtained followed by the same scan sequence after intravenous contrast administration. Additional T1-weighted post-contrast sequences in the coronal and sagittal plane are very helpful. In small or irregular tumours slice thickness should be correspondingly small.

Conventional spin echo-techniques are preferred to all kinds of gradient echo sequences, because flow-related enhancement of cerebral vessels by gradient echo- sequences may cause problems in differentiation from meningeal enhancement and the extent and degree of enhancement may be of a lesser order than conventional T1-weighted imaging.

The administration of Gadolinium should follow the general rule of a slow intravenous injection of 0.1mmol/kg bodyweight Gadolinium. The post-contrast scan should not be started until after the full injection of the contrast medium.

Due to the availability of different Gadolinium containing contrast-media it should be observed to always apply equivalent amounts of Gadolinium.

Generally, follow-up scanning should be comparable with prior examinations as it can be very hard to make direct comparisons between studies using different imaging planes and machines.

Segmentation

The segmentation and the corresponding tumour volume are needed for the simulation model in gliomas and lung cancer. It is of utmost importance that the segmentation is done as precisely as possible. An accurate segmentation depends on the person doing the segmentation and on the software used. To have reproducible segmentation data it is necessary to use a tool that will do the segmentation semi-automatically and reproducible so that the inter-personal variance can be minimized as much as possible. For that purpose a tool is under development by FORTH fulfilling these criteria. This tool will help to standardize the segmentation process and will make it reproducible and user independent.

The second problem of segmentation is the fact, that no one can really define the edge between tumour and normal tissue. What seems to be the border of a tumour in imaging studies is not proven by histology. Even different weighted MRI scans of the same tumour at the same time will show different shapes of the same tumour (fig. 4.3). As a result the imaging has to be standardized, especially if follow-up MRIs of a patient will be compared and used for the validation of the simulation model.

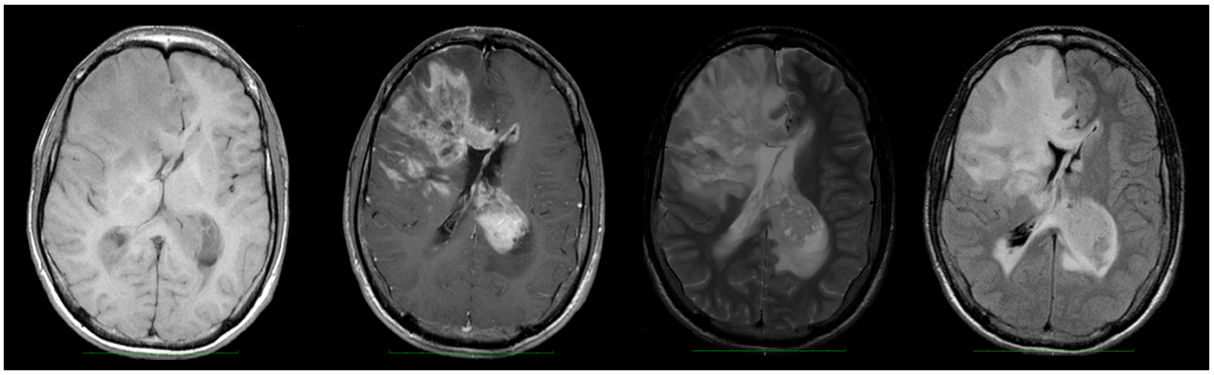

Fig 4.3. Glioblastoma at the time of diagnosis. From left to right: T1, T1 + contrast enhancement (gadolinium), T2, T2 flair weighted images.

For glioma there is no possibility to compare macroscopic histology with the imaging scans. This will be the case for lung cancer in few patients and will be done in ContraCancrum.

Most of the tumours are heterogeneous, meaning that it is also necessary to render necrotic areas in a tumour. This is a task that is not easy to reproduce even with the help of semiautomatic tools. For that purpose the following model will be used:

After segmentation a histogram of signal intensities within the tumour can be done for each slice. It is possible to define different areas of the histogram that correspond to different areas in the tumour. With the help of upper and lower threshold values by using the segmentation tool one is able to narrow the signals in the histogram to the different heterogeneous areas in the tumour. This will allow calculating the area of each of the different heterogeneous areas for each slice. Using the same threshold values for each slice will allow calculating also the volume of the different heterogeneous parts within the tumour. Such a tool allows also displaying the shape of the different heterogeneous areas within the tumour automatically.

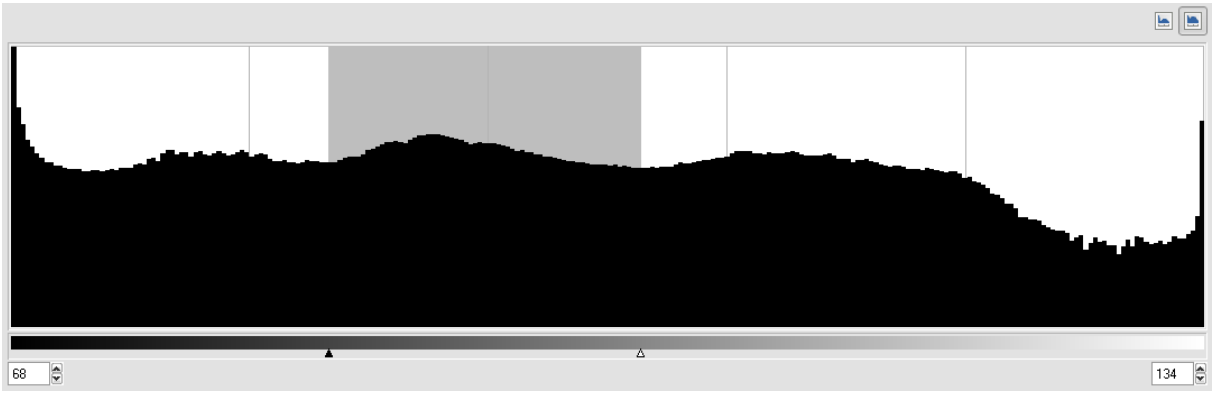

Fig. 4.4. Histogram of signal intensity in a tumour.

If this procedure is done in different tumours, where the imaging parameters are the same the so rendered heterogeneous areas of the tumour can be correlated to large-section histology imaging if possible. It might be possible that specific groups of signal intensities are correlated to a specific histology. This will be tested during ContraCancrum. Such a correlation might help in the future to correlate MRI imaging with histology.

Molecular Genetic data

Molecular genetic data will be generated analyzing tumour tissue and serum of the patient. The following data will be used for the 'in silico' models in ContraCancrum:

Autoantibodies against tumour specific antigens	glioma and lung cancer
Sequencing of EGF receptor	lung cancer
cDNA expression	glioma The preference would be for the Affymetrix Platform HGU133plus2. This will allow consistancy with both TCGA data and REMBRANDT data sources.

These data will be stored in a repository for molecular data. For more details regarding molecular data see chapter 3, 6 and 7.

Data security

All regulations regarding data security are fulfilled within ContraCancrum. Only pseudonymized or anonymized data will be used and stored. The process of pseudonymisation is done at the University Hospital in Homburg, where the clinical, imaging and molecular data are created. There will be no link available between personal data and data used in ContraCancrum. The process of anonymization or pseudonymisation is given in figure 4.5.

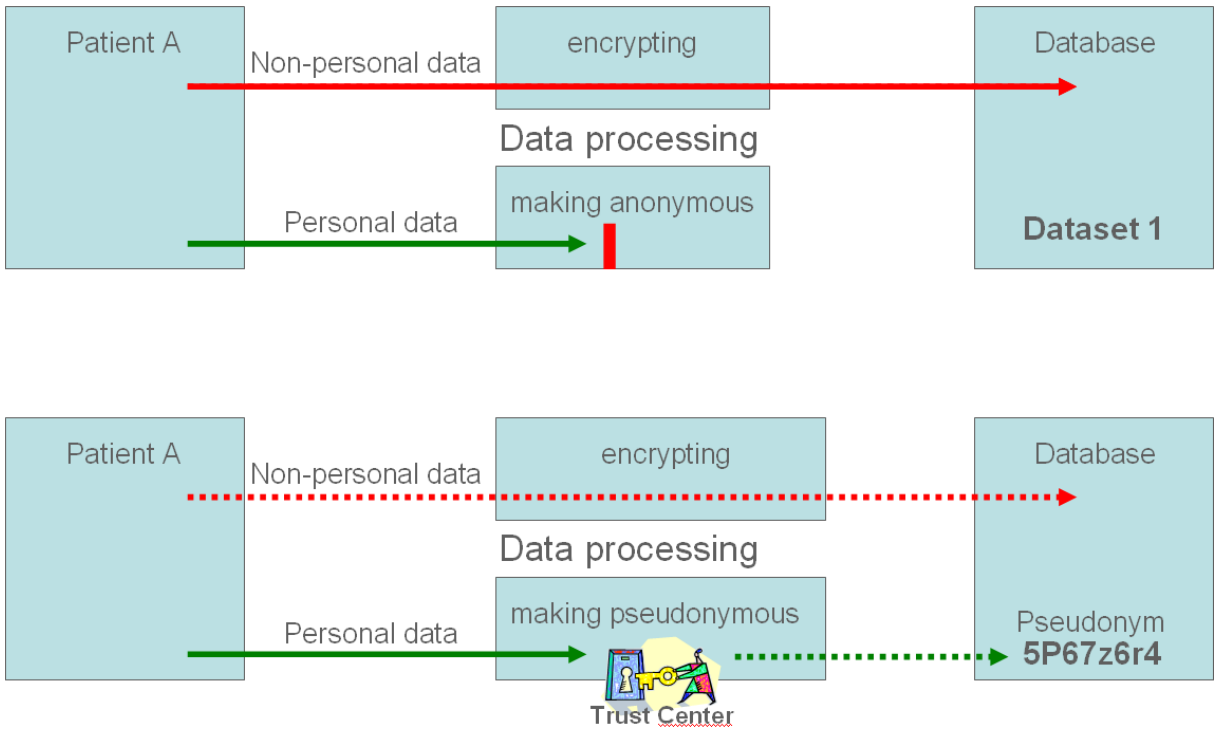

Fig. 4.5. Process of anonymization or pseudonymisation

ContraCancrum Clinical Studies and 'In Silico' Modelling Scenarios

5. Variants of Medulloblastoma as examples of distinct entities with typical clinical, imaging and molecular data

This chapter is mainly based on the excellent paper of Gilbertson et al.[11].and McManamy et al.[12]

Medulloblastoma is the most common malignant brain tumour in children. The World Health Organisation (WHO) classification of 2007 of nervous system tumours recognizes anaplastic, large cell, desmoblastic/nodular variants and Medulloblastoma with extensive nodularity (MBEN). The incidence of desmoplastic medulloblastoma, which is characterized by a nodular architecture and a network of internodular collagen fibers, varies substantially between series (5%–25%), but it appears to be greater in infants and adult patients than in children. The nodular/desmoplastic medulloblastoma appears to have clinical, imaging, genetic and biological characteristics that set it apart from other variants of this tumour[1] (fig. 5.1).

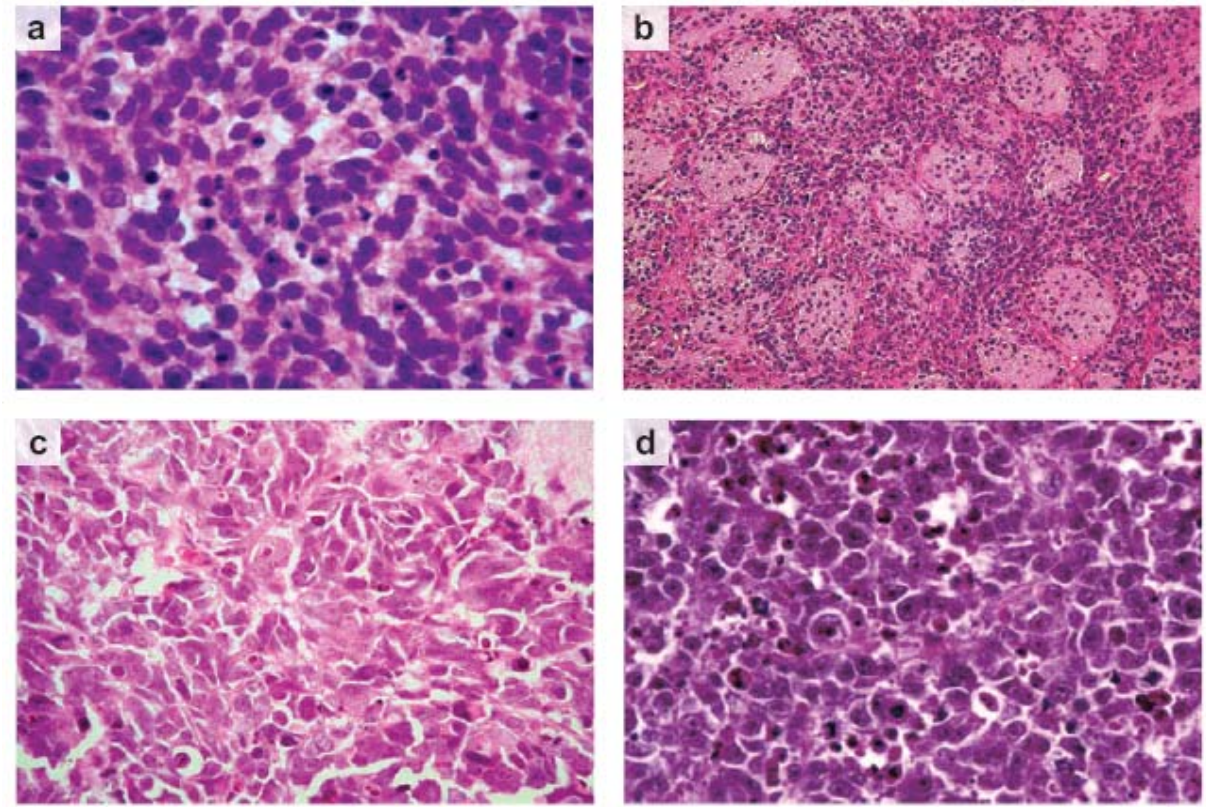

Fig 5.1: Histologic variants of medulloblastoma. (*a*) The classic medulloblastoma is composed of sheets of small uniform cells with a high nuclear-to-cytoplasmic ratio. (*b*) The nodular/desmoplastic medulloblastoma combines nodules of differentiated neurocytic cells with a low growth fraction and desmoplastic internodular zones of moderately pleomorphic cells with a high growth fraction. (*c*) The anaplastic medulloblastoma contains pleomorphic cells with polyhedral forms and a high growth fraction. Abundant apoptosis and examples of cell wrapping are evident. (*d*) The large-cell medulloblastoma contains groups of large uniform cells with vesicular nuclei and a single nucleolus. Anaplasia characterizes other regions of this variant. From Gilbertson et al.[11].

[11] Richard J. Gilbertson, David W. Ellison: The Origins of Medulloblastoma Subtypes. Annu. Rev. Pathol. Mech. Dis. 3:341–365, 2008

[12] Charles S. McManamy; Jane Pears; Claire L. Weston; Zoltan Hanzely; James W. Ironside; Roger E. Taylor; Richard G. Grundy; Steven C. Clifford; David W. Ellison; on behalf of the Clinical Brain Tumour Group, Children's Cancer and Leukaemia Group, UK: Nodule Formation and Desmoplasia in Medulloblastomas—Defining the Nodular/Desmoplastic Variant and Its Biological. Brain Pathology 17:151-164, 2007

Nevoid basal cell carcinoma (NBCCS) was first delineated in 1960 and belongs to the so-called phakomatoses[13]. The major clinical manifestations are multiple basal cell carcinomas of the skin, dyskeratotic palmar, and plantar pits, odontogenic keratocysts of the jaw, and fused or markedly splayed ribs[14]. Patients with NBCCS are at risk for medulloblastomas (MB), mostly with desmoplastic differentiation[15],[16]. The incidence of medulloblastomas in NBCCS was found to be between 3–5%. NBCCS is an autosomal dominant disorder linked to 9q22.3-q31 caused by inactivation mutations in the PTCH gene[17], a human homologue of the Drosophila segment polarity gene patched[18]. NBCCS-associated MB are reported to occur in young infants, have a better prognosis and to be predominantly of the desmoplastic variant or belong to the group of medulloblastoma with extensive nodularity[19].

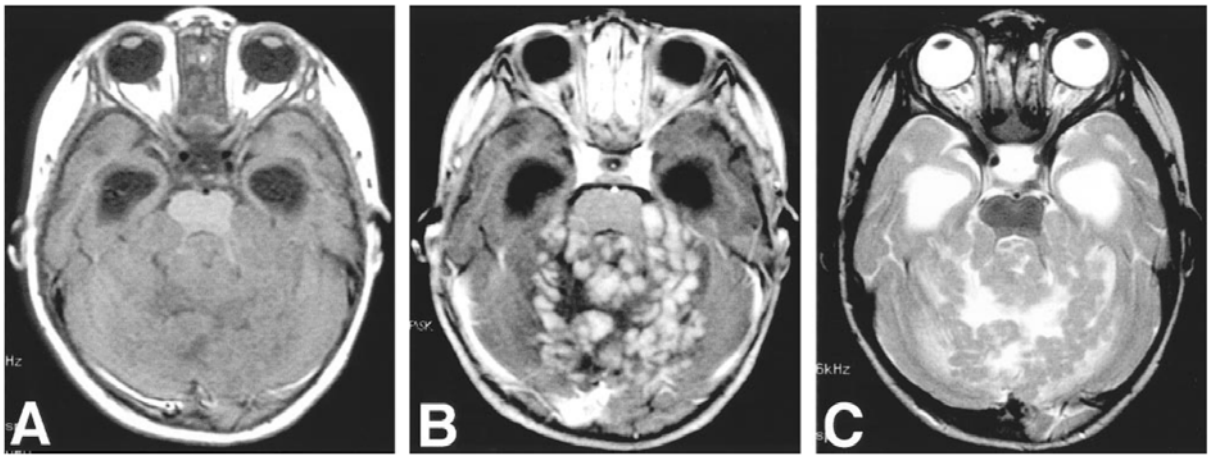

Fig 5.2: Axial view MR images show an equivocal nodularity with surrounding high T2-weighted signal intensity in the tumour. *A*, T1-weighted image (380/10/2 [TR/TE/NEX]). *B*, Contrast-enhanced T1-weighted image (380/10/2) reveals a grape-like enhancing tumour occupying the upper part of the posterior fossa. *C*, T2-weighted image (4500/100/2). From Yuichiro et al., 2002[10]

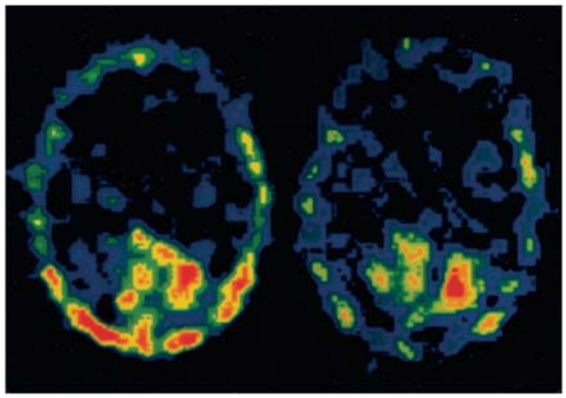

Fig. 5.3: Axial view iodine-123 metaiodobenzylguanidine SPECT scans, obtained 30 minutes (*left*) and 6 hours (*right*) after an IV injection of iodine-123 metaiodobenzylguanidine at a dose of 111 MBq, shows extremely high uptake in the tumour. From Yuichiro et al., 2002[10]

[13] Gorlin RJ, Goltz RW.Multiple basal-cell epithelioma, jaw cysts and bifid rib: A syndrome. N Engl J Med 262:908–912, 1960

[14] Kimonis VE, Goldstein AM, Pastakia B, et al. Clinical manifestations in 105 persons with nevoid basal cell carcinoma syndrome. Am J Med Genet 69:299–308, 1997

[15] Evans DG, Farndon PA, Burnell LD, et al. The incidence of Gorlin syndrome in 173 consecutive cases of medulloblastoma. Br J Cancer 64:959–961, 1991

[16] Schofield D, West DC, Anthony DC, et al. Correlation of loss of heterozygosity at chromosome 9q with histological subtype in medulloblastomas. Am J Pathol 146:472–480, 1995

[17] Johnson RL, Rothman AL, Xie J, et al. Human homolog of patched, a candidate gene for the basal cell nevus syndrome. Science 272:1668–1671, 1996

[18] Giangaspero F, Perilongo G, Fondelli MP, et al. Medulloblastoma with extensive nodularity: A variant with favorable prognosis. J Neurosurg 91:971–977, 1999

[19] Aliani S, Brunner J, Graf N, Altmeyer K, Niedermayer I, Strowitzki M: Medulloblastoma with extensive nodularity in nevoid cell carcinoma syndrome. Med Pediatr Oncol 40:266-272, 2003

The typical nodular imaging of the MBEN is given in the following figures (5.2, 5.3). These figures are from Yuichiro et al., 2002[20].

The potential cellular and molecular origins of medulloblastoma subgroups is given in figure 5.4. Fore more details also regarding the development of the normal cerebellum and the pathways that are involved see Gilbertson et al. 2008[21].

Conclusions:
Medulloblastomas with extensive nodularity represent a variant that is characterized by:

1. occurrence in very young children
2. a peculiar grapelike appearance on neuroimaging
3. a common molecular genetic pathway
4. an apparently favourable outcome.

"Childhood tumours containing cells that are morphologically and functionally similar to normal progenitor cells provide fertile ground for investigating the links between development and cancer. In this respect, integrated studies of normal cerebellar development and the medulloblastoma, a malignant embryonal tumour of the cerebellum, have proven especially fruitful. Emerging evidence indicates that the different precursor cell populations that form the cerebellum and the cell signaling pathways that regulate its development likely represent distinct compartments from which the various subtypes of medulloblastoma arise. Definitive characterization of each medulloblastoma subtype will undoubtedly improve treatment of this disease and provide important insights to the origins of cancer."[1]

[20] Yuichiro Naitoh, Toshio Sasajima, Hiroyuki Kinouchi, Shigeki Mikawa, and Kazuo Mizoi: Medulloblastoma with Extensive Nodularity: Single Photon Emission CT Study with Iodine-123 Metaiodobenzylguanidine. *Am J Neuroradiol* 23:1564–1567, 2002

[21] Richard J. Gilbertson, David W. Ellison: The Origins of Medulloblastoma Subtypes. Annu. Rev. Pathol. Mech. Dis. 3:341–365, 2008

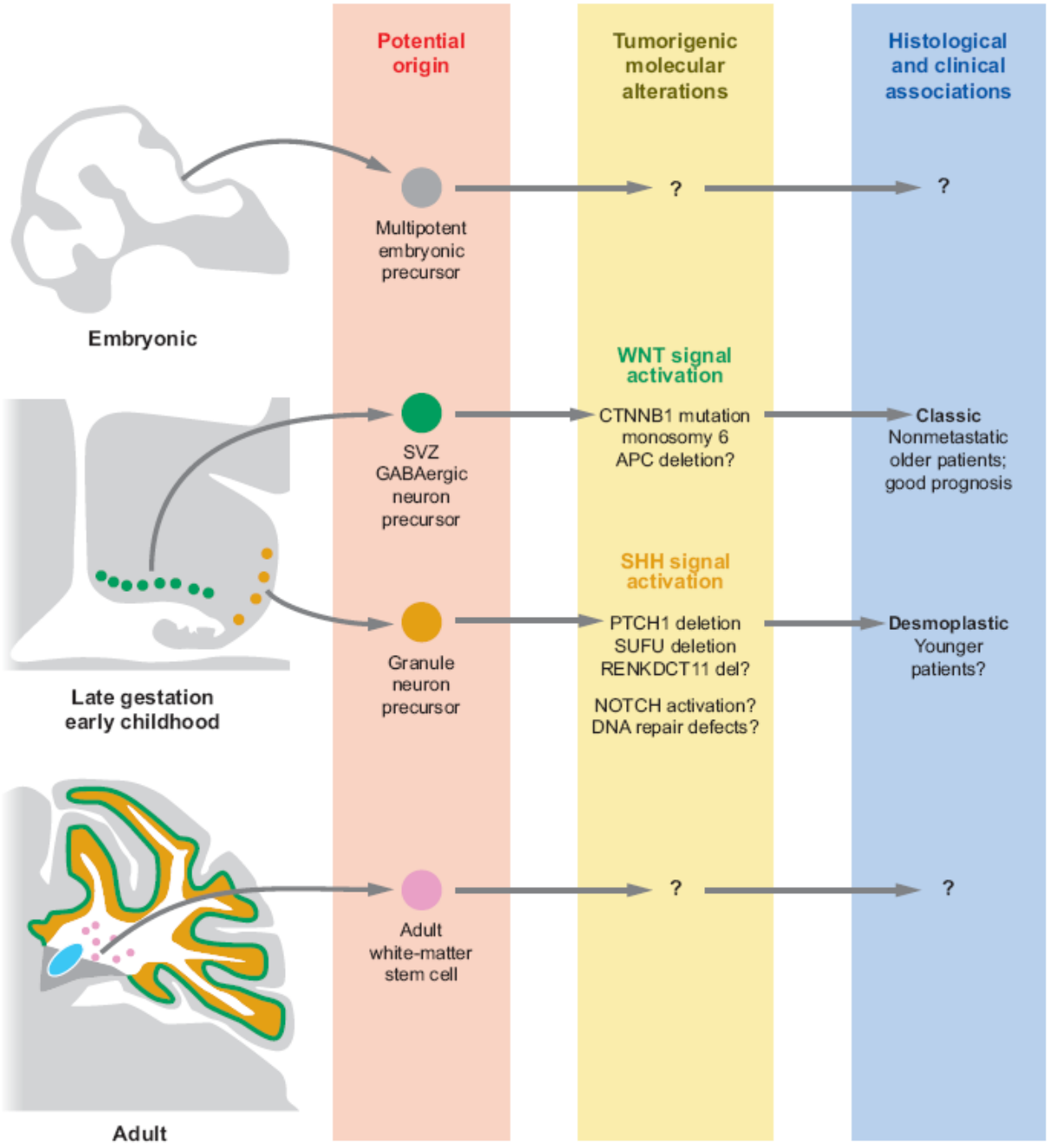

Fig 5.4: Cancer stem cells of medulloblastoma may arise from multipotent precursor cells of the developing embryo, although the types of mutations to which these cells may be susceptible and the forms of the disease to which these tumours may give rise remain unclear. Medulloblastomas that develop with activating mutations in the WNT and SHH pathways are mutually exclusive and are predominantly classic, good prognosis and desmoplastic, mixed prognosis tumours, respectively. White-matter stem cells may be cells of origin of some adult medulloblastomas, but this remains to be shown definitively. From Gilbertson et al., 2008[1].

6. Glioma

Clinical background of gliomas

The following introduction (6.1 and 6.2) is based on:

1. Louis D, Ohgaki H, Wiestler OD, Cavenee WK (eds) (2007) WHO Classification of Tumours of the Central Nervous System. IARC Press, Lyon
2. CBTRUS (2008). Statistical Report: Primary Brain Tumours in the United States, 2000–2004. Published by the Central Brain Tumour Registry of the United States
3. Ohgaki H, Kleihues P (2005) Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. J Neuropathol Exp Neurol 64:479-489.
4. Ohgaki H, Kleihues P (2005) Epidemiology and etiology of gliomas. Acta Neuropathol 109: 93–108.
5. Louis D. (2006) Molecular Pathology of Malignant Gliomas. Annu. Rev. Pathol. Mech. Dis.. 1:97–117
6. Claes A, Idema AJ, Wesseling P (2007) Diffuse glioma growth: a guerilla war. Acta Neuropathol 114:443–458

Gliomas are tumours that arise from glial cells, and include astrocytomas, glioblastomas, oligodendrogliomas, ependymomas, mixed gliomas, malignant gliomas NOS, and neuroepithelial tumours. The broad category glioma represents 36% of all primary brain and CNS tumours. Sixty–one percent of gliomas occur in the frontal, temporal, parietal, and occipital lobes of the brain. (CBTRUS, Figure 6.1).

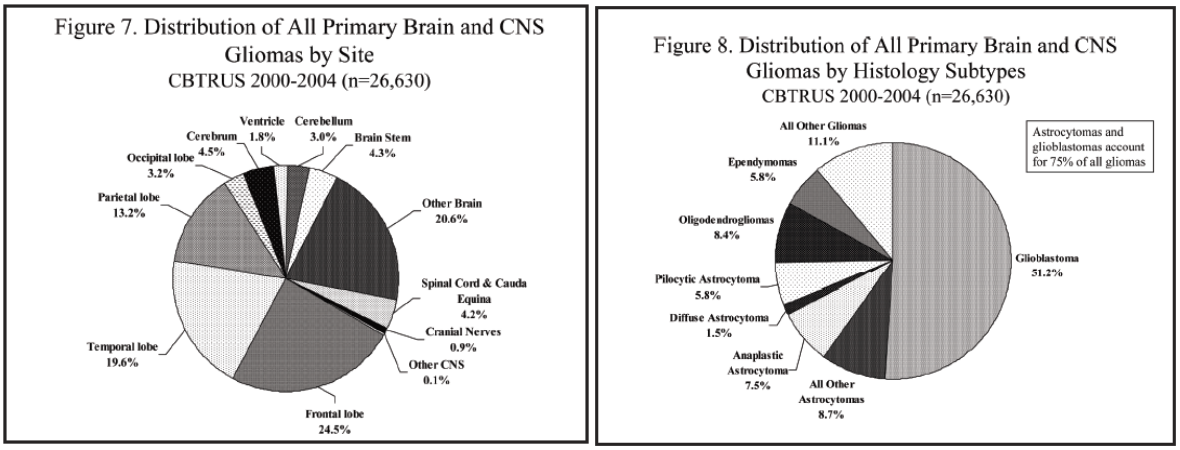

Figure 6.1: Distribution of gliomas by location and histological subtypes (from CBTRUS (2))

The WHO classifies gliomas into four malignancy grades and two main histological subtypes, astrocytomas and oligodendrogliomas. (Louis 2007 (1)):

WHO grade I gliomas or pilocytic astrocytomas (PA) are relatively circumscribed, slowly growing tumours. They represent 5-6% of all gliomas. PA is the most common glioma in children and mostly located in cerebellum (67%). As a rule, PA are macroscopically somewhat discrete. Thus, when anatomy permits, eg cerebellum or cerebral hemispheres, many can be removed in toto. As a group, PA are remarkable in maintaining their WHO grade I status over years and even decades. As a rule, alterations over time are in the direction of regressive change rather than anaplasia. Long patient survival is the rule. Although the lesion may eventually prove fatal, there are a few long term studies for the ultimate outcome of the patients.

The diffuse astrocytoma of WHO grade II (DA) represents a diffusely infiltrating astrocytoma that typically affects young adults. DAs show a high degree of cellular differentiation and slow growth and possess an intrinsic tendency for malignant progression to anaplastic astrocytoma (AA) and ultimately glioblastoma (GBM). DA account for 10-15% of all astrocytic brain tumours. The peak incidence in young adults is between 30-40yr. (10% <20yr, 60% between 20-45, 30% > 45yr, mean 34yr). The total length of disease is mainly influenced by the dynamics of malignant progression to GBM, which tends to occur after a mean interval of 4-5 yrs. There is currently no validated factor that unambiguously predicts in individual paitents whether and how soon malignant progression to anaplastic astrocytoma and GBM is likely to occur. Young age at diagnosis has been consistenly predictive of a more favourable clinical course of patients with low-grade astrocytoma, while large tumour size appears to be a negative predictor. Gross total resection is significantly associated with longer survival.

Anaplastic astrocytoma of WHO grade III (AA) is a diffusely infiltrating malignant astrocytoma in adults, that is histologically characterized by nuclear atypia, increased cellularity, and significant proliferative activity. AAs may arise from WHO II astrocytomas or *de novo* and have an intrinsic tendency to progress to GBM. Mean age of manifestation lies between 45 and 51 yrs. AAs represent the intermediate stage on the route of progression to GBM. They show a strong tendency to progress to GBM. The pace of progression is variable, but population-based studies suggest a mean time interval of approx. 2 yrs. Similar to WHO II astrocytoma and GBM, increasing age is a negative prognostic factor.

The WHO grade IV astrocytoma or glioblastoma multiforme (GBM) is the most frequent primary brain tumour in adults and the most malignant neoplasm in the brain with predominant astrocytic differentiation. Histopathological features include nuclear atypia, cellular pleomorphism, mitotic activity, vascular thrombosis, microvascular proliferation and necrosis. Most GBM (approx. 95%) manifest rapidly de novo. GBMs cannot be completely resected due to their invasive nature. Despite progress in radio/chemotherapy, less than half of the patients survive more than one year, with older age being the most significant adverse prognostic factor. GBMs are the most frequent brain tumour, accounting for approx. 12.15% of all intracranial neoplasm and 60-75% of astrocytic tumours. GBMs can manifest at any age, but preferentially affect adults, with a peak incidence at between 45 and 75 yrs (mean: 61 yrs). The clinical history of the disease is usually short (less than 3 months in more than 50% of cases), unless the neoplasm has developed from a lower grade astrocytoma (secondary GBM). The majority of GBMs (>90%) develop very rapidly with a short clinical history (<3 months), without clinical or histopathological evidence of a preexisting, less malignant precursor lesion (primary GBM). Primary GBM typically affects older patients (mean 62yrs), while secondary GBM manifest younger. Time to progression from WHO II to WHO IV varies considerably, with time intervals ranging from less than 1 year to more than 10 years with a mean interval of 4-5 years. Survival of patients with secondary GBM is significantly longer (median survival time 7.8 months), than for those with primary GBM (4.7 months), but this is likely to be the reflection of younger age of secondary GBM patients.

Despite progress in surgery, radio- and chemotherapy of brain tumours, the overall survival of patients with GBM remains extremely poor. Virtually all therapeutic trials have shown that younger patients (<50yrs) have a significant better prognosis. The presence and extent of necrosis are associated with shorter survival. GBMs are highly resistant to therapy, with only marginal survival increases in a small fraction of patients, even after aggressive surgical resection, external beam radiation therapy (both conformal and whole brain) and maximal tolerated doses for chemotherapy with agents such as temozolomide or nitrosourea. Therapeutic resistance is due to:

1. poor drug delivery because of partial blood-brain-barrier preservation and high tumour interstitial pressure
2. genome instability produced by point mutations, LOH, chromosome deletion and rearrangements, gene amplifications, and epigenetic gene silencing which leads to broad genotypic and phenotypic heterogeneity resulting in clonal populations of cells resistant to any single therapeutic modality
3. invasive properties of GBM cells enabling malignant cells to cross the corpus callosum, spread even to the brain stem and spinal cord, and reside behind a completely intact blood-brain-barrier
4. the presence of a population of neural stem cell-like cells that may harbour resistance mechanisms that are distinct from those of the majority of non-stem-like tumour cells and that may contribute to cellular heterogeneity and
5. retention of abundant DNA repair machinery that abrogates effectiveness of chemotherapy and radiotherapy.

The WHO further defines oligodendrogliomas (either WHO II or WHO III), as well as mixed gliomas or oligoastrocytomas, that show oligodendroglial and astrocytic components (either WHO II or WHO III). Glioma patients whose tumours show oligodendroglial component have a longer median survival rate.

Incidence rates for histologic subtypes of CNS tumours are not readily available, since most cancer registries give combined data for all brain tumours. More detailed information is collected in certain regions; in the United States, registration of all brain neoplasms, including benign lesions, in the SEER database has recently become mandatory (CBTRUS, McCarty et al. 2000). Incident cases included in this population-based study are summarized in Table 1. With 3.55 new cases per 100,000 persons per year, adjusted to the European Standard Population, the glioblastoma is the most frequent histologic type, and accounts for 69% of all incident cases of astrocytic and oligodendroglial tumours. The incidence rate of glioblastomas in the USA, adjusted to the US Standard Population, is 29.6 new cases per million populations per year (CBTRUS). These incident cases do not include secondary glioblastomas that progressed from low-grade or anaplastic gliomas, since only the first diagnosis is considered as an incident case. Incidence rates of all astrocytic and oligodendrogliomas combined, adjusted to the European and US Standard Populations were 5.27 and 5.17 per 100,000 persons per year, respectively (Figure 6.2) (Ohgaki et al (3)).

Table 1 Population-based data of incidence rates, age and sex, and survival of patients with gliomas

Tumor	WHO grade	Region	Incidence rates[a]	M:F ratio	Mean age at diagnosis	Survival						References
						Median (months)	Mean (months)	1 year	2 years	5 years	10 years	
Pilocytic astrocytoma	I	USA	0.23	1.09	17			95%	93%	89%	86%	http://www.cbtrus.org
		Zurich	0.39	1.12	20		142	100%	100%	100%	96%	Burkhard et al. [14]
Diffuse astrocytoma	II	USA	0.13	1.46	47			73%	60%	45%	34%	http://www.cbtrus.org
		Zurich	0.26	1.7	41	67	77	92%	88%	58%	26%	Okamoto et al. [102]
Anaplastic astrocytoma	III	USA	0.49	1.20	50			60%	43%	28%	19%	http://www.cbtrus.org
		Zurich	0.25	1.19	44	20	30	65%	43%	11%	7%	Unpublished data
Glioblastoma	IV	USA	2.96	1.26	62			28%	8.2%	2.9%	1.7%	http://www.cbtrus.org
		Zurich	3.39	1.28	61	4.9	7.3	18%	3.3%	1.2%	0.2%	Ohgaki et al. [100]
Oligodendroglioma	II	USA	0.34		42			88%	80%	66%	47%	http://www.cbtrus.org
		Zurich	0.27	0.92	40	139	106	98%	96%	78%	51%	Okamoto et al. [102]
Anaplastic oligodendroglioma	III	USA	0.10	1.15	46			75%	57%	38%	25%	http://www.cbtrus.org
		Zurich	0.11	2.33	49	16	37	50%	45%	30%	7.5%	Unpublished data
Oligoastrocytoma	II	Zurich	0.10	1.0	40	79	85	95%	90%	70%	49%	Okamoto et al. [102]
Anaplastic oligoastrocytoma	III	Zurich	0.08	0.77	46	18	30	62.5%	43.8%	12.5%	0%	Unpublished data
Mixed glioma[b]	II / III	USA	0.12	1.21	40			84%	72%	54%	39%	http://www.cbtrus.org
Ependymoma / Anaplastic ependymoma	II / III	USA	0.23	1.29	35			86%	79%	66%	55%	http://www.cbtrus.org

[a]Incidence rates (per 100,000 person per year) in 1992–1997 in United States adjusted to 2000 US population (data for USA are from CBTRUS data [20]); Incidence rates in Zurich, Switzerland in 1980–1994 adjusted to 2000 US population
[b]Oligoastrocytoma (WHO grade II) and anaplastic oligoastrocytomas (grade III) are combined

Figure 6.2: Incidence rates of glioma subtypes (from Ohgaki et al. (4))

Except for patients with pilocytic astrocytoma, who had excellent outcome irrespective of radiotherapy (Ohgaki et al. 2003), survival of patients with an astrocytic tumour is still very poor. The median survival time (MST) of patients with a low-grade astrocytoma was 5.6 years, with anaplastic astrocytoma 1.6 years and with glioblastoma only 4.9 months (Figure 6.3). The observed survival rates of glioblastoma patients at the population level were 42.4% at 6 months, 17.7% at one year, 3.3% at 2 years, and 1.2% at 3 years (Ohgaki et al., 2004). This is consistent with population-based data from Canada, which showed that only 15 out of 689 glioblastoma patients (2.2%) diagnosed during 1975–1991 survived 3 years or longer (Scott et al., 1999). The presence of oligodendroglial components in low grade gliomas is associated with longer survival: oligodendroglioma patients (MST = 11.6 years; 51% at 10 years) survived longer than patients with oligoastrocytomas (MST = 6.6 years; 49% at 10 years) or fibrillary astrocytoma (MST = 5.9 years; 31% at 10 years) (Okamoto et al. 2004). For oligodendrogliomas the MST was 11.6 years for grade II (Okamoto et al. 2004) and 3.5 years for grade III (unpublished data).(Ohgaki et al (3)).

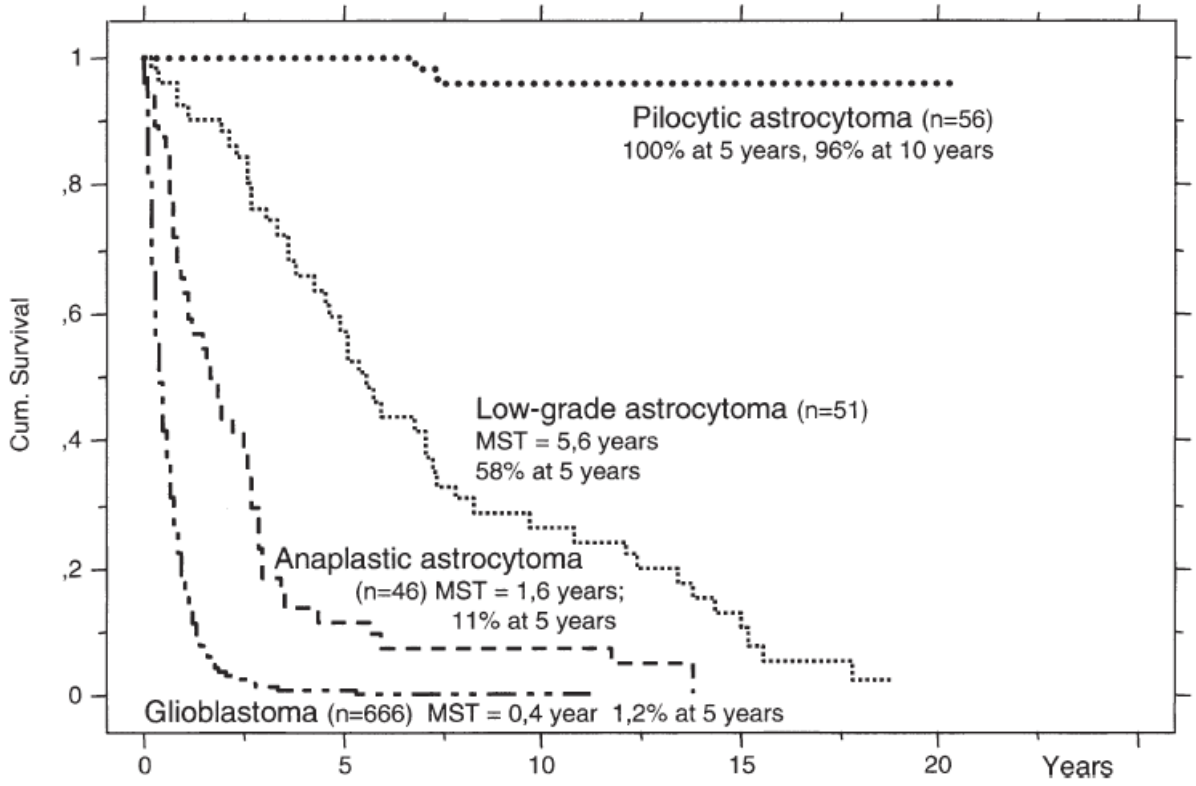

Figure 3: Survival of glioma patients (from Ohgaki et al. (3))

Molecular biology of gliomas

An understanding of glioma phenotype must also take into account the observations that astrocytic tumours are associated with *TP53* mutations, and oligodendrogliomas with 1p and 19q loss. It is possible that astrocytomas and oligodendrogliomas arise from the same cells of origin, but different genetic events drive differentiation along different lines; i.e., *TP53* mutation drives or allows astrocytic differentiation and 1p and19q loss drives or allows oligodendroglial differentiation (Figure 6.4*a*). Alternatively, astrocytomas and oligodendrogliomas may arise from different cells of origin that are oncogenically permissive only for particular events; i.e., disruption of the p53 pathway is only oncogenic as an initiating event in particular precursor cells, and oncogenesis in such precursor cells necessitates astrocytic differentiation (Figure 6.4*b*). Note that oligodendroglial tumours with 1p and 19q loss preferentially affect particular areas of the brain, raising the possibility that oncogenesis in specific precursor cell populations in different brain regions must proceed along distinct genetic pathways to reach the common phenotypic endpoint of oligodendroglioma (Zlatescu et al. 2001). (from Louis et al (5))

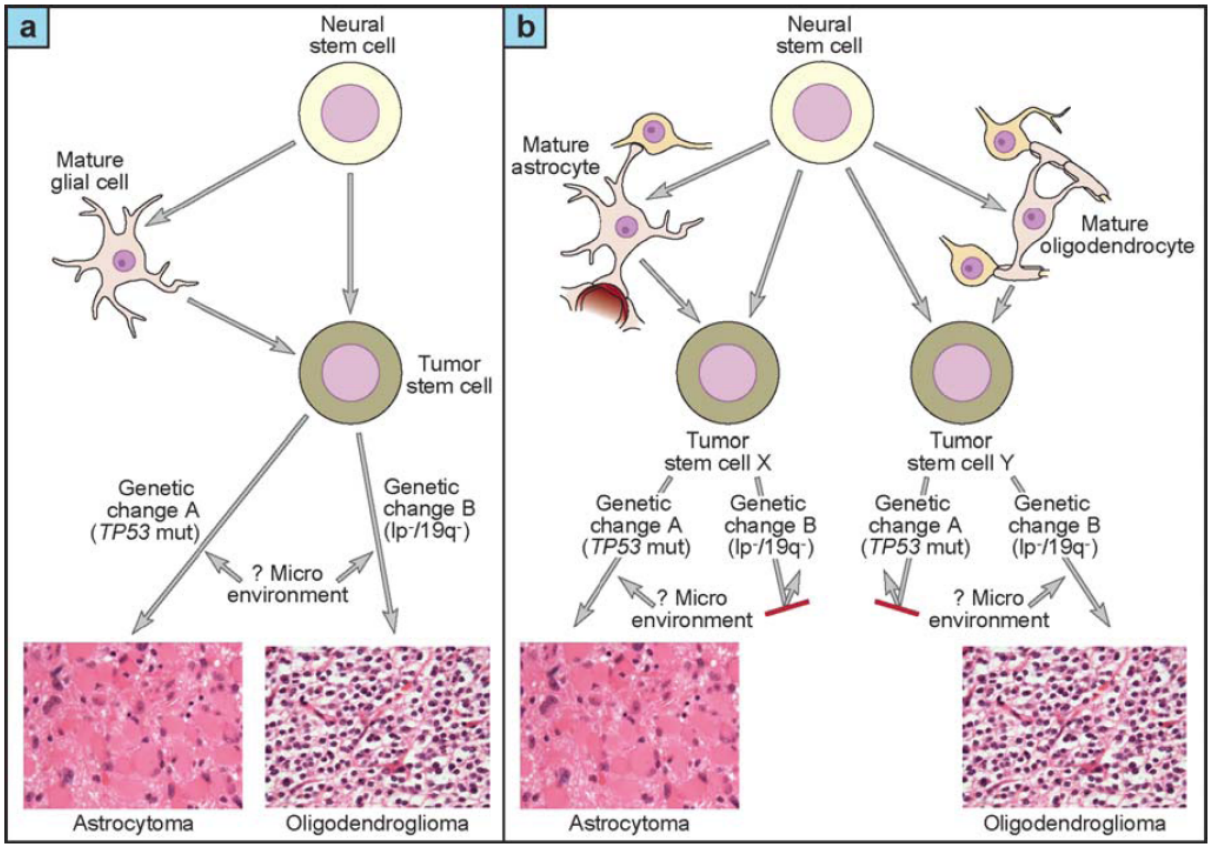

Figure 6.4: Glioma differentiation (from Louis et al. (5)) Stem cells and tumour differentiation. (*a*) A single type of tumour stem cell can arise either from a neural stem cell or from a mature glial cell. The end phenotype of astrocytoma or oligodendroglioma depends upon the activation of particular cellular pathways, either as a result of undergoing one set of genetic changes (e.g., *TP53* mutation) or another (e.g., 1p and 19q loss), possibly influenced by the tumour micro-environment. (*b*) Possible different types of tumour stem cells for various tumour types. In this model, one tumour stem cell (X) is permissive for neoplastic transformation only in the setting of particular genetic changes (e.g., *TP53* mutation), perhaps partially as a result of micro-environmental factors; other genetic changes (e.g., 1p and 19q loss) are lethal, hence the restriction to an astrocytic phenotype. Another stem cell (Y), perhaps influenced by micro-environmental factors, undergoes tumourigenesis only in the setting of other genetic changes (e.g., 1p and 19q loss), yielding an oligodendroglial tumour.

Gliomas have a remarkable tendency to infiltrate the surrounding brain, which confounds therapeutic attempts at local control (Figure 6.5). These invasive abilities are often apparent in low-grade as well as high grade tumours, implying that the invasive phenotype is acquired early in tumourigenesis. The patterns of brain invasion in gliomas are the stereotypic so-called secondary structures of Scherer that are seen in histological evaluation. For instance, there is preferential invasion along white-matter tracts: Many gliomas cross the corpus callosum to form butterfly lesions; other gliomas remain confined to the white matter, stopping abruptly at the gray-white-matter junction. Other characteristic migratory patterns include preferential growth around neurons in the gray matter (perineuronal satellitosis), perivascular growth, and subpial spread. (Louis et al. (5))

In diffuse gliomas, the cells preferentially invade along myelinated fibres in white matter tracts (intrafascicular growth), and subpial, perivascular, and perineuronal accumulation of tumour cells is frequently encountered (Giese et al., 1996) (Figure 6.6).

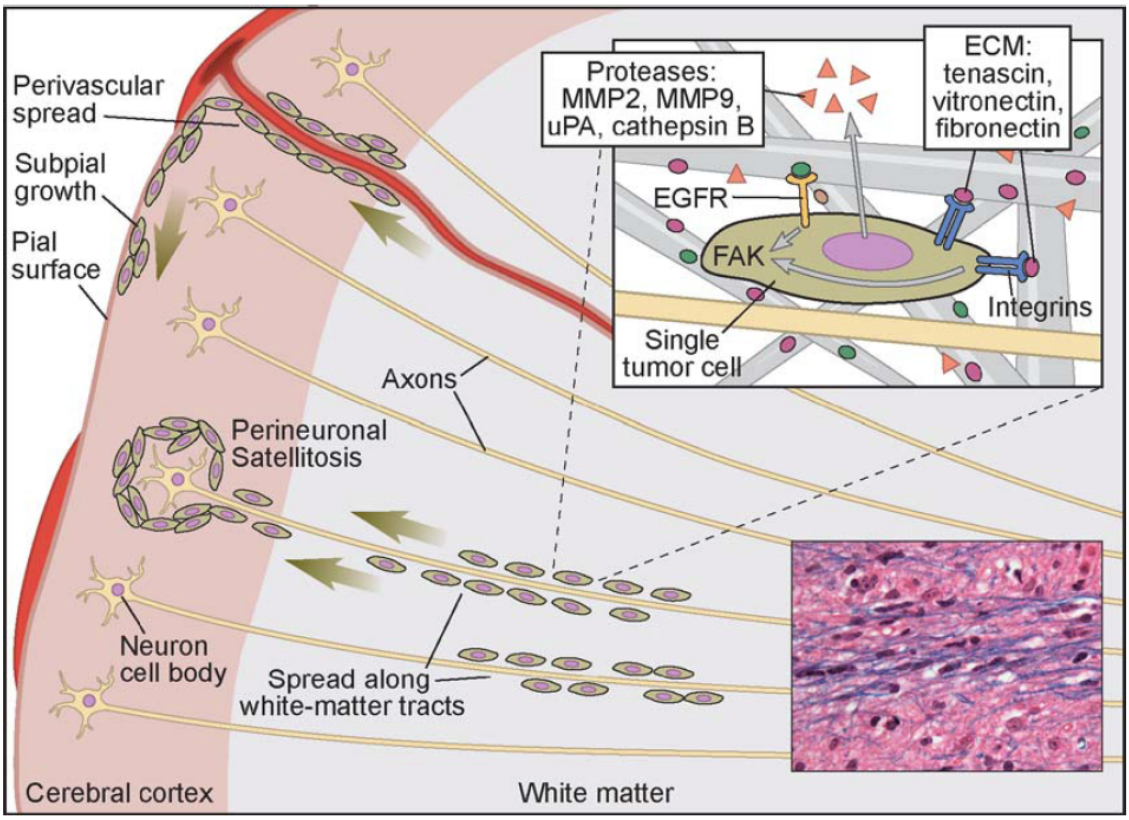

Figure 6.5: Glioma cell migration (from Louis et al. (5)) Malignant glioma cells show preferential invasion along white-matter tracts, around neurons and blood vessels, and in the subpial region. The photomicrograph (*lower right corner*; Luxol fast blue H&E stain, 400X) illustrates individual elongated, hyperchromatic tumour nuclei oriented along myelinated axons (which stain bright blue with the Luxol fast blue stain). The inset at the top right illustrates molecular events relating to the invasion of single cells: elaboration of proteases such as matrix metalloproteinases MMP2 and MMP9, urokinase-type plasminogen activator (uPA), and cathepsin B; expression of integrins that interact with extracellular matrix (ECM) components such as tenascin, vitronectin, and fibronectin that are themselves expressed by tumour cells; and activation of focal adhesion kinase (FAK)-mediated cellular signalling pathways via epidermal growth factor receptor (EGFR) or integrin signalling

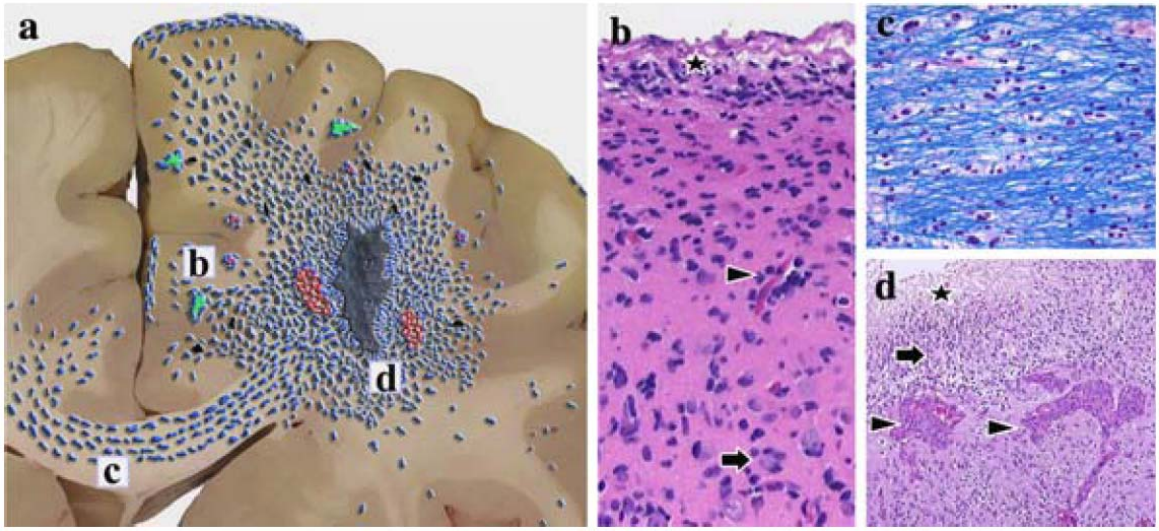

Figure 6.6: Glioblastoma growth pattern (from Claes et al. (6)) Schematic representation of the growth pattern of a GBM (**a**), including the following secondary structures of Scherer: perivascular accumulation of tumour cells (example in area indicated by *b*; vessels in *red*, tumour cells in *blue*), perineuronal satellitosis (*b*; neurons in *green*), subpial growth of tumour cells (**b**), and intrafascicular growth in the corpus callosum (**c**). Mitotic tumour cells are depicted in *black*. Furthermore, in GBMs necrosis (*dark grey area*) surrounded by pseudopalisading tumour cells and adjacent florid/glomeruloid microvascular proliferation (**d**) are often present. Images **b–d** on the *right* represent the histology of these features: in **b** *asterisk* indicates subpial growth, *arrow* indicates perineuronal satellitosis, *arrowhead* indicates perivascular accumulation of tumour cells; image **c** shows increased cellularity with diffuse infiltration of tumour cells in the relatively well preserved myelinated tracts of the corpus callosum; in image **d** *asterisk* indicates area of necrosis, *arrow* indicates peri-necrotic pseudopalisading tumour cells, *arrowheads* indicate glomeruloid microvascular proliferation [**b**, **d**: H&E staining, **c**: combined Luxol Fast Blue and H&E staining; original magnification 200 (**b**, **c**) and 100 (**c**)]

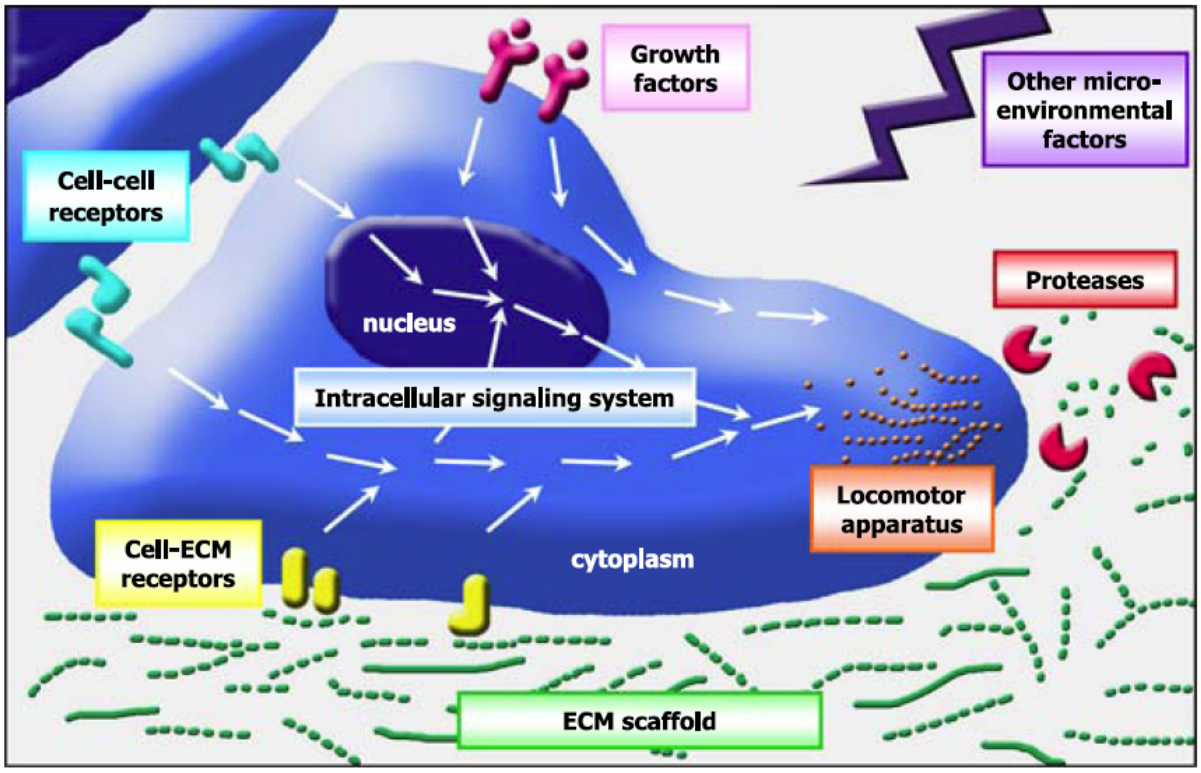

Figure 6.7: Schematic overview of factors and mechanisms important for difuse infltration of glioma cells in the neuropil (from Claes et al. (6)). In this scheme the protrusion on the right side of the cell represents the lamellipodium at the front. (for further information, see Claes et al.)

The following aspects relevant for glioma growth pattern can be recognized (figure 6.7):

1. an intracellular system that coordinates all incoming and outgoing signals via a complex set of pathways,
2. a locomotor apparatus in which the actin cytoskeleton plays a crucial role,
3. a scaffold (ECM, surface of cells/cell processes) on which the glioma cells can travel,
4. cell–ECM and/or cell–cell receptors that allow direct interaction with the ECM and cellular microenvironment,
5. tools to remove obstacles like ECM degrading proteases,
6. growth factors that guide the way, and
7. other stimulatory or permissive microenvironmental factors (e.g., chemokines derived from inflammatory cells).

Genetic hallmark of low-grade diffuse astrocytomas is frequent TP53 mutation (>60%). That frequency does not significantly increase during malignant progression to secondary GBM indicating that this is an early event. The most frequent genomic imbalance is gain of chromosome 7q and amplification of 8q. Other observed genetic changes are LOH on 22q in 17%, and deletions of chromosome 6 in 14%. LOH on 22q on one or more loci is found in 27-33% of WHO II astrocytomas. The most frequent copy number aberrations are gains on 7q, 5p, 9 and 19p, and losses on 19q 1p, and Xp. A third of low-grade astrocytomas show p14ARF promotor methylation and about 50% show MGMT promotor methylation.

Anaplastic astrocytoma show high frequency of TP53 mutations and LOH on 17p (50-60%), similar to that of diffuse astrocytoma. LOH of10q is observed in 35-60% of AA, PTEN mutations in 18-23%. LOH of 22q is shown in a similar frequency as in WHO II astrocytomas (20-30%). LOH of 19q is significantly more frequent than in DA (46%). A third of AAs show LOH of 6q. EGFR amplification is uncommon in anaplastic astrocytoma (<10%)

The pattern of genetic changes, in particular the high frequency of p53 mutations (>70%), would be compatible with the assumption that such tumours progressed rapidly from DAs. (Louis et al. (1)) (see also Figure 6.8)

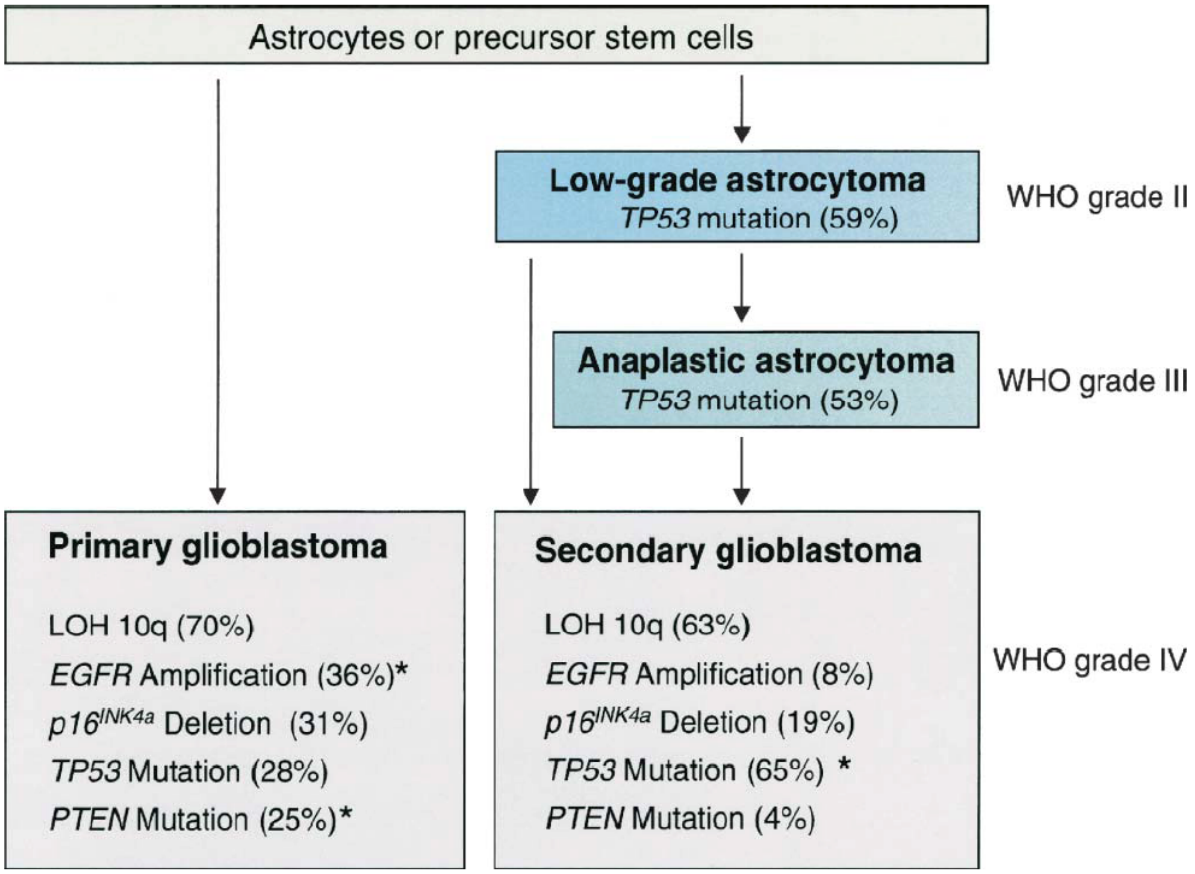

Figure 6.8: Genetic pathways of glioma progression (from Ohgaki et al. (3)) Genetic pathways to primary (de novo) and secondary glioblastomas. LOH 10q is frequent in both primary and secondary glioblastomas. TP53 mutations are early and frequent genetic alterations in the pathway leading to secondary glioblastomas. Asterisks (*) indicate genetic alterations that are significantly different in frequency between primary and secondary glioblastomas.

Humeral response against glioma

We will further elucidate the complexity of the humeral response against glioma. We will screen an array encompassing 1827 antigen-expressing clones with sera of glioma patients and of persons without known disease. Analysis will be conducted by newly developed computer aided image analysis software that was tailored to the evaluation of protein macroarrays exhibiting seroreactivity pattern.

The following outline describes the general form of results that will be obtained for the glioma model. A similar approach has been conducted for meningioma, a benign intracranial tumour.

Screening procedure

Recently we screened 38.016 recombinant *E. coli* clones of the hex1 library [Büssow et al.] with sera of patients with different cancers and inflammatory diseases. We identified 1827 reactive clones that were subarrayed in duplicates and screened with 53 sera of meningioma patients and 60 sera of persons without known disease. In brief, macroarrays were washed twice with TBSTT (TBS, 0.05 % Tween20, 0.5 % Triton-X100) and 4 times with TBS. Membranes were blocked 2 hours in TBST (TBS, 0.05 % Tween20) with 3 % non-fat dry milk and subsequently incubated over night with sera diluted 1:1000 in TBST/3 % dry milk. Sera were retained and stored for a second incubation round. Filters were washed twice in TBST and subsequently incubated in stripping solution at 70°C. After 2 times washing in TBST and 4 times in TBS, macroarrays were again blocked for 2 h with TBST/3 % dry milk and incubated with the same serum dilution over night. For the detection of bound autoantibodies, arrays were washed thrice in TBST and subsequently incubated with secondary antibody (1:1000 rabbit anti-human IgG, IgA, IgM-Cy5 (H+L) (Dianova) in TBST/3 % dry milk). After washing 4 times in TBST and 2 times in TBS, membranes were dried over night and scanned with Typhoon 9410 scanner (GE Healthcare).

Image analysis and statistics

Evaluation of spot intensity was carried out by a novel computer-aided image analysis procedure. In brief, the protein array was segmented in target areas so that each area contained a single protein spot. The pixels contained in a target area were clustered in foreground and background pixels using k-means clustering. The dark foreground was extracted from the pale background by applying the so-called black top-hat operator that is established in the field of image analysis and the mean intensity of all foreground pixels was assigned to each spot. Since each expressed protein was spotted in duplicates on the macroarray, the mean intensity of the two replicates was assigned to each protein. Thus, the image analysis of each macroarray resulted in an autoantibody profile consisting of 1827 integer intensity values ranging from 0 to 255, the standard range of values in a grey scale image. Please note that our fully automated procedure assigned 'not-available' values to spots where no appropriate spots can be detected. Additionally, we computed for each clone the number of 'not-available' spots in all samples, e.g. serum analyses that did not yield intensity values by the automated analysis system. Clones with more than 10 such not-available values were excluded from further analysis.

We carried out standard quantile normalization to minimize array-to-array variations. For the different classification tasks, we used linear Support Vector Machines. Mean sensitivity, specificity and accuracy were computed by 100 repetitions of standard 10-fold cross validation. Additionally, we performed 100 classification runs with randomly permuted class labels to test for overtraining. For each seroreactive clone we calculated the "area under the receiver-operator-characteristics-curve (ROC) value" (AUC) as a measure of its information content for the corresponding classification task. The ROC curve shows the sensitivity as a function of (1-specificity) while the discrimination threshold is varied. To study the diagnostic potential of each seroreactive clone, we considered the putative intensity values ranging from 0 to 255 as discrimination thresholds. For a given clone c and threshold h, we considered meningioma sera as true positive (TP) if clone c had an intensity value larger or equal h. If c had an intensity value smaller h, meningioma sera were typed as false negative (FN). Control sera with intensity value above the threshold were considered false positive (FP), control sera with intensity value below the threshold as true negative (TN). We determined sensitivity (TP/(TP+FN)) and specificity (TN/(TN+FP)) for all possible thresholds h and computed the ROC curve and subsequently the AUC value of the considered antigen. Since AUC values of 0 and 1 indicate a perfect separation, we considered antigens with AUC values below 0.3 or above 0.7 as informative for a separation task.

To determine how frequent each clone reacts with a certain type of sera, e.g., sera of meningioma patients, we only considered reactivities with intensity values larger or equal 50 as positive.

Classification of meningioma sera using protein macroarrays

We screened 53 sera of patients with meningioma and 60 sera of persons without known disease in the following referred to as normal sera. The sera were tested for autoantibodies against 1827 immunogenic *E. coli* clones. The clones represented 509 human in-frame peptide sequences and 1318 human out-of-frame sequences. Reactivity of serum autoantibodies against the clones was measured by a novel automated image analysis system. Following quantile normalization we excluded clones that were mostly not recognized by the image analysis system: In detail, we computed for each of the 1827 clones the number of serum tests that did not yielded intensity values. We excluded 410 clones with more than 10 'not-available' values. Using linear Support-Vector Machines we were able to discriminate meningioma sera from normal sera with a specificity of 95.62 % (95 % CI: 95.16-96.07), a sensitivity of 91.83 % (95 % CI: 91.47-92.19) and an accuracy of 93.84 % (95 % CI: 93.52-94.16). The mean specificity, sensitivity and accuracy of these classifications together with the 95 % confidence intervals (CI) are shown in Table 6.1. As control, we carried out 100 permutation tests with randomly permutated the class labels. We obtained a mean specificity, sensitivity and accuracy of approximately 50 % for this classification.

Table 6.1:

classification	specificity [%]	sensitivity [%]	accuracy [%]
Meningioma vs healthy	95,6 [95,17-96,07]	91,8 [91,47-92,19]	93,8 [93,52-94,16]
random	52,9 [51,61-54,19]	44,3 [42,84-45,65]	48,8 [47,67-50,01]

Information content of seroreactive clones

We ranked 1417 seroreactive clones according to the information content that they contributed to the classification task. We computed the AUC value for each clone as detailed in Material and Methods. Approximately 95 % of the seroreactive clones show AUC values between 0.3 and 0.7, e.g., they are considered as non-informative. An overview on the distribution of AUC values among the clones is provided in Table 6.2. Out of the 1417 seroreactive clones, 60 clones including 23 in-frame clones are considered informative for the classification of meningioma sera versus normal sera. The in-frame clone that shows the best AUC value (0.202) for the classification of meningioma patients versus normal controls is exemplarily shown in figure 6.9.

Table 6.2:

AUC range	Meningioma vs healthy	
	No. of clones (all)	No. of clones (in-frame)
0.0-0.1	0	0
0.1-0.2	1	0
0.2-0.3	41	21
0.3-0.4	285	92
0.4-0.5	516	149
0.5-0.6	423	107
0.6-0.7	133	20
0.7-0.8	18	2
0.8-0.9	0	0

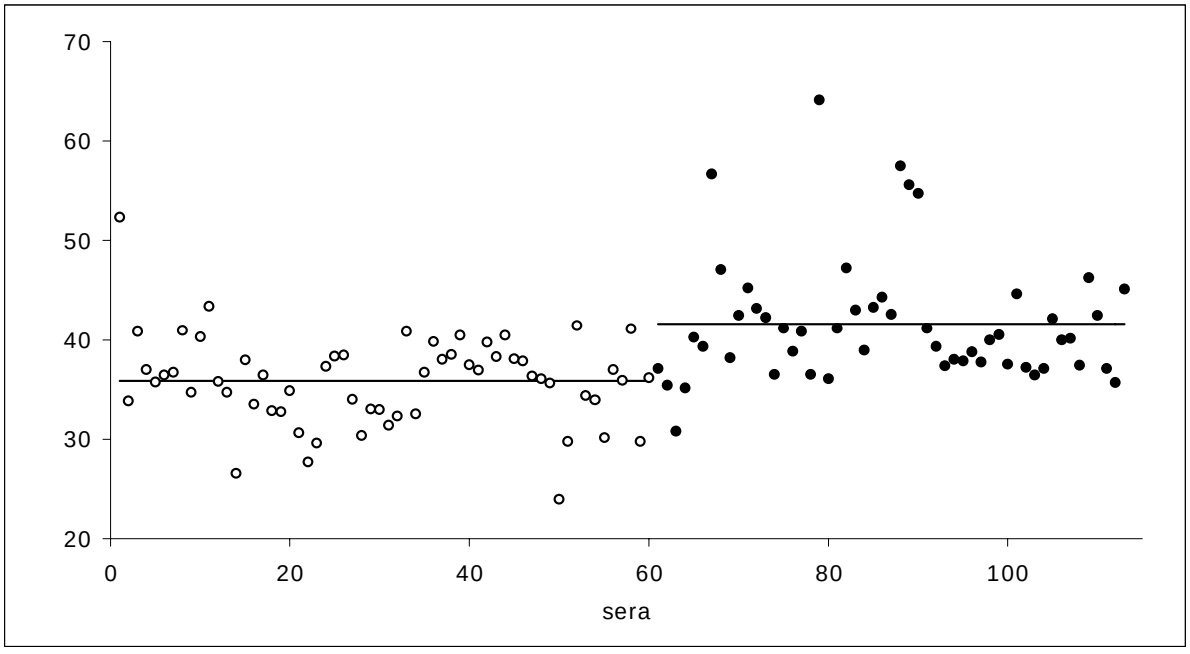

Figure 6.9: Intensity values of seroreactivity against the best clone for the classification of meningioma versus healthy are provided for healthy controls (empty circle) and for meningioma sera (full circle) Frequency of seroreactive clones The graph shows the intensity values of seroreactivity (y-axis) in the tested sera (x-axis). Mean seroreactivity value is indicated as horizontal line.

We determined how frequent each seroreactive clone reacted with serum autoantibodies of patients and normal controls. We considered only reactions that yield an intensity value above 50 in our automated image analysis system. In total, we found 1417 antigenic clones including the 391 in-frame clones. Following normalization we computed the mean reactivity of each seroreactive clone in patients and controls (Table 6.3). Out of the 1417 clones, 37 clones including 6 in-frame clones reacted with more than 60 % of meningioma sera and 43 clones including 7 in-frame clones reacted with more than 60 % of healthy sera. Among those clones that reacted with at least 20 % of meningioma sera, we identified 56 clones including 20 in-frame clones that react at least twice as frequent with the sera of meningioma patients as with normal sera. The highest frequency was detected for a clone that reacted with 49.06 % of meningioma sera but also with 23.33 % of control sera.

Table 6.3:

Frequency range	Meningioma		healthy		Meningioma (twice)	
	No. of clones (all)	No. of clones (*in-frame*)	No. of clones (all)	No. of clones (*in-frame*)	No. of clones (all)	No. of clones (*in-frame*)
0 %	105	40	176	62	-	-
0 – 10 %	578	171	552	169	141	70
10 – 20 %	321	85	305	78	67	28
20 – 30 %	184	43	163	36	35	**12**
30 – 40 %	119	30	93	22	15	**6**
40 – 50 %	49	14	55	8	3	**2**
50 – 60 %	24	2	30	9	1	0
60 – 70 %	20	3	27	5	2	0
70 – 80 %	11	2	11	2	0	0
80 – 90 %	5	1	4	0	0	0
90 – 100 %	1	0	1	0	0	0

Data used for the glioma model

In deliverable D4.1 chapter 5.1 the data used for the adaptation, optimization and validation of the models for glioma tumour growth and response to treatment are listed. The following classification of data is done:

- A1. Treatment data
- A2. Normal tissue complication data
- A3. Clinical data
- A4. Imaging data
- A5. Segmentation data
- A6. Histopathological data
- A7. Molecular data

For details see D4.1 chapter 5.1. ObTiMA, developed by ACGT, will be used as data management system. For further details see chapter 8 of this deliverable. There are more items provided than given in D4.1. These tables are part of the prospective trial for gliomas, in which more data are collected. Molecular data will not be stored on CRFs in ObTiMA. They will be provided in a repository for genetic data after pseudonymisation and annotation to the Gene Ontology. Segmentation data are stored in an imaging repository after pseudonymisation.

Scenario and Workflow of the glioma model

The workflow of the glioma model is given in figures 6.10 and 6.11. At the time of diagnosis clinical data, imaging data and histological data will be collected and pseudonymized. Segmentation of the tumour will be done using the segmentation tool described in chapter 4. Autoantibodies against glioma antigens, cytogenetic and cDNA data will be analyzed. All these data will be used for the *in silico model*. The model will simulate the tumour response of the given treatment of the patient. A correlation between histology and imaging data and molecular data is mainly impossible, because there will no gross total resection of a glioma possible. Therefore a fusion of the histological picture with the imaging data of MRI is most unlikely. The correlation to molecular biology to histology and imaging is impossible for the same reason. During treatment of the patient toxicity data are collected and used for the prediction of side effects on normal tissue.

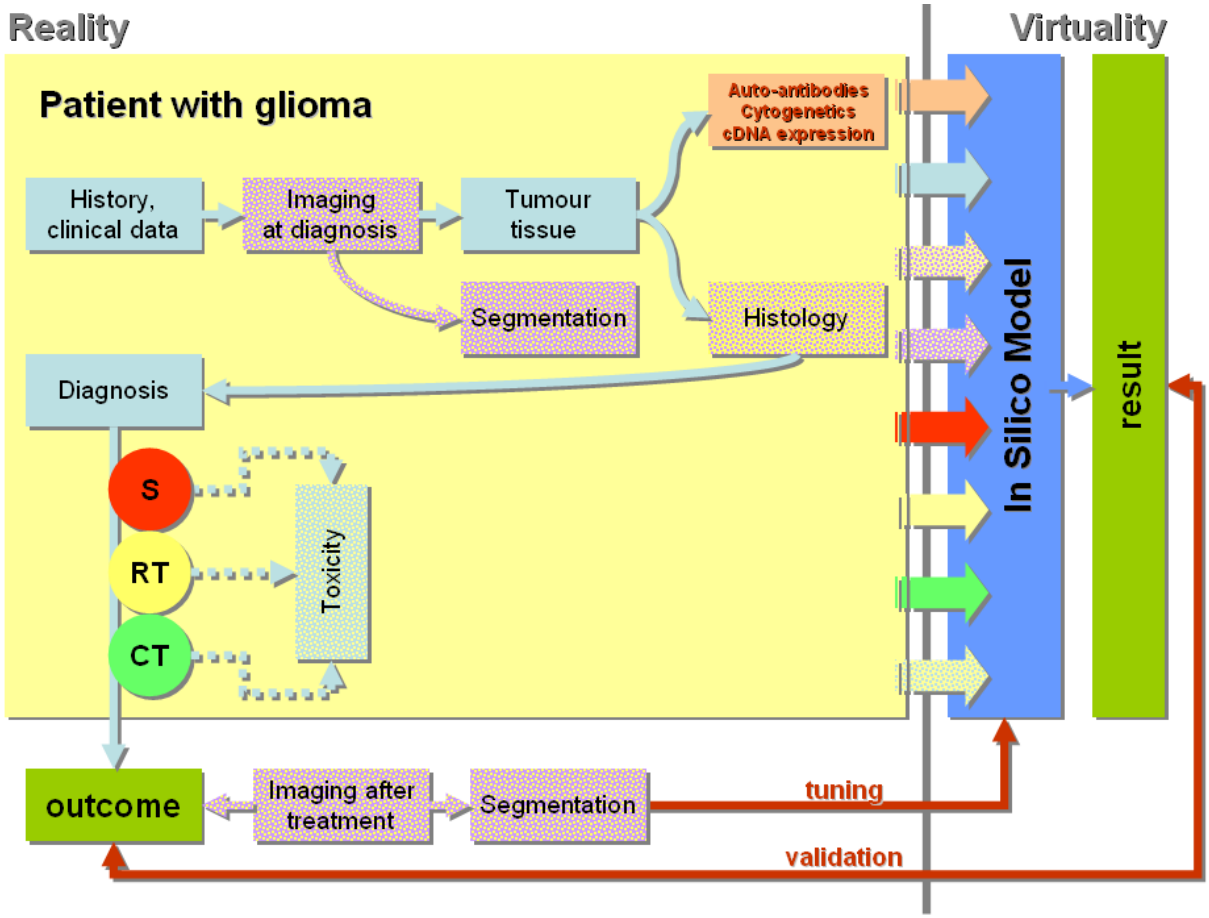

Figure 6.10: Workflow of the glioma scenario. S: surgery; RT: radiotherapy; CT: chemotherapy.

Validation of the glioma model

At the end of treatment the outcome of the patient regarding tumour volume will be used for validation of the model. The tumour of the patient will be segmented and the tumour volume will be calculated. The difference between the real tumour volume found in imaging studies of the patient and the simulated tumour volume has to tune the *in silico model*. Differences greater than 10 % between the real tumour volume and the simulated volume are unacceptable. In this case optimization of the model is needed. The model can be considered as helpful for clinicians if in at least 50 patients a correct result (less than 10 % deviation between reality and virtuality) is predicted. In this situation the model might be tested as a decision making tool for patients with glioma in prospective clinical trials.

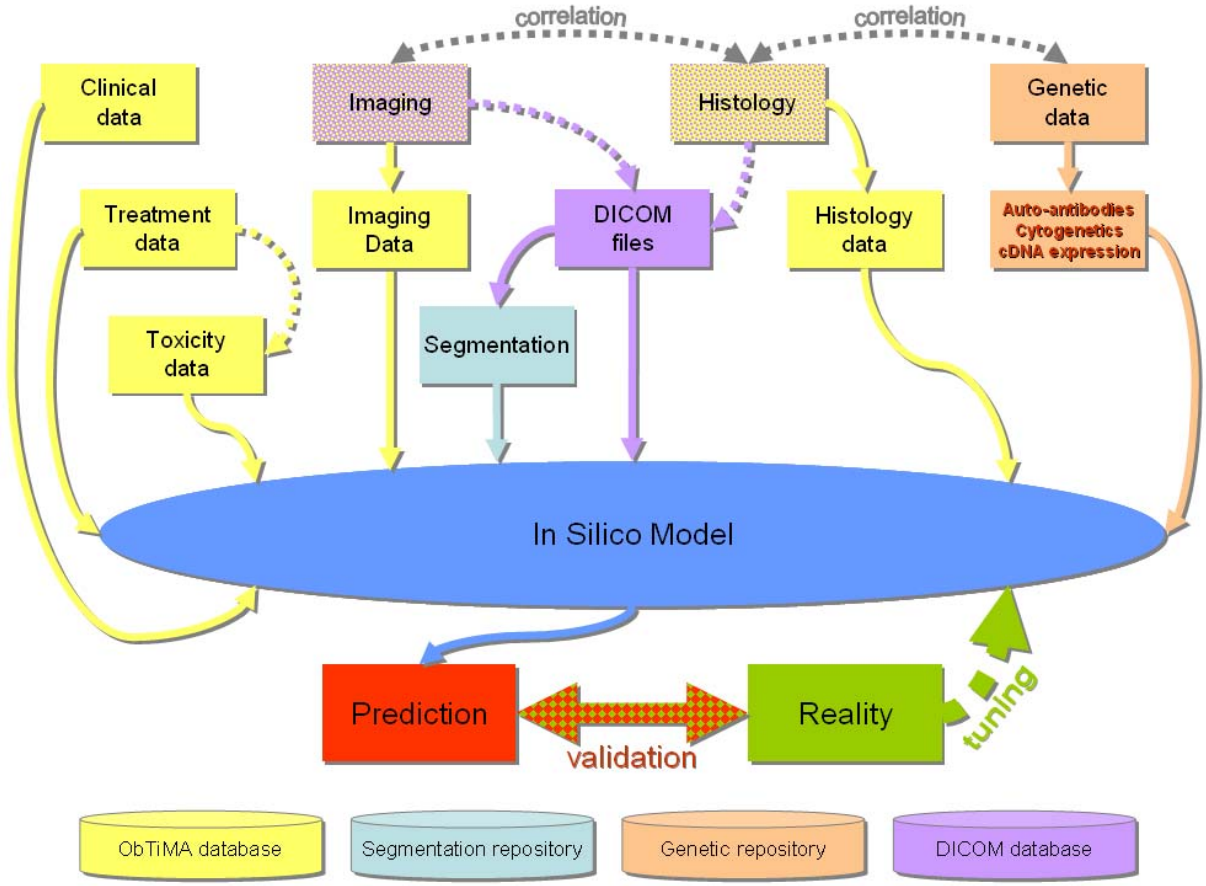

Figure 6.11: Workflow of data of the glioma scenario.

7. Lung cancer

The following introduction (7.1 and 7.2) is based on:

1. Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC (eds) (2004) WHO Classification of Tumours. Pathology and genetics of tumours of the lung, pleura, thymus and heart. IARC Press, Lyon
2. Goldstraw P, Crowley J, Chansky K et al. (2007) The IASLC lung cancer staging project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. J Thorac Oncol 2:706-714
3. Rusch VW, Asamura H, Watanabe H et al.(2009) The IASLC lung cancer staging project. A proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. J Thorac Oncol 4:568-577
4. Risch A, Plass Chr (2008) Lung cancer epigenetics and genetics. Int. J. Cancer: 123, 1–7.
5. Peto R (2004). WHO Mortality statistics with UN population estimates, 1950-2000. www.ctsu.ox.ac.uk
6. Borczuk C, Toonkel RL, Powell CA (2009) Genomics of Lung Cancer. Proc Am Thorac Soc 6: 152–158

Clinical background of lung cancer

Lung cancer is the leading cause of cancer deaths among adults worldwide. Over 1.2 million new cases and 1.1 million deaths worldwide were predicted in 2004. The more than one million deaths represent almost 18% of all cancer deaths worldwide. Thus, lung cancer is the most common and deadliest cancer in the world. The 5-year survival rate in Europe is only 10%, which is not much better than the recorded rate of 8.9% noted in developing countries.

The clinical symptoms and signs of pulmonary cancer depend on the tumour location and extent. Data from the United States show that 16% of the patients presented with localized disease, 37% with regional disease and 39% with distant disease while 8% of the patients were unstaged.

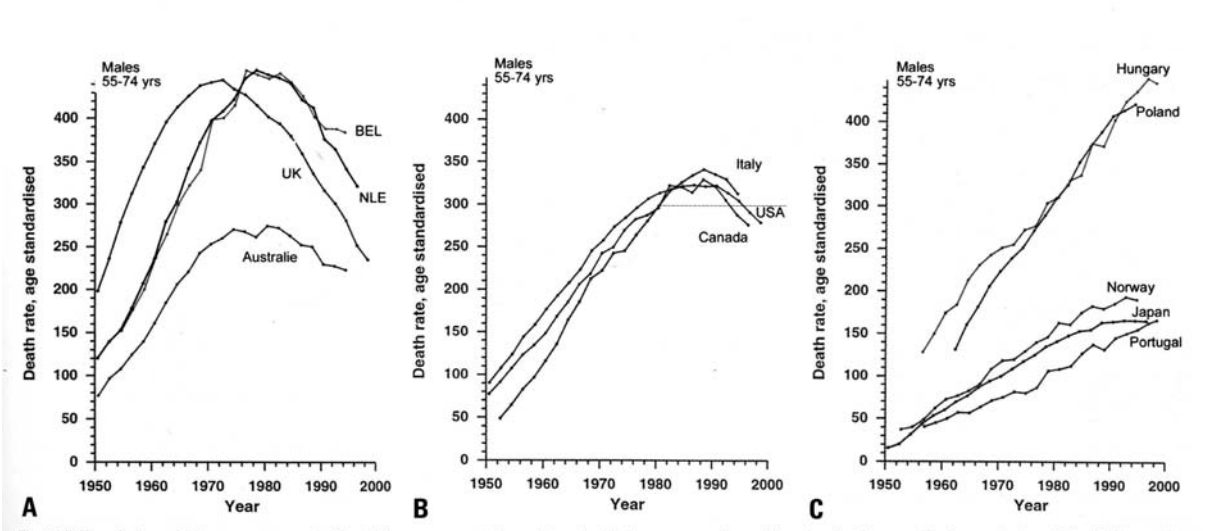

Fig. 1.02 Trends in male lung cancer mortality. **A** In some countries with early, high rates, a substantial reduction in mortality began in the 1970s (UK) or 1980s (Belgium, Netherlands, Australia). **B** In other countries (Italy, USA, Canada), the decline started in the 1990s. **C** Failure to achieve a significant reduction in tobacco consumption until recently has in some countries caused rising lung cancer mortality without apparent levelling off in males at ages 55-74. From R. Peto et al. {1589A}.

Figure 7.1: Lung cancer mortality from 1950-2000 (from Peto (5))

Central and peripheral located tumours differ in their clinical appearance. While central tumours, arising in the large airways, cause cough, wheezing and haemoptysis, peripheral tumours can cause chest pain, dyspnoea and cough. Atelectasis with subsequent pneumonia mainly occur in central tumours. Due to the close connection with the visceral pleura, peripheral tumours can spread within the thorax. A variety of clinical complaints can occur, as pleural effusions, pleural seeding of tumour cells, vascular obstruction as superior vena cava syndrome or Pancoast syndrome, if apical lung cancer invades the parietal pleura and the lower brachial plexus. Beyond, severe complaints can occur from bulky lymph node metastasis obstructing the oesophagus and from pericardial or cardiac involvement leading to tamponade or constriction. Lymphangitic and haematogenous dissemination is common in lung cancer. At least one third of the lung cancer patients present symptoms of extrathoracic manifestations.

Almost all lung cancers are carcinomas (other histologies comprise well under 1%). In the combined data from the series, small cell carcinomas comprise about 20% of cases and large cell /undifferentiated carcinomas about 9%. But for the other histological types, the proportions differ by sex: squamous cell carcinomas comprise 44% of lung cancers in men, and 25% in women, while adenocarcinomas comprise 28% cases in men and 42% in women. Incidence rates, and the estimated rates by histological subtype have been reported for 30 populations for which a relatively high proportion of cases had a clear morphological diagnosis. Among men, only in certain Asian populations (Chinese, Japanese) and in North America (USA, Canada) does the incidence of adenocarcinoma exceed that of squamous cell carcinoma. In women, however, adenocarcinoma is the dominant histological type almost everywhere, except for Poland and England where squamous cell carcinomas predominate, and Scotland where small cell carcinoma is the most frequent subtype.

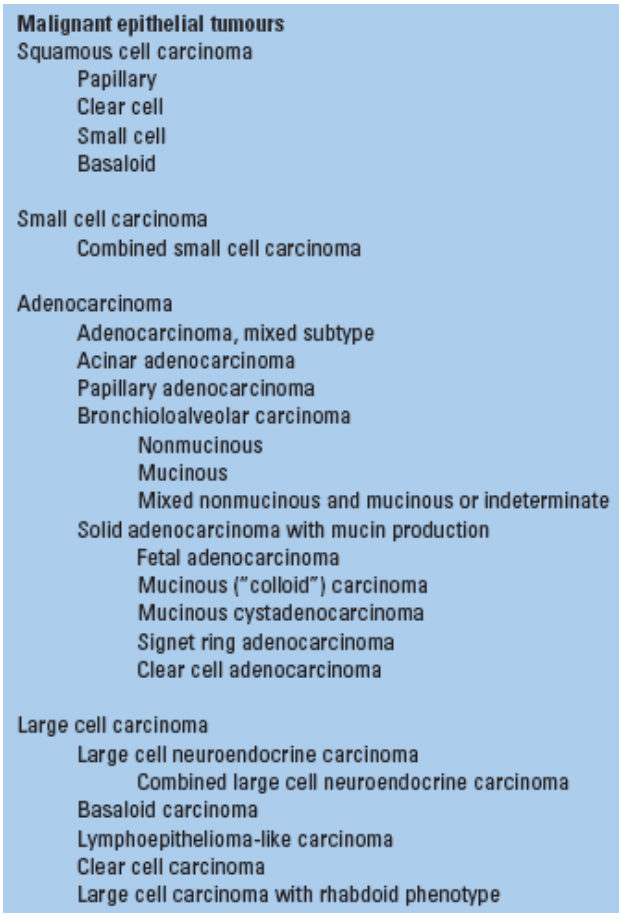

Figure 7.2: WHO histological classification of tumours of the lung (from Travis et al. (1))

Adenocarcinomas are particularly predominant in Asian females (72% cancers in Japan, 65% in Korea, 61% in Singapore Chinese). The differences in histological profiles are strongly influenced by the evolution of the epidemic of smoking-related lung cancer over time.

Staging of Lung Cancer is currently done according to the two main histological types of lung cancer, the NSCLC (non-small cell lung cancer) type – squamous cell carcinoma, adenocarcinoma, large cell carcinoma - and the SCLC (small cell lung cancer) type.

The internationally accepted TNM staging system is recommended for the NSCLC types. The stage of the disease is important for prognosis and treatment planning. Pathologic staging is based on the pathologic evaluation of sampled tissues according to the TNM system. For patients in whom surgical resection is attempted, there are surgical protocols for sampling the lymph node stations, including superior mediastinal nodes (numbered 1-4), aortic nodes (numbered 5 and 6), inferior mediastinal nodes (numbered 7-9) and nodes associated with the lobectomy specimen labelled “N1” nodes (numbered 10-14).

For the Staging of SCLC the TNM staging classification is generally not utilized in SCLC, as it does not predict well for survival. SCLC is usually staged as either limited or extensive disease. The consensus report of the International Association for the Study of Lung Cancer (IASLC) modified the older VALG classification in accordance with the revised TNM system: Limited disease is restricted to one hemithorax with regional lymph node metastases including hilar ipsilateral and contralateral, mediastinal ipsilateral and contralateral, supraclavicular ipsilateral and contralateral, ipsilateral pleural effusion (independent of cytology). Limited disease is equivalent to stage I - III of the TNM system.
Extensive disease is defined as disease beyond the definition of limited disease, equivalent to stage IV in the TNM system.

Recently a proposal for a new TNM classification of lung cancer was published which combines NSCLC and SCLC staging. It is very likely that this proposal will be included in the 7th edition of the TNM classification appearing in late 2009.

Squamous cell carcinoma

Squamous cell carcinoma (SCC) is defined as a malignant epithelial tumour showing keratinisation and/ or intercellular bridges that arises from bronchial epithelium.

Over 90% of squamous cell lung carcinomas occur in cigarette smokers. The majority of squamous cell lung carcinomas arise centrally in the mainstem, lobar or segmental bronchi. Up to 50% of SCCs occur in the peripheral lung. Macroscopically the tumours are white or grey and contain focal carbon pigment deposits in the centre. The periphery of the tumours often appears with star-like retractions. Clinically SCCs are best demonstrated by CT scan.

SCCs tend to be locally aggressive involving adjacent structures by direct contiguity. Metastases to distant organs are much less frequent than in pulmonary carcinomas of other histologic types. Thus, locoregional recurrence after surgical resection is more common in squamous cell carcinoma than in other cell types. The prognosis of SCC is, stage for stage, significantly better than for adenocarcinoma, reaching an 80% five year survival rate at stage 1 (T1 N0 M0). Up to now histologic appearance of necrosis seems to be an important prognostic factor, while other histologic parameters do not show a high correlation with prognosis. A clear-cut survival difference exists between the surgical cases and the rest (70% of the patients).

Currently, the stage of disease and the performance status at diagnosis remain the most powerful prognostic indicators for survival for primary squamous cell carcinoma.

Adenocarcinoma

Adenocarcinoma (AC) of the lung is defined as a malignant epithelial tumour with glandular differentiation or mucin production, showing acinar, papillary, bronchioloalveolar or solid with mucin growth patterns or a mixture of these patterns.

Adenocarcinoma is in many countries now more frequent as compared to SCC. Some types of adenocarcinomas, i.e. the bronchioloalveolar subtype, develop more frequently than any other

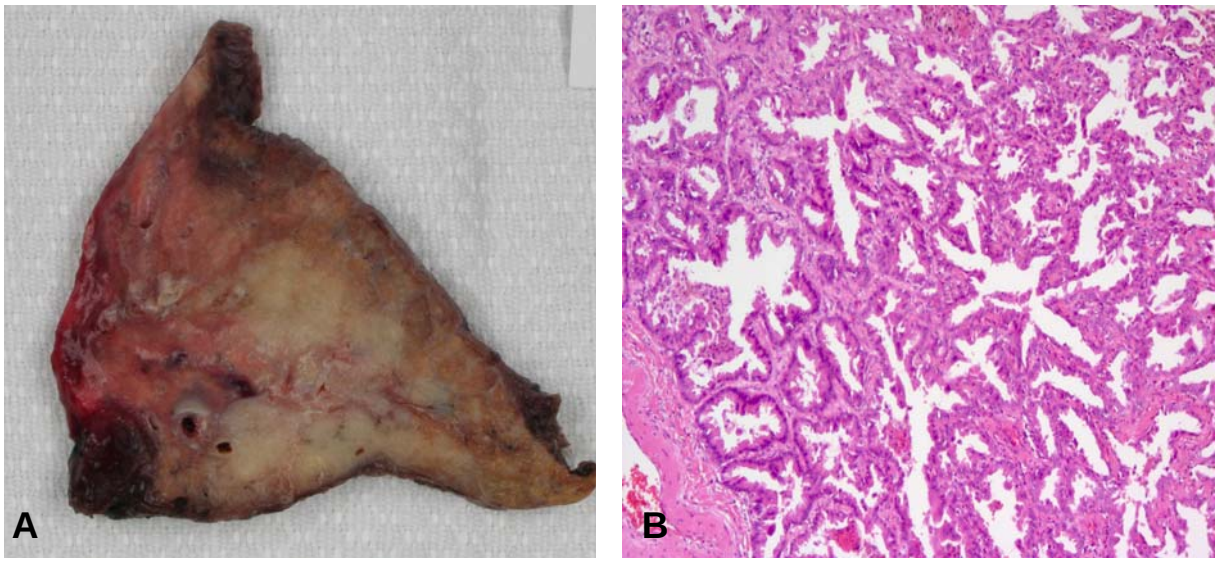

Figure 7.3: Adenocarcinoma of the lung. Resection specimen of a peripheral adenocarcinoma (A), histological subtype: acinar adenocarcinoma (B, hematoxylin-eosin)

histologic type of lung cancer in individuals (particularly women) who have never smoked. ACs typically appear at the periphery of the lung either as single or as multiple nodules. In contrast to other histologic subtypes, in adenocarcinomas of the lung aerogenous dissemination occurs, especially in the bronchioloalveolar subtype.

The most frequent type of pulmonary adenocarcinomas is the “adenocarcinomas mixed subtype”, representing approximately 80% of resected adenocarcinomas. The major individual histologic patterns/ subtypes are acinar, papillary, bronchioloalveolar, and solid adenocarcinoma with mucin production.

Unfavourable prognostic factors are a high histological grade, the occurrence of vascular invasion, an increased mitotic activity, relatively few tumour infiltrating lymphoid cells, and extensive tumour necrosis.

Large cell carcinoma

Large cell carcinoma (LCC) is an undifferentiated non-small cell carcinoma lacking the cytologic and architectural features of SCC, adenocarcinomas and small cell carcinomas. Together with large cell neuroendocrine carcinoma it accounts for up to 12% of all lung cancers and occurs preferentially in smokers.

LCC typically present as large, peripheral masses, invading the visceral pleura, chest wall, or adjacent structures. Sectioning reveals a soft, pink-tan tumour with frequent necrosis.

Most LCC show a stage III-IV at diagnosis, frequently metastases occur to the hilar or mediastinal nodes while haematogenous dissemination occur to the liver, bone, brain. As compared to the classic LCC, combined large cell neuroendocrine carcinoma, large cell carcinoma with rhabdoid phenotype and basaloid carcinoma have a worse prognosis.

Overall, the prognosis of LCC, i.e. large cell neuroendocrine carcinoma is not significant different as compared to SCLC if stratified by stage.

Small cell carcinoma

Small cell carcinoma of the lung (SCLC) is a malignant epithelial tumour consisting of small cells with scant cytoplasm, ill-defined cell borders, finely granular nuclear chromatin, and absent or inconspicuous nucleoli. The cells are round, oval and spindle-shaped. Nuclear moulding is prominent. Necrosis is typically extensive and the mitotic count is high. In combined small cell carcinomas SCLC occur together with any histologic type of non-small cell carcinoma.

SCLCs typically occur in a central location of the lung with nodal involvement but are often disseminated (e.g. to bone marrow and liver) at the time of primary diagnosis. Later brain metastases appear. Paraneoplastic syndromes are a typical phenomenon of SCLC. Sometimes the primary tumour is very small and can not be detected on radiographic studies. A few SCLC are observed in the peripheral lung. They are radiographically indistinguishable from other pulmonary neoplasms. Macroscopically and histologically SCLC are soft, friable perihilar masses that show extensive necrosis, typically spreading along bronchi in a submucosal and circumferential fashion, often involving lymphatics.

The tendency for widespread dissemination at presentation has led to small cell carcinoma being staged as limited versus extensive disease rather than using the TNM system. SCLC is by definition high grade, and shows a high mitotic rate, thus grading is inappropriate. In small biopsies the differential diagnosis includes lymphoid infiltrates, other neuroendocrine tumours, other “small round blue cell tumours” (SRBCT), and primary or metastatic non-small cell carcinomas. Immunohistochemical staining for cytokeratin vs. leukocyte common antigen as well as neuroendocrine markers and TTF-1 or corresponding cytologic specimens may be helpful. Occasionally, the morphologic separation of SCLC from NSCLC can be difficult.

The prognosis of SCLC is mainly depending on the stage of the disease, which is separated into two categories: the limited and the extensive stage.

The former is defined by SCLC disease restricted to one hemithorax with regional lymph node metastases, including hilar, ipsi-, and contralateral mediastinal, or supraclavicular nodes or to patients with contralateral mediastinal lymph nodes and supraclavicular lymph nodes and to patients with ipsilateral pleural effusion (benign or malignant). All patients with disease beyond the limited stage are classified as extensive stage disease.

Molecular biology of lung cancer

Cancer gene signatures and lung development

Currently the molecular genesis of lung cancer is seen as a multistep process originating in pluripotent stem and progenitor cells. These cell types have shown to be capable of differentiation into one or several histologic cell types. Several data point to significant influences of gene transcriptional activation and/or gene repression in early tumour cells recapitulating important parts of embryonic lung development. The hypothesis that lung cancer arises from aberrant expression of genes involved in lung development is supported by gene expression studies demonstrating similarities between signatures obtained from human lung tumours and signatures characteristic of normal lung development. Hierarchical cluster analyses of NSCLC showed that tumours can be segregated by molecular profiles according to their histologic type and differentiation. Beyond, it could be shown, that the main histotypes of lung cancer allow identification of numerous genes with known important function in embryonic lung development. Thus, a number of molecular similarities were found between genes expressed during the early stage of lung development and expressed in large cell carcinoma (LCC), while genes expressed by adenocarcinomas mirror those expressed during the later terminal sac and alveolar stages. Taken together, these observations suggest that poor differentiation is linked to molecular parameters of early development representing lung stem and progenitor cell programs, and that gene signatures of these phenotypes are important for lung cancer differentiation, progression, and clinical outcome.

Disruption of normal p53 gene function, usually by point mutations, is frequent in all types of lung cancers. Disruption of the RB gene pathway is universal in lung cancers. While mutations of the RB

gene are the usual method of disruption in SCLC, they are rare in NSCLC. In NSCLC the mechanism of disruption is via the upstream pathway. In particular, inactivation of p16Ink4 as demonstrated by immunohistochemistry occurs via epigenetic or genetic mechanisms.

Squamous Cell Carcinoma

Molecular genetic studies have shown that multiple genetic loci contribute to sporadic lung cancers. The molecular abnormalities are found in both growth-promoting oncogenes and growth-suppressing tumour suppressor genes. Squamous cell carcinoma commonly shows distinct molecular genetic characteristics. ErbB (EGFR, HER2/neu, KRAS) pathway abnormalities are common in non-small cell carcinoma, e.g. an average of 84% of squamous cell carcinomas are EGFR positive, while HER2/neu expression and KRAS mutations are relatively rare in SCCs. Most squamous cell carcinomas demonstrate large 3p segments of allelic loss, whereas most adenocarcinomas and preneoplastic/ preinvasive lesions have smaller chromosome areas of 3p allele loss.

Gene expression profiles of SCC can be helpful to differentiate SCC from other subtypes, in order to identify prognostic factors specific to SCC. Additionally it can be an aid in distinguishing primary pulmonary SCC from metastatic SCC of the head and neck (HNSCC). Genes encoding cytokeratins, e.g. cytokeratin 14, and cell adhesion proteins, e.g. components of desmosomes, hemi-desmosomes, and gap junctions, such as integrins ITGB4, desmocollin, and desmoplakin, are able to distinguish SCC from adenocarcinoma. In certain clinical situations primary SCC of the lung must be distinguished from metastatic HNSCC. Using multigene classifiers pulmonary SCC can be separated from SCC of the tongue or other head and neck primaries with an accuracy of up to96% accuracy.

Adenocarcinoma

Adenocarcinoma, the most frequent histologic subtype of NSCLC, is characterized by multiple somatic DNA alterations. Several chromosomal loci of lung adenocarcinoma specimens show alterations of copy numbers. The top focal regions of amplification and deletion included 14q13.3, 12q15, 8q24.21, 7p11.2, and 8q21.13. Additionally novel alterations such as amplification of the transcription factor TTF-1 on chromosome 14q13.3 were identified. It appears that these DNA amplification events are important in mediating lung cancer initiation, differentiation, and progression.

The genetic alterations in adenocarcinoma include point mutations of dominant oncogenes, such as the K-ras gene, and tumour suppressor genes such as p53 and p16Ink4. KRAS mutations occur in approximately 30% of adenocarcinomas. They are rare in other lung cancers, they are more common in cancers arising in smokers. Mutations result in constant downstream signalling resulting in proliferative stimuli. MYC, Cyclin D1 and EGFR15 are amplified and over-expressed in 2.5–10%, 5% and 6% of NSCLC, respectively. C-erbB2 (Her-2/neu) or BCL2 over-expression are involved in _25% of cases. Systematic resequencing of oncogenes identified novel mutations in, for example, BRAF, present in about 2% of adenocarcinoma patients and restricted to tumours that did not show KRAS mutations. More recently, mutations in the EGFR gene were detected, and the mutation status correlated with response to small molecule kinase inhibitors (e.g., gefitinib or erlotinib). A subset of NSCLC patients appears to express a transforming EML4-ALK fusion gene.

Mutations have also been described in the putative precursor lesion, atypical adenomatous hyperplasia. p53 mutations are also a negative prognostic factor for limited stage adenocarcinoma. Increased expression of p27, one of the cell cycle regulators, correlates with better tumour differentiation and more favourable prognosis. p16Ink4 inactivation by multiple mechanisms occurs frequently in adenocarcinomas and may be smoking related. LKB1/STK11, the gene responsible for Peutz-Jeghers syndrome, is reported to be frequently inactivated in adenocarcinoma of the lung.

Expression profiles can be used for subdividing lung carcinomas into several groups and for discrimination of primary cancers from metastases of extrapulmonary origin. The abnormal expression of genes involved in maintaining the mitotic spindle checkpoint and genomic stability contributes to the molecular pathogenesis and tumour progression of tobacco smoke-induced adenocarcinoma of the lung. Alterations in cell cycle genes have also been identified in lung adenocarcinoma using gene expression profiling. Gene expression profiles revealed by microarray analyses have been found to be

of prognostic significance in adenocarcinomas. High relative expression levels of surfactant protein genes and several other shared genes, including BENE, cytochrome b5, and elenium-binding protein 1, characterized adenocarcinomas as bronchioloalveolar carcinomas and appear to form a clear and distinct branch within the whole group of pulmonary adenocarcinomas. More recently, a risk index compiling the relative expression of 50 genes was developed to identify high or low risk groups of Stage I adenocarcinomas that correlated inversely with patient survival.

Recent research has focused on characterizing the molecular mechanisms and clinical implications of adenocarcinoma invasion, the initial step of metastasis. Paralleling malignancies in other organs, such as breast and cervix, where tumours are defined as non-invasive (in situ carcinoma), microinvasive (microscopic invasion), or as invasive carcinomas, the extent of the invasive component seen in lung adenocarcinoma is associated with clinical outcomes. The clinical importance of lung adenocarcinoma invasion is supported by several recent studies indicating that the risk of death in non-invasive BAC tumours is significantly lower than that of pure invasive tumours and in tumours with less than 0.6 cm of fibrosis or linear invasion. These results suggest that the prognosis of BAC is favourable, and suggest a similarly favourable prognosis for the subset of AC-Mixed subtype tumours with limited invasion (< 0,6 mm). Among the genes differentially expressed in the progression from BAC to invasive tumours was the type II transforming growth factor b receptor (TGFbRII), which was less highly expressed by AC-Mixed and solid invasive tumours compared with BAC Among potential mediators identified was the chemokine CCL5 (RANTES [regulated on activation, normal T cell expressed, and presumably secreted]), which is upregulated by invasive tumours and TGFbRII knockdown cells. RANTES is involved in immunoregulatory and inflammatory processes and is transcribed and secreted not only by T cells, other inflammatory cells, and stromal cells, but also by tumour cells and normal bronchial epithelium. Inhibition of RANTES signalling was found to be associated with abrogation of tumour invasion, suggesting that RANTES is required for invasion in TGFbRII repressed lung adenocarcinoma cells. The clinical significance of this pathway is further supported by the finding that tumour expression of both RANTES and CCR5 by a large panel of lung adenocarcinoma is associated with patient survival.

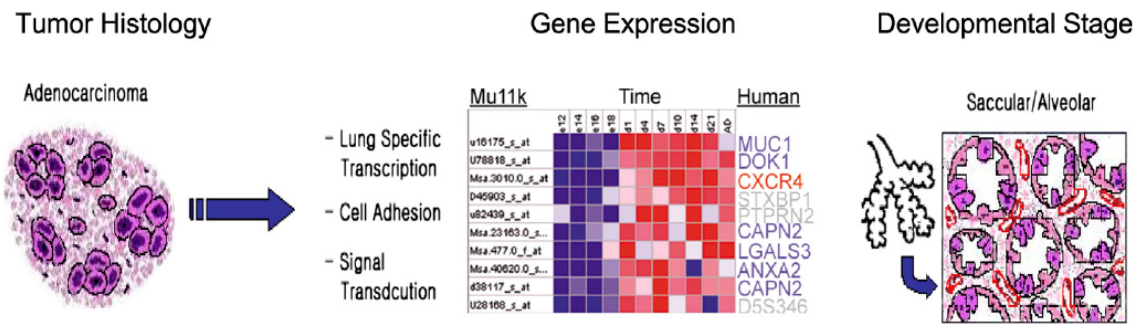

Figure 7.4: Gene expression of human lung adenocarcinoma with genes temporally expressed during lung development.

Large cell carcinomas

Cytogenetics shows that large cell carcinoma of the lung is mostly an aneuploid neoplasm with the highest mean chromosome number and DNA content of all lung cancer types being in the near triploid range or above. Accordingly, the karyotypes are complex and indicate a high chromosomal instability. The tumours harbour imbalances that have been associated with progression and metastasis formation, e.g. amplifications of 1q21-q22, 8q and deletion of 3p12-p14, 4p, 8p22-p23, 21q. On the molecular genetic level large cell carcinomas share the molecular and genetic alterations commonly seen in NSCLC, since it is a poorly differentiated tumour issued from the same stem cells, exposed to the same carcinogens.

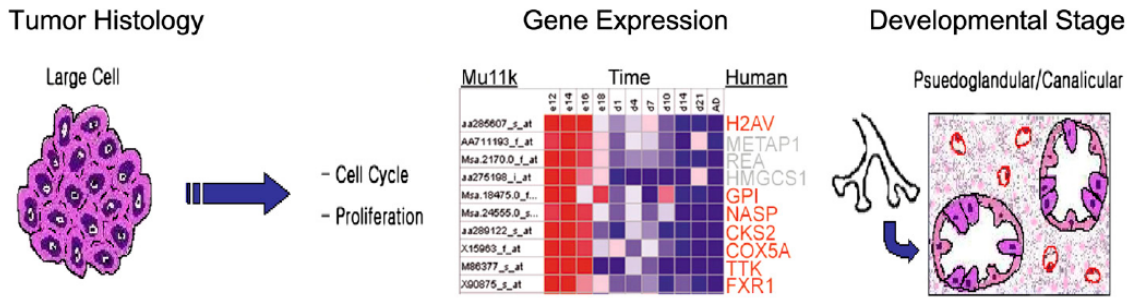

Figure 7.5: Gene expression of human lung large cell carcinoma with genes temporally expressed during lung development.

KRAS mutations, p53 mutations and Rb pathway alteration occur with the same frequency as in other NSCLC. Fas is downregulated but its ligand FasL is strongly upregulated.

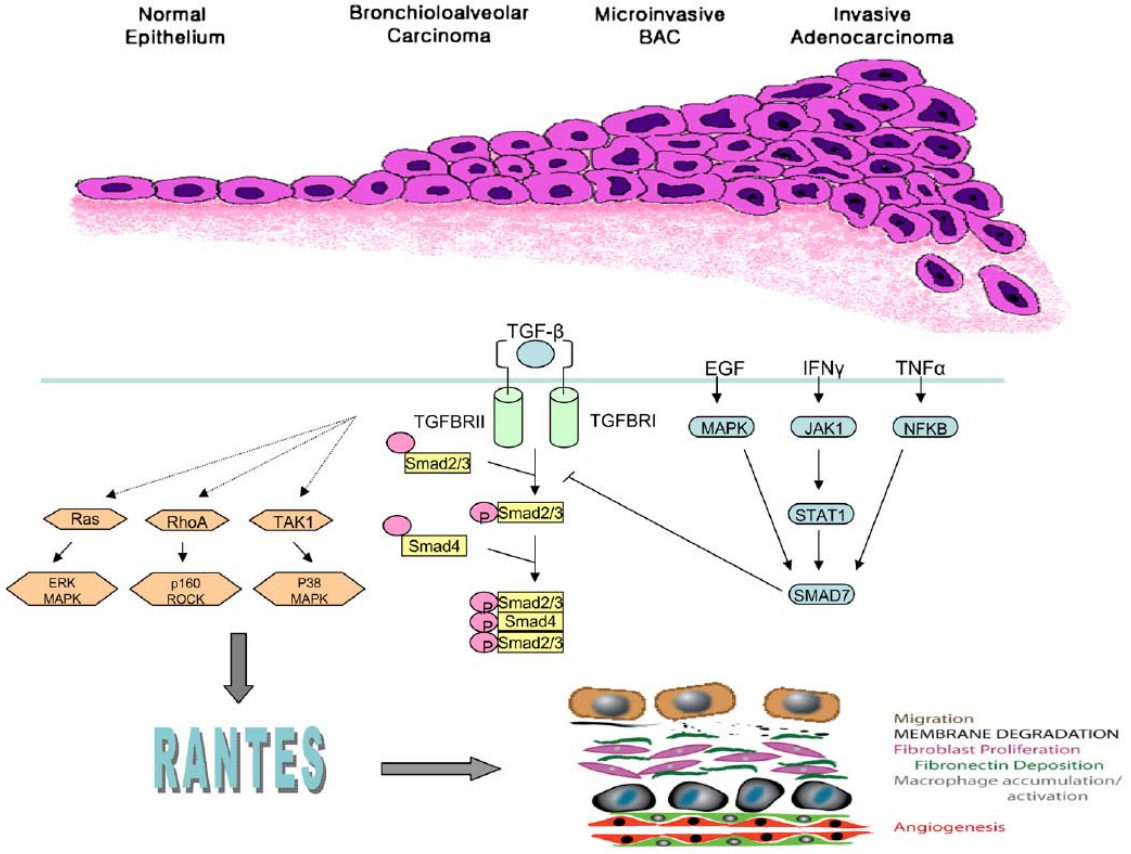

Figure 7.6: Classification of lung adenocarcinoma reflects progression along a spectrum of invasiveness spanning epithelial atypia, bronchioloalveolar adenocarcinoma (BAC), BAC with foci of microinvasion, and invasive adenocarcinoma. In cells with repression of TGFb-RII (lower panel), signaling through the Smad pathway is abrogated (right). Invasiveness is mediated through noncanonical signaling pathways that require RANTES (left).

Data used for the lung cancer model

In deliverable D4.1 chapter 5.2 the data used for the adaptation, optimization and validation of the models for lung cancer growth and response to treatment are listed. The following classification of data is done:

- B1. Treatment data
- B2. Normal tissue complication data

- B3. Clinical data
- B4. Imaging data
- B5. Segmentation data
- B6. Histopathological data
- B7. Molecular data

For details see D4.1 chapter 5.2. ObTiMA, developed by ACGT, will be used as data management system. For further details see chapter 8 of this deliverable. There are more items provided than given in D4.1. These tables are part of the prospective trial for lung cancer, in which more data are collected. Molecular data will not be stored on CRFs in ObTiMA. They will be provided in a repository for genetic data after pseudonymisation and annotation to the Gene Ontology. Segmentation data are stored in an imaging repository after pseudonymisation.

Scenario and Workflow of the lung cancer model

The workflow of the lung cancer model is given in figures 7.7 and 7.8. At the time of diagnosis clinical data, imaging data and histological data will be collected and pseudonymized. Segmentation of the tumour will be done using the segmentation tool described in chapter 4. Autoantibodies against lung cancer antigens and EGFR sequencing data will be analyzed. All these data will be used for the *in silico model*. The model will simulate the tumour response of the given treatment of the patient. A correlation between histology and imaging data and molecular data will be done in several cases to fuse the histological picture with the imaging data of MRI. The correlation to molecular biology to histology and imaging is possible if the origin of the specimen from the tumour is known. During treatment of the patient toxicity data are collected and used for the prediction of side effects on normal tissue.

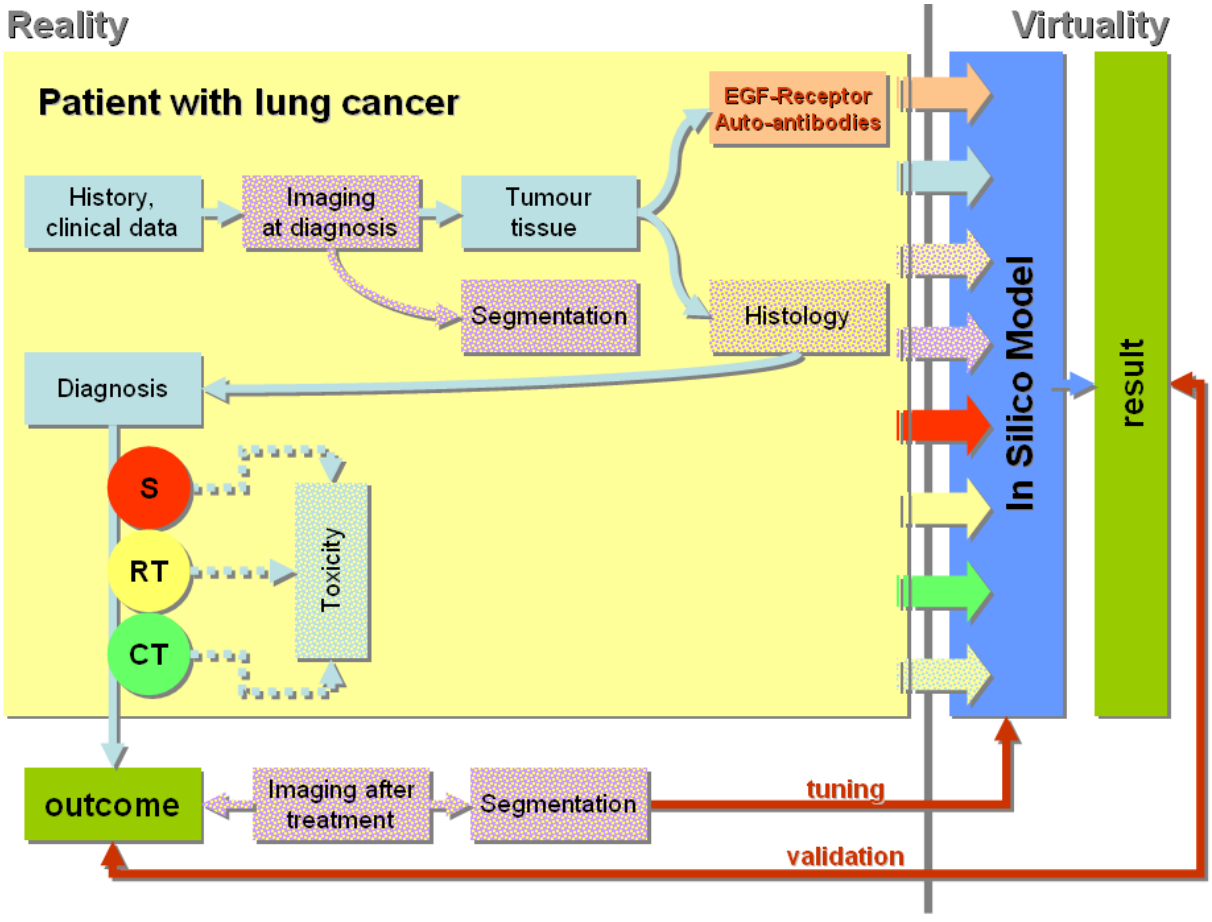

Figure 7.7: Workflow of the lung cancer scenario. S: surgery; RT: radiotherapy; CT: chemotherapy.

Validation of the lung cancer model

At the end of treatment the outcome of the patient regarding tumour volume will be used for validation of the model. The tumour of the patient will be segmented and the tumour volume will be calculated. The difference between the real tumour volume found in imaging studies of the patient and the simulated tumour volume has to tune the *in silico model*. Differences greater than 10 % between the real tumour volume and the simulated volume are unacceptable. In this case optimization of the model is needed. The model can be considered as helpful for clinicians if in at least 50 patients a correct result (less than 10 % deviation between reality and virtuality) is predicted. In this situation the model might be tested as a decision making tool for patients with lung cancer in prospective clinical trials.

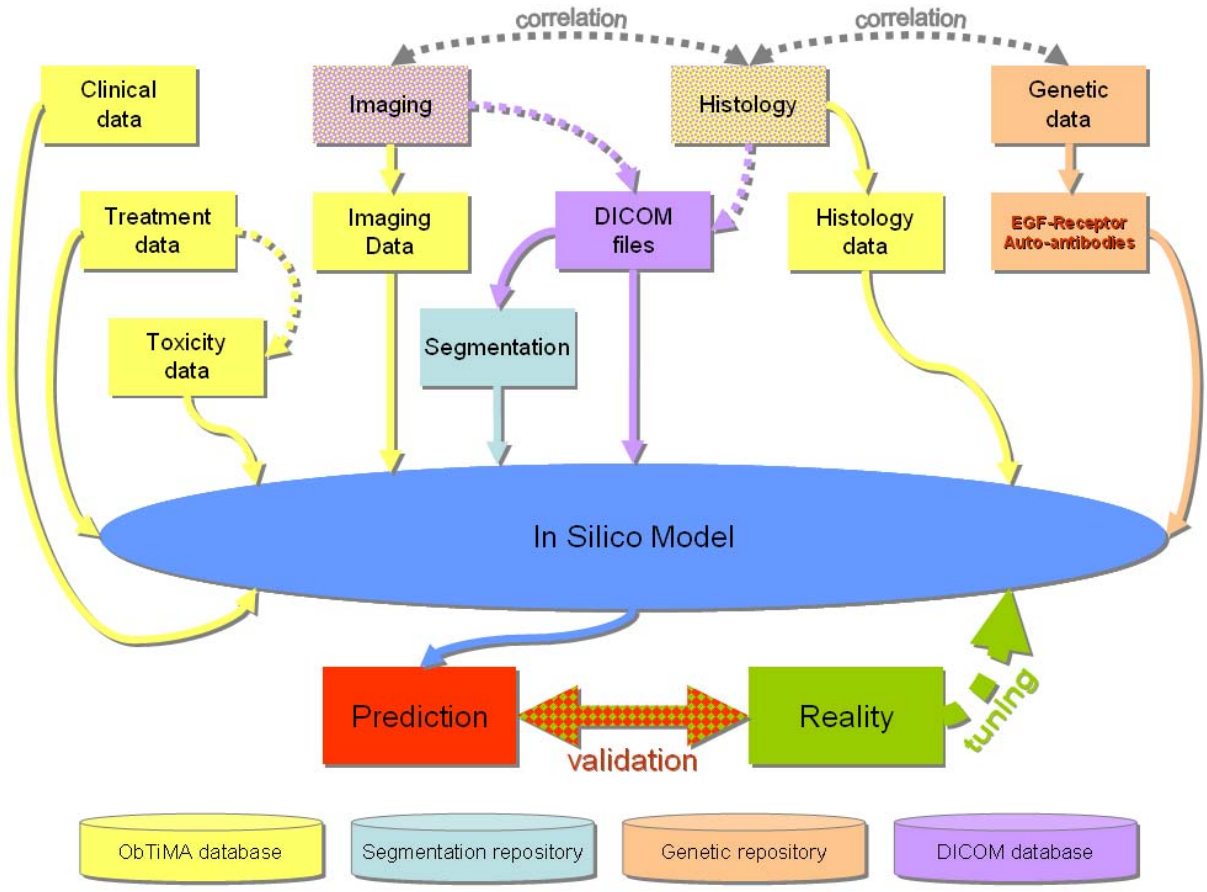

Figure 7.8: Workflow of data of the lung cancer scenario.

ContraCancrum clinical data

8. Clinical and imaging database

In the following chapter the clinical and imaging database is described that will be used for the management of data in the scenarios and trials of ContraCancrum. ObTiMA[22] will be the tool that will be used for that purpose. The ACGT MO[23] will be expanded to lung cancer and glioma.

The basic dataset includes the data that are needed for each trial. These datasets belong to the basic module of a clinico-genomic trial and is an adaptation of the Dataset used in ACGT[21].

Description

The following tables include different categories. In the first column the context is described. This defines the header for the category in the second column. All the categories are directly linked to the context. The type describes the format of the data. It is divided into the following data types:

- A: alphanumeric
- I: Integer
- R: real number
- D: Date
- M: Memo

The number indicates the number of characters of the item.

If the item is coded the column Format/Coding/Options shows the description of the items that can be collected. If in the following tables ”done” is written, it mean that people just have to say if the specific item was done or not. The Column for the Ontology Reference is left empty in the following tables but indicates that each category will be matched to the Ontologies if ObTiMA is used.

Clinical database

8.1.1 Personal data

All personal data will be anonymized before used in ContraCancrum. Only the treating physician has the right to see personal data.

Person identifying data

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Personal Data	Family Name	A50		
	First Name	A50		
	Birth Name	A50		
	Title	I2	-1: not yet known 0: no title 1: Mr. 2: Mrs. 3: Dr. 4: Prof.	

[22] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. 15th September 2007

[23] http://bioportal.bioontology.org/ontologies/38787 (last accessed 13th April 2009)

Gender		I2	-1: not yet known 1: male; 2: female; 3: ambiguous;	
Date of birth		D8	DDMMYYYY	
of birth	Country	I2	ISO 3166-1-alpha-2 code	
	ZIP Code	A10		
	City	A50		

Valid home address of Patient

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Valid home address	Street	A50		
	ZIP Code	A10		
	City	A50		
	Country	I2	ISO 3166-1-alpha-2 code	
Email		A50	Including @	
Phone number	Number	A25		
	Description	I2	-1: not yet specified 1: Main line 2: Fax 3: Mobile	

Family data

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Mother	Date of birth	D4	YYYY	
Father	Date of birth	D4	YYYY	
Patient	Number of siblings	I2		
Multiples	Multiples	I2	-1: not yet known 1: no 2: yes	
	Type	I2	-1: not yet known 1: no 2: twin 3: triplets 4: more	
	Genetics	I2	-1: not yet known 1: monozygotic 2: dizygotic	

8.1.2 Common basic data

Identification Numbers of Patient

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Patient-ID (PID) in running trial	ID	A8		
2nd PID	ID	A8		
Identification number of the treating hospital	ID	A25		
Identification number in the trial	ID	A25		

Address of Hospital / Institution / Laboratory / Company

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Hospital / Institution	Name		A50		
Institution ID	ID		I5		
Type of Hospital / Institution / Laboratory / Company	University		I2	-1: not yet known 1: no 2: yes	
	Type		I2	-1: not yet known 1: University Hospital 2: Specialized Hospital 3: General Hospital 4: Research Institute 5: IT Institute 6: legal/ethics Institute 7: Laboratory 8: EC/IRB 9: Regulatory Body 10: Pharma. Company 11: IT Company 12: other	
		If other, specify	A50		
Actual used Information System	Name		A25		
Director	ID		I5	User-ID [2.4.3.3]	
Responsible IT person	ID		I5	User-ID [2.4.3.3]	
Address	Street		A50		
	ZIP Code		A10		

	City	A50		
	Country	A50		
		I2	ISO 3166-1-alpha-2 code	
Homepage		A50	http://	
Email		A50	Including @	
Phone	Number	A25		
	Description		-1: not yet specified 1: Main line 2: Fax 3: Mobile	

Data of User

Context	Category	Type	Format/ Coding/Options	Ontology Reference
User ID	ID	I5		
Role in the trial	ID	I2	-1: not yet known 0: Family Doctor/general Practitioner 1: local Physician 2: Data Manager 3: Chairman of the trial / PI 4: Reference Physician 5: Statistician 6: Study Nurse 7: Ethicist 8: Jurist / Lawyer 9: Basic Researcher 10: Computer Scientist 11: Member of DSMC 12: Member of Ethical Committee 13: Member of regulatory body 14: Member of Insurance company 15: Member of Pension Office 16: other	
	If other, specify	A50		
Speciality	ID	I2	-1: not yet known 0: not further specified 1: Adult Oncologist 2: Paediatric Oncologist 3: Radiotherapist 4: General Surgeon 5: Paediatric Surgeon 6: Orthopaedic Surgeon 7: Neurosurgeon 8: Heart Surgeon 9: Pathologist	

			10: Radiologist 11: Molecular Biologist 12: Geneticist 13: Gynaecologist 14: Urologist 15: Dermatologist 16: Paediatrician 17: Internist 12: Otorhinolaryngologist 13: Neurologist 14: Bioinformatician 15: Physicist 16: Philosopher/Ontologist 17: other	
	If other, specify	A50		
Master data of treating physician	Name	A50		
	First Name	A50		
	title	I2	-1: not yet known 0: no title 1: Mr. 2: Mrs. 3: Dr. 4: Prof. 5: MD 6: PhD 7: MD/PhD 8: other	
	If other: specify	A10		
Address	ID	I5	Institution-ID	
	Street	A50		
	ZIP Code	A10		
	City	A50		
	Country	I2	ISO 3166-1-alpha-2 code	
Email		A50	Including @	
Phone	Number	A25		
	Description	I2	-1: not yet specified 1: Main line 2: Fax 3: Mobile	

Admission to the treating hospital

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Admission	Date	D8	DDMMYYYY	
	with relapse	I2	-1: not yet known 1: No, primary disease 2: Yes	
ID of the institution	ID	I5		
Treatment in the hospital	Function	I2	-1: not yet known 1: Primary Therapy 2: after primary therapy in the home country 3: After treatment in a foreign country	

Family history

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Siblings- Prior diseases (multiple entries are possible)	Type	I2	-1: not yet known 1: malignant disease 2: Immunodeficiency 3: haematological disease 4: chronic viral infection 5: chromosomal Aberration 6: syndrome 7: other	
Mother- Prior diseases (multiple entries are possible)	Type	I2	-1: not yet known 1: malignant disease 2: Immunodeficiency 3: haematological disease 4: chronic viral infection 5: chromosomal Aberration 6: syndrome 7: other	
Father- Prior diseases (multiple entries are possible)	Type	I2	-1: not yet known 1: malignant disease 2: Immunodeficiency 3: haematological disease 4: chronic viral infection 5: chromosomal Aberration 6: syndrome 7: other	
Other family members- Prior diseases (multiple entries are	Type	I2	-1: not yet known 1: malignant disease 2: Immunodeficiency 3: haematological disease	

possible)			4: chronic viral infection 5: chromosomal Aberration 6: syndrome 7: other	

Actual history

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Actual history	Reason for diagnosis	I2	-1: not yet known 1: symptoms related to malignant disease 2: Preventive examination 3: accidentally, without symptoms	
Address at the time of diagnosis	Street	A50		
	ZIP Code	A10		
	City	A50		
	Country	I2	ISO 3166-1-alpha-2 code	

Clinical Status at time of diagnosis

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Karnofsky Index	Code	I1		
	Text	A254		

Case history

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Prior diseases	Type	I1	Multiple answers possible	
	Begin	D8		
	End	D8		
Diagnosis	Text	A254		
Diagnostic key	Code	A20	ICD - 10	
	Text	A254		

Actual diagnosis

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Main diagnosis	Text	A254		
	Date of diagnosis	D8	DDMMYYYY	
ICD-O-3 Morphology	Code	A20		
	Text	A254		
ICD-O-3 Topography	Code	A20		
	Text	A254		
ICD-10	Code	A20		
	Text	A254		

Concomitant diseases at time of primary diagnosis

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Main diagnosis	Text	A254		
	Date of diagnosis	D8	DDMMYYYY	
ICD-O-3 Morphology	Code	A20		
	Text	A254		
ICD-O-3 Topography	Code	A20		
	Text	A254		
ICD-10	Code	A20		
	Text	A254		

Diagnosis for trial inclusion

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Diagnosis	Date of diagnosis	D8	DDMMYYYY	
	Text	A254		

	Diagnosis confirmed by	I2	-1: not yet defined 1: clinical – no imaging 2: Specific diagnosis – with imaging 3: cytological 4: histological 5: autopsy	
ICD-O-3 Morphology	Code	A20		
	Text	A254		
ICD-O-3 Topography	Code	A20		
	Text	A254		
ICD-10	Code	A20		
	Text	A254		
Microphotographs	Picture provided	I2	-1: not yet known 1: no 2: yes	
	file stored at	A50		
Description	Text	Memo		
	Amount of dead cells	I3	%	
	Amount of proliferating cells (Ki67)	I3	%	

Confirmation of the diagnosis by reference institution

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Examination for confirmation of diagnosis	done	I2	-1: not yet known 1: no 2: yes	
Findings for confirmation of diagnosis	Status	I2	-1: not yet known 1: not available 2: available	
Available Findings for confirmation of diagnosis	Diagnosis confirmed	I2	-1: not yet known 1: no 2: yes	
Examination institute	Function	I2	-1: not yet known 1: no reference center 2: reference center	
Main data of the institution	Institution ID	I5		
	Name	A50		

	Street	A50		
	Postal code	A10		
	City	A50		
	Country	A50		
Email		A50	Including @	
Phone	Number	A25		
	Description	I2	-1: not yet specified 1: Main line 2: Fax 3: Mobile	

8.1.3 Basic trial data

Trial and possible amendments

Context		Category	Type	Format/ Coding/Options	Ontology Reference
Trial		Name	A254		
		Abbreviation	A20	Abbreviation of the Trial	
		ContraCancrum Number	I5		
		Characteristics	I2	-1: not yet known 1: treatment trial 2: preventive trial 3: diagnostic trial 4: screening trial 5: quality of life trial 6: epidemiological trial	
			I2	-1: not yet known 1: prospective 2: retrospective	
			I2	-1: not yet known 1: randomized 2: non-randomized	
			I2	-1: not yet known 1: unicenter 2: multicenter	
EudraCT No.		Number	A14	YYYY-NNNNNN-CC	
EC/IRB	Name	Institution ID	I5		
	Approval	Done	I2	-1: not yet known 1: no 2: yes	
	Date of approval	Date	D8	DDMMYYYY	

Regulatory Body	Name	Institution ID	I5			
	Approval	Done	I2	-1: not yet known 1: no 2: yes		
	Date of approval	Date	D8	DDMMYYYY		
Insurance	Name	Institution ID	I5			
	Approval	Done	I2	-1: not yet known 1: no 2: yes		
	Date of approval	Date	D8	DDMMYYYY		
	No.	Number	A50			
Funding	Name	Institution ID	I5			
	Approval	Done	I2	-1: not yet known 1: no 2: yes		
	Date of approval	Date	D8	DDMMYYYY		
	No.	Number	A50			
Sponsor		User ID	I5			
Chairman / PI		User ID	I5			
Deputy Chairman		User ID	I5			
Data Manager		User ID	I5			
Statistical centre		Institution ID	I5			
Statistician		User ID	I5			
Study committee		User ID	I5		multiple entries possible	
Reference Centres		Institution ID	I5			
		Contract done	I2	-1: not yet done 1: no 2: yes		
		Date of contract	D8	DDMMYYYY		
DSMC		User ID	I5			
Dates	Start of trial	Date	D8			
	patient recruitment – Start	Date	D8			
	patient recruitment – Closure	Date	D8			
	End of trial	Date	D8			

Duration of follow-up	Timeframe		I3	in months	
Sample size	Number		I4		
Statistical analysis	Date		D8	multiple entries possible	
	Number of patients		I5		
	Number of events		I5		
Monitoring of the trial	Done		I2	-1: not yet done 1: no 2: yes	
	If yes: Date		D8	Multiple entries possible	
	Monitor at that date Institution-ID		I5	Multiple entries possible	
	Monitor(s) at that date User-ID		I5	Multiple entries possible	
	Protocol of the visit at that date		M	Memo Multiple entries possible	
Publication rules	Done		I2	-1: not yet done 1: no 2: yes	
	If yes, Text		M	Memo	
Amendments	Done		I2	-1: not yet done 1: no 2: yes	
	If yes, please specify	Name of amendment	A50	multiple entries possible	
		Date	D8		

Basic data of the patient in the trial

Context	Category	Type	Format/ Coding/Options	Ontology Reference
ContraCancrum No.	ID	I5		
ID in running trial	ID	A8		
Inclusion into the trial	Status	I2	-1: not yet known 1: no 2: yes	
Informed consent	Given by parents	I8	-1: not yet known 1: no 2: yes	
	Given by patient	I8	-1: not yet known 1: no 2: yes	
	Status	I2	-1: not yet known 1: protocol patient 2: trial patient 3: study patient	

Protocol patient: treated according to protocol
Trial patient: randomised patient
Study patient: registered in the study, but only observed

Stratification

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Stratification	Date	D8	DDMMYYYY	
Risk group	Name	I1	Risk group	
Treatment arm	Name	I1	Treatment arm	

Randomisation

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Randomisation possible	Patient can be randomised	I2	-1: not yet known 1: no 2: yes	
Randomisation	Done	I2	-1: not yet known 1: no 2: yes	
Date of Randomisation	Date	D8	DDMMYYYY	

Not Randomised	Reason		I2	-1: not yet defined 1: no Informed consent 2: Refusal by hospital 3: Medical reasons 4: organisational reasons	
		If medical or organisational reason, specify	A254		
Treatment arm assigned by randomisation	Name		I1	Treatment arm	
Treatment arm assigned without randomisation	Name		I1	Treatment arm	
Treatment arm given	Name		I1	Treatment arm	

8.1.4 Basic Treatment data

Previous treatment of the main disease

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Previous therapy	Institution-ID [2.4.3.2]		I5	[where treatment was done]	
	User-ID [2.4.3.3]		I5	[responsible physician]	
	Done		I2	-1: not yet known 1: no 2: yes	
Treatment block	Begin		D8	DDMMYYYY	
	End		D8	DDMMYYYY	
	Type		I2	-1: not yet known 1: Chemotherapy 2: Radiotherapy 3: Surgery 4: Stem cell-transplantation 5: Other Therapy	
		If other please specify	A254		

Treatment after the beginning of Protocol therapy

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Chemotherapy	Done	I2	-1: not yet known 1: no 2: yes	
Radiotherapy	Done	I2	-1: not yet known 1: no 2: yes	
Surgery	Done	I2	-1: not yet known 1: no 2: yes	
Stem cell-transplantation	Done	I2	-1: not yet known 1: no 2: yes	
Other treatment	Done	I2	-1: not yet known 1: no 2: yes	

Protocol treatment

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Protocol therapy	Institution-ID		I5	[where treatment is given]	
	User-ID		I5	[responsible physician]	
	Begin		D8	DDMMYYYY	
	End		D8	DDMMYYYY	
	First kind of treatment		I2	-1: not yet known 1: Chemotherapy 2: Radiotherapy 3: Surgery 4: Stem cell-transplantation 5: other Therapy	
		If other, please specify	A254		
Treatment block	Name		I2	Scheme	

Physical examination at begin of therapy

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Patients' Height	Result	I3	cm	
Patients' Weight	Result	R3,2	kg	
Patients' Body surface area	Result	R1,2	m^2	

8.1.5 Chemotherapy and other tumour specific treatments

Recommended / Planned Treatment

Context	Category	Type	Format/ Coding/Options	Ontology Reference Number
Recommended / Planned Treatment block	Begin Date	D8	DDMMYYYY	
	End Date	D8	DDMMYYYY	

Therapeutic agents in planned (chemo-) therapy

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Chemotherapeutic agents or other drugs	Application		I2	-1: not yet known 1: intravenous (i.v.) 2: per infusionem (p.i.) 3: intrathecal (i.th.) 4: oral (p.o.) 5: intramuscular (i.m.) 6: subcutaneous (s.c.) 7: rectal 9: other application	
		If other, please specify	A50		
	Date		D8	DDMMYYYY	
	Time		D4	HH:MM	
	Duration, if per infusionem		R2.1	hours	
	Dose	Amount	R8.2		
		unit	I2	-1: not yet known 1: mg 2: units	
	Active component		I5	Drug ID 2.4.6.3	

Table of Medical Drugs

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Medical drug	Drug-ID		I5		
	Generic Name		A50		
	Code		A10		
	Type		I2	-1: not yet known 1: chemother. agent 2: immune modulator 3: Supportiva 4: other	
		If other, specify	A25		

Administered chemotherapy

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Administered chemotherapy	Begin	D8	DDMMYYYY	
	End	D8	DDMMYYYY	

Therapeutic agents in administered (chemo-) therapy

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Chemotherapeutic agents or other drugs	Application		I2	-1: not yet known 1: intravenous (i.v.) 2: per infusionem (p.i.) 3: intrathecal (i.th.) 4: oral (p.o.) 5: intramuscular (i.m.) 6: subcutaneous (s.c.) 7: rectal 9: other application	
		If other, specify	A50		
	Date		D8	DDMMYYYY	
	Time		D4	HH:MM	
	Duration given		R2.1	hours	
	Dose	Amount	R8.2		
		unit	I2	-1: not yet known 1: mg 2: units	
	Active component		I5	Drug ID	
	Pharmacokinetics[24]	Peak serum level	R6.2		
		Central volume (Vc)	R6.2		
		Tissue volume (Vt)	R6.2		
		Volume distribution (Vd)	R6.2		
		Clearance	R6.2		
		Elimination rate constant (Kel)	R6.2		
		Half-life (t1/2)	R6.2		
		AUC	R6.2		

[24] For calculation of pharmacokinetics: http://www.rxkinetics.com/pktutorial/1_1.html

8.1.6 Surgery

Surgical treatment

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Surgical treatment	Institution-ID [2.4.3.2]		I5		
	Main Surgeon User-ID [2.4.3.3]		I5		
	Date		D8	DDMMYYYY	
	Description		A254		
	Complications	Description	A254		
		MedDRA	I10	MedDRA Code	
Coding of surgical procedure according	Code		A10	OPS Version 2007 LOINC®	
(multiple entries possible)	Code Text		A254		

8.1.7 Radiotherapy

Administered irradiation

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Administered irradiation	Begin	D8	DDMMYYYY	
	End	D8	DDMMYYYY	
Type of irradiation	Code-1	I2	-1: not yet known 1: Teletherapy 2: Brachytherapy 3: Radioisotope therapy	
	Code-2	I2	-1: not yet known 1: conventional 2: conformal 3	
	Code-3	I2	-1: not yet known 1: conventional 2: hyperfractionated 3: hyperfractionated accelerated 4: Proton Therapy 5: Electron Therapy 9: other	

Radiation field

Context	Category			Type	Format/ Coding/Options	Ontology Reference
Radiation field (for multiple fields including boost filed multiple records are possible)	Institution-ID [2.4.3.2]			I5		
	Main Radiotherapist User-ID [2.4.3.3]			I5		
	Begin			D8	DDMMYYYY	
	End			D8	DDMMYYYY	
	Localisation of the field			A254	Description of the field	
	side			I2	-1: not yet known 0: not applicable 1: left; 2: right 3: bilateral; 4: midline	
	size	anterior	length	R2.1	cm	
			width	R2.1	cm	
		posterior	length	R2.1	cm	
			width	R2.1	cm	
	kind of irradiation			I3	-1: not yet known 1: percutaneous 2: Brachytherapy 3: Radionuclide 4: Stereotactic 5: other radiation	
		If other, please specify		A50		
	No. of irradiation days			I2		
Administered irradiation Dosage	Overall dose			R2.1	Gy	
	Single dose			R1.1	Gy	
	Number of single doses			I3		
Percutaneous radiation of a field	Kind of radiation			I2	-1: not yet known 1: Photons 2: Electrons 3: Neutrons 4: Protons 5: other	
		If other, please specify		A50		
Brachytherapy	Application			I2	-1: not yet known 1: Interstitial 2: Intracavity 3: Seed Implantation 4: Flab Radiation 5: other application	
		If other, please specify		A50		

8.1.8 Course of primary tumour

Evaluation of primary tumour response

Context	Category	Type	Format/ Coding/Options	Ontology Reference
PID in running trial	ID	A8		
Evaluation of response (multiple entries possible)	Date	D8	DDMMYYYY	
	Result	I2	-1: not yet known 1: Complete Remission 2: Partial Remission 3: No Changes 4: Progressive Disease 5: not determinable	
	Method	I2	-1: not yet defined 1: clinical – no imaging 2: Specific diagnosis – with imaging 3: cytological 4: histological 5: autopsy	

Relapse

Context	Category				Type	Format/ Coding/Options	Ontology Reference
PID in running trial	ID				A8		
Relapse (multiple entries possible)	occurrence				I2	-1: not yet known 1: no 2: yes	
	occurrence	If yes, specify	kind		I2	-1: not yet known 1: local relapse 2: metastasis 3: combined relapse 4: systemic relapse	
	occurrence	If yes, specify	side	Code	A20	ICD-O-3 topographic (multiple entries possible)	
	occurrence	If yes, specify	side	Text	A254		
	Date				D8	DDMMYYYY	

8.1.9 Second malignancy

Context	Category	Type	Format/ Coding/Options	Ontology Reference
PID in running trial	ID	A8		
Second malignancy	Code	I2	-1: not yet known 1: no 2: yes	
Main diagnosis	Text	A254		
	Date of diagnosis	D8	DDMMYYYY	
ICD-O-3 Morphology	Code	A20		
	Text	A254		
ICD-O-3 Topography	Code	A20		
	Text	A254		
ICD-10	Code	A20		
	Text	A254		

8.1.10 Late effects

Context	Category	Type	Format/ Coding/Options	Ontology Reference
PID in running trial	ID	A8		
Late effect	Category	I2	Category list	
	Infection related		-1: not yet known 1: no 2: yes	
Diagnosis	Date	D8	DDMMYYYY	
	Text	A254		
	Comment	A254		
	ICD-10 Code	A20		
	ICD-10 Text	A254		
Causality	Related to the diagnosis of the patient	I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	

	Treatment related		I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	
		If treatment related, specify	I2	-1: not yet known 1: chemotherapy 2: irradiation 3: surgery 4: infection 5: 2nd malignancy 6: other	
		If other, specify	A25		
Classification			I2	-1: not yet known 1: death 2: life threatening 3: unplanned or prolonged Hospitalisation 4: persistent or significant disability/ incapacity	
General Karnofsky Score	Code		I1		
	Text		A50		
MedDRA	Code		I8		
	Term		A254		
CTCAE v3.0	Grade		I1		
	Text		A254		

8.1.11 Death of patient

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Patient	Dead	I2	-1: not yet known 1: no 2: yes	
	Date	D8	DDMMYYYY	
Course of disease	Status	I2	-1: not yet known 1: before treatment 2: before 1st CR 3: in 1st CR 4: after relapse 5: after 2nd Malignancy	
Autopsy	Done	I2	-1: not yet known 1: no 2: yes	

Cause of death

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Cause of death		I2	-1: not yet known 0: not determinable 1: treatment 2: malignant disease 3: Infection 4: Accident 5: Suicide 6: others	
	If other, specify	A25		
Causality	If related to the malignant disease	I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	
	If treatment related	I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	
	If treatment related – specify	I2	-1: not yet known 1: chemotherapy 2: irradiation 3: surgery 4: infection 5: other	
	If treatment related – If other, specify	A25		
	Cause as result of secondary malignancy	I2	-1: not yet known 1: no 2: yes	
Diagnosis Cause of death	Text	A 254		
	ICD-10 Code	A20		
	ICD-10 Text	A 254		
Autopsy	Done	I2	-1: not yet known 1: no 2: yes	
	Cause of death confirmed by autopsy	I2	-1: not yet known 1: no 2: yes	
	Diagnosis of autopsy – Text / Protocol	M	Memo	
	Diagnosis of autopsy – ICD-10 Code	A20		
	Diagnosis of autopsy – ICD-10 Text	A 254		

8.1.12 State of documentation

Context	Category	Type	Format/ Coding/Options	Ontology Reference
Patient	Lost to follow up	I2	-1: not yet known 1: no 2: yes	
Last follow-up done	Date	D8	DDMMYYYY	
Last information about the patient	Date	D8	DDMMYYYY	
CRFs (multiple entries possible)	Name of CRF	A50		
	Code of CRF	I10		
	Status	I2	-1: not yet known 1: no data 2: incomplete data 3: complete data not verified 4: complete data verified on side 5: complete data verified technically 6: complete data verified medically	
	Repository code	I10		

Dataset for SAE and SUSAR Reporting

The 'ACGT Trial Builder'[25] will use the Common Terminology Criteria for Adverse Events (CTCAE v3.0) of the NCI (http://www.fda.gov/cder/cancer/toxicityframe.htm), which is a descriptive terminology. It will be utilized for Adverse Event (AE) reporting. The main categories and the grading of version 3.0 are shown in the following table. CTCAE v3.0 grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life-threatening and death. This grading system inherently places a value on the importance of an event, although there is not necessarily 'proportionality' among grades (a Grade 2 is not necessarily twice as bad as a Grade 1). Some adverse events (AEs) are difficult to 'fit' into this 5-point schema, but altering the general guidelines of severity scaling would render the system useless for comparing results between trials, an important purpose of the system. All grades are not appropriate for all AEs.

The International Conference on Harmonization (ICH) develops requirements for drug regulatory reporting globally and has chosen MedDRA (Medical Dictionary for Regulatory Activities) terminology to be the standard. To facilitate data transfer to regulatory agencies, CTCAE v3.0 terms are mapped to the current MedDRA version. The mapping can be found on the CTEP (Cancer Therapy Evaluation Program) Homepage (http://ctep.cancer.gov/). A PDF file describing the implementation in a PDMS is found at: http://ctep.cancer.gov/forms/CTCAE_Imp.pdf.

Definitions of Different Types of Adverse Events

An AE is any untoward medical occurrence in a clinical investigation patient administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavourable and unintended sign, finding, symptom, syndrome, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product. Laboratory abnormalities and changes in vital signs, 12-lead ECG, and telemetry are considered AEs only if they result in withdrawal from the study, necessitate therapeutic intervention, and/or the investigator considers them to be AEs.

SAEs include any untoward medical occurrence that at any dose:

- results in death
- is life-threatening [A]
- requires hospitalization or prolongation of existing hospitalization [B]
- results in persistent or significant disability or incapacity
- is a congenital anomaly/birth defect, or
- is another medically important condition [C]

[A] The term "life-threatening" in the definition of "serious" refers to an event in which the patient is at risk of death at the time of the event; it does not refer to an event, which hypothetically might have caused death if it were more severe.

[B] Hospitalization for convenience does not constitute an SAE.

[C] Medically important conditions that may not result in death, be life-threatening, or require hospitalization may be considered as SAEs when, based upon appropriate medical judgment, they may jeopardize the patient or may require intervention to prevent one of the outcomes listed in the definition above. Examples of such events are intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias, or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

[25] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. 15th September 2007

Category	Grading					
	0	1	2	3	4	5
Allergy/Immunology	No Adverse Event (absent) or within normal limits	Mild Adverse Event (minor; no specific medical intervention; asymptomatic laboratory findings only, radiographic findings only; marginal clinical relevance)	Moderate Adverse Event (minimal intervention; local intervention; noninvasive intervention)	Severe and undesirable Adverse Event (significant symptoms requiring hospitalization or invasive intervention; transfusion; elective interventional radiological procedure; therapeutic endoscopy or operation)	Life-threatening or disabling Adverse Event (complicated by acute, life-threatening metabolic or cardiovascular complications such as circulatory failure, haemorrhage, sepsis. Life-threatening physiologic consequences; need for intensive care or emergent invasive procedure; emergent interventional radiological procedure, therapeutic endoscopy or operation)	Death related to Adverse Event
Auditory/Ear						
Blood/Bone Marrow						
Cardiac Arrhythmia						
Cardiac General						
Coagulation						
Constitutional Symptoms						
Death						
Dermatology/Skin						
Endocrine						
Gastrointestinal						
Growth and development						
Haemorrhage/Bleeding						
Hepatobiliar/Pancreas						
Infection						
Lymphatics						
Metabolic/Laboratory						
Musculoskeletal/Soft Tissue						
Neurology						
Ocular/Visual						
Pain						
Pulmonary/upper respiratory						
Renal/Genitourinary						
Secondary Malignancy						
Sexual / Reproductive function						
Surgery/Intra-Operative Injury						
Syndromes						
Vascular						

Please note: The term "severe" is often used to describe the intensity (severity) of an event (as in mild, moderate, or severe pain); the event itself may be of relatively minor medical significance (such as severe headache). This use of “severe” is not the same as "serious," which is based on patient/event outcome or action criteria usually associated with events that pose a threat to the patient’s life or vital functions. Seriousness (not severity) serves as a guide for defining regulatory reporting obligations.

Toxicity Table

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Toxicity	Begin		D8	DDMMYYYY	
	End		D8	DDMMYYYY	
Description of Toxicity	Text		M	Memo	
Comment	Text		M	Memo	
Causality	Related to the diagnosis of the patient		I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	
	Treatment related		I2	-1: not yet known 1: definitely 2: probable 3: possible 4: unlikely 5: unrelated	
		If treatment related, specify	I2	-1: not yet known 1: chemotherapy 2: irradiation 3: surgery 4: infection 5: 2nd malignancy 6: other	
		If other, specify	A25		
Classification			I2	-1: not yet known 1: death 2: life threatening 3: unplanned or prolonged Hospitalisation 4: persistent or significant disability / incapacity 5: congenital anomaly/birth defect 6: another medically important condition	
General Karnofsky Score	Code		I1		
	Text		A50		
MedDRA	Code		I8		
	Term		A254		
CTCAE v3.0	Grade		I1		
	Text		A254		

Imaging database

A dataset for imaging studies can be divided into different parts[26]:

- Dataset for Teleradiology
- Dataset for imaging study done
- Dataset for diagnosis done by imaging studies

In the following these datasets are described. These datasets list data that are useful in ACGT trials using imaging studies.

Dataset for Teleradiology

Context		Category	Type	Format/ Coding/Options	Ontology Reference
Imaging study ID		ID	I15		
Institution ID sending images		Institution-ID	I5		
Institution ID receiving images		Institution-ID	I5		
Imaging study	Date send	Date	D8	DDMMYYYY	
	Anonymised	Done	I2	-1: not yet known 1: no 2: anonymised 3: pseudonymised	
	Send via Trust Centre	Done	I2	-1: not yet known 1: no 2: yes	
	Number of images send	Number	I3		
Contract between sender and receiver of imaging studies		Done	I2	-1: not yet known 1: no 2: yes	
Informed consent from patient for Teleradiology		Done	I2	-1: not yet known 1: no 2: yes	

[26] ACGT: D2.2. User requirements for an ontology based clinical data management system and for the Trial Builder. 15th September 2007

Dataset for imaging study

Context	Category		Type	Format/ Coding/Options	Ontology Reference
Institution ID doing imaging study	Institution-ID		I5		
Responsible radiologist	User-ID		I5		
Patient Pseudonym	Pseudonym		A25		
Imaging ID	Imaging-ID		A25		
Imaging study	Date done		D8	DDMMYYYY	
	Time		I2	-1: not yet known 1: at diagnosis 2: during treatment 3: after treatment 4: during follow-up 5: at relapse 6: emergency 7: other	
		If other, specify	A50		
	Type		I2	-1: not yet known 1: Sonography 2: X-ray 3: CT 4: MRI 5: Angiography 6: Scintigraphy 7: PET 8: other	
		If **CT**, specify			
		Techniques	I2	-1: not yet known 1: without contrast 2: with contrast 3: without and with contrast 4: spiral CT 5: at relapse	
		Slice thickness	I2	mm	
		Voltage	I5	kV	
		Ampere	I5	mAs	
		Quality	I2	-1: not yet known 1: bad 2: good 3: intermediate	

		If **MRI**, specify:			
		Sequences	I1	-1: not yet known 1: T1 2: T1 + contrast 3: T2	
		Fatsat	A50	Fatsat	
		IR - Sequences	A50	IR - Sequences	
		Other	A50	Other	
		Axis	I1	1: sagittal 2: coronar 3: transversal	
		Slice thickness	I2	mm	
		fMRI	I1	-1: not yet known 1: no 2: yes	
		SPECT	I1	-1: not yet known 1: no 2: yes	
		PET	I1	-1: not yet known 1: no 2: yes	
		If other **Scintigraphy**, specify			
		Type	A50		
		Angiography			
		Type	A50		
		If **Sonography**, specify			
		Type	A50		
		If **other**, specify			
		Name	A50		
		Quality	I2	-1: not yet known 1: bad 2: good 3: intermediate	
Side / Region / Field of imaging study	Description		A12	FMA Ontology	
	Name		A50		
Reason for imaging study			I2	-1: not yet known 1: primary tumour 2: metastatic disease 3: SAE / SUSAR 4: function diagnosis 5: other	
	If other, specify		A50		

Dataset for diagnosis done by imaging studies

Context			Category		Type	Format/ Coding/Options	Ontology Reference
Institution doing imaging study			Institution-ID		I5		
Responsible radiologist			User-ID		I5		
Patient Pseudonym			Pseudonym		A25		
Imaging-ID			ID		A25		
Diagnosis			Diagnosis related to		I2	-1: not yet known 1: primary diagnosis 2: relapse 3: 2nd malignancy 4: SAE / SUSAR 5: other	
			If other, specify		A50		
	ICD-O-3	Morphology	Code		A20		
			Text		A254		
		Topography	Code		A20		
			Text		A254		
	ICD-10		Code		A20		
			Text		A254		
Result	Primary tumour		Tumour volume		I2	-1: not yet known 1: not measurable 2: measurable	
					I5	ml	
			Calculation method for tumour volume		A50		
			Segmentation	done	I2	-1: not yet known 1: not measurable 2: done for T2 MRI 3: done for T1 MRI with gadolinium 4: necrotic area T1 MRI with gadolinium 5: other areas	Multiple entries possible
				If other area, specify area	A50		
				If other area: method	A50		

			tool used	A50			
			file stored at	A50			
		Morphological findings					
		Structure		I2	-1: not yet known 1: homogenous 2: inhomogeneous 3: cystic 4: not determinable		
		echogenicity, density, signal intensity		I2	-1 : not yet known 1 : iso- 2 : hypo- 3 : hyper		
		Cysts		I2	-1: not yet known 1: no 2: yes, partly cystic 3: yes, completely cystic		
		Tumour is infiltrating		I2	-1: not yet known 1: no 2: yes 3: possibly		
		Calcification		I2	-1: not yet known 1: no 2: yes		
		Necrosis		I2	-1: not yet known 1: no 2: yes, < 1/3 of tumour 3: yes, > 1/3 of tumour		
		Alteration of vessels		I2	-1: not yet known 1: no 2: yes 3: possibly		
		tumour bleeding		I2	-1: not yet known 1: no 2: yes		
		Tumour rupture		I2	-1: not yet known 1: no 2: yes, major rupture 3: yes, minor rupture		
		Malignant effusion		I2	-1 : not yet known 1 : no 2 : yes		
		If yes, specify		A50			
	Metastasis	Metastatic disease		I2	-1: not yet known 1: no 2: yes, single 3: yes, multiple		

			diagnosed with	I2	-1: not yet known 1: sonography 2: X-ray 3: CT 4: MRI 5: Angiography 6: scintigraphy 7: PET 8: other	Multiple entries possible	
	ICD-0-3	Topography of metastasis	Code	A20	Multiple entries possible		
			Text	A254			

9. Conclusions

The simulation modules described in this deliverable are based on the top-down approach being developed by the *In Silico Oncology Group* (www.in-silico-oncology.iccs.ntua.gr) under the lead of G. Stamatakos[27],[28]. The top-down approach uses clinical observations and the knowledge about the behaviour of a cancer as a whole[29]. This approach tries to identify subsystems based on physiological and biological findings that are required to build a reproducible model of a specific cancer. An iterative process continues to find the highest granularity of the system in making the model as accurate as possible in reflecting reality. Doing so, the model is kept under surveillance by the overall behaviour of the entire system[30],[31]. In contrast a bottom-up approach assembles all known parts of a system starting with genes and proteins and brings them into a formal structure until a model of the system is attained[31],[32]. The disadvantage of the bottom-up approach is the fact that the discovery of each new component needs a reconfiguration of the whole model[30],[31]. For further information available biosimulation software is summarized by Ho et al.[33] and Deisboeck et al.[34], who gives an excellent overview of '*in silico'* cancer modeling by reviewing selected studies on modeling the progression and therapy of highly malignant brain tumours.

A description of the entire field of '*in silico'* cancer research is beyond the scope of this article. The aim of '*in silico'* oncology is to develop patient specific computer simulation models of malignant tumours and normal tissues in order to optimize the planning of various therapeutic schemes. Ultimately, the aim is to contribute to the process of effectively treating cancer and to contribute to the understanding of the disease at the molecular, cellular, organ and body level.

From a clinical point of view it is expected that cancer growth and response to different treatments can be simulated. In simulating response to treatment in a given cancer this knowledge is significant for assessing better methods for treatment efficacy as the RECIST criteria[35],[36] provide. It might be time to improve traditional measures of clinical response as trial end points and to evaluate the activity on rare and resistant cancer cells. If simulation experiments are to be of more help for a clinician than providing a prediction of changes in tumour volume and shape, the response of treatment of the small fraction of resistant cancer cells will be of utmost importance in the future.

Such experiments might help clinicians in the future to find the best way of treating an individual patient by simulating different treatments in the computer before starting the treatment in reality. Two preconditions are of utmost importance before one can rely on '*in silico*' oncology models[37],[38]:

1. every '*in silico'* method has to be part of a clinico-genomic trial
2. every prediction of an '*in silico'* method has to be compared with the reality

Before establishing simulation models it is necessary to define the needed data in a first step, including data from the tumour (molecular biology, pathology, imaging), from the patient (clinical data)

---

[27] Dionysiou DD et al. A four-dimensional simulation model of tumour response to radiotherapy in vivo: parametric validation considering radiosensitivity, genetic profile and fractionation. J Theor Biol 2004; 230:1–20

[28] Stamatakos GS et al. A four-dimensional computer simulation model of the *in vivo* response to radiotherapy of glioblastoma multiforme: studies on the effect of clonogenic cell density. Br J Radiol 2006; 79:389–400

[29] Tsiknakis M et al. A semantic grid infrastructure enabling integrated access and analysis of multilevel biomedical data in support of post-genomic clinical trials on Cancer, IEEE Transactions on Information Technology in Biomedicine, Special issue on Bio-Grids, 2008; 12:191-204

[30] Friedrich CM, Paterson TS. In silico predictions of target clinical efficacy. Drug Disc Today 2004; 3:216-222

[31] Michelson S. In silico prediction of clinical efficacy. Current Opinion in Biotechnology 2006; 17:666-670

[32] Noble D. Modeling the heart – from genes to cells to the whole organ. Science 2002; 295:1678-1682

[33] Ho RL, Bartsell LT. Biosimulation software is changing research. Biotechnol Annu Rev 2004; 10:297-302

[34] Deisboeck TS et al. In silico cancer modeling: is it ready for prime time? Nat Clin Pract Oncol. 2009; 6:34-42

[35] Bogaerts J et al. Individual patient data analysis to assess modifications to the RECIST criteria. EJC 2009; 45:248-260

[36] Therasse P et al. New guidelines to evaluate the response to treatment in solid tumours. J Natl Cancer Inst 2000; 92:205-16

[37] Graf N, Hoppe A, Desmedt Ch, Lunzer A, Belleman R, Jacques J, Tsiknakis M, Stamatakos G : 'in silico' oncology for clinical decision-making in the context of nephroblastoma. Klin Pädiatr accepted, 2009

[38] Graf N, Hoppe A. What are the expectations of a Clinician from In Silico Oncology? In: Marias K, Stamatakos G (Hrsg). Proc. 2nd International Advanced Research Workshop on In Silico Oncology, Kolympari, Chania, Greece, 25-26 Sept. 2006, pp. 36-38 (http://www.ics.forth.gr/bmi/2nd-iarwiso/) (last access: 3.5.2009)

and from the possible treatment (pharmacokinetics of drugs that will be used, the treatment schema) as well as from literature and open-source databases. To make the simulation predictions precise and realistic it is crucial to get as much information as possible from each of the different categories. The amount of data will be restricted by the availability of tumour material, imaging data and clinical data. Therefore '*in silico'* oncology must always be integrated into or be part of a clinico-genomic trial, where data management including data security and anonymisation or pseudonymisation of data, along with tumour banking, are well established. In addition the trial is always reviewed by an ethical committee and fulfils all other GCP criteria to get approval by regulatory authorities[37],[38].

ANNEXES

Appendix 1: Prospective clinical trial for glioma

1. Organisation

1.1. Title of the study: Prediction of treatment response in glioma patients using 'In Silico' Modeling

1.2. Core Committee

Neuroradiology:	Prof. Dr. Wolfgang Reith Klinik für diagnostische und interventionelle Neuroradiologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 90.4 D-66421 Homburg/Saar
Neurosurgery:	Prof. Dr. Wolf-Ingo Steudel Klinik für Neurochirurgie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 90 D-66421 Homburg/Saar
Patholoy:	Prof. Dr. Wolfgang Feiden Institut für Neuropathologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 90.3 D-66421 Homburg/Saar
Chemotherapy protocol:	PD Dr. Ralf Ketter Klinik für Neurochirurgie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude D-66421 Homburg/Saar
Radiotherapy protocol:	Prof. Dr. Christian Rübe Klinik für Strahlentherapie und Radioonkologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 49, 51 D-66421 Homburg/Saar
Biostatistics:	PD Dr. Stefan Gräber Institut für Medizinische Biometrie, Epidemiologie und Medizinische Informatik Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 86 D-66421 Homburg/Saar

1.3. Primary study objectives

1.3.1. Offer a uniform, standardized concept for the treatment of patients affected by a glioma.

1.3.1.1. Radiotherapy: Use of modern techniques for planning and treatment

1.3.1.2. Chemotherapy: Use of Temozolomide during and after irradiation

1.3.2. Prediction of treatment response in glioma patients using 'In Silico' Modeling and optimization of the model by validation

2. Eligibility criteria

2.1. Age: children, adolescents and adults independent of age

2.2. Histology: Gliomas (ICD O-Code: M9380 - M9480, ICD-10 code: C71)

2.3. Primary tumour localization: intracranial

2.4. Associated conditions: Patients are eligible for the trial regardless of the presence of associated genetic diseases or syndromes like Neurofibromatosis NF I or NF II.

2.5. Primary tumour diagnosis: The tumour should not be pretreated following surgery

2.6. Informed consent: The patient and/or his legal guardian (parents) have to have declared their written informed consent to the trial.

2.7. Randomization: There is no randomization within the trial.

3. Exclusion Criteria

3.1. Primary tumour localization: diffuse intrinsic tumours of the Pons independent of histology

3.2. Dissemination: These patients will be excluded from the trial

3.3. Pretreatment: Patients treated with chemo- or radiotherapy prior to entering the study will not be included in the trial. Previous treatment with steroids is not considered a chemotherapeutic treatment.

3.4. Preexisting impairments of health status, making the conduct of the study impossible or ethically unwise.

4. Overview of protocol treatment

4.1. Basic Protocol Scheme
All patients with gliomas, eligible according to the above mentioned eligibility criteria (point 2.), should be entered into the current trial and follow the same general strategy concerning non-surgical therapy.

4.2. Treatment
The indication for non-surgical therapy in a patient with glioma following diagnosis is based upon the extent of surgical resection and the histological diagnosis.

4.3. Chemotherapy
Temozolomide is the only chemotherapy given in this protocol

4.3.1. Induction therapy
During the induction phase Temozolomide is given in combination with radiotherapy

4.3.2. Consolidation therapy
During consolidation therapy temozolomide is given up to 1 year or to the time of progression.

4.4. Radiotherapy
Patients receiving radiotherapy shall be treated according to modern treatment planning and application recommendations concerning fields and doses (total and per fraction).

	Total dose	Dose per fraction	Treatment time
Brain	54	1,8 Gy	6 weeks
Spine	50,4	1,8 Gy	5 ½ weeks

4.5. Specific study parameters

Autoantibodies against glioma specific antigens will be determined.

Clinical, imaging, histological and molecular data will be used for the 'in silico' model.

5. Study end points.

All trial patients:	Feasibility of validation of an 'in silico' model
	Overall survival, progression free survival following diagnosis
	Progression free survival, event free survival, overall survival
	Response to non-surgical therapy
	Long term sequelae, health status, quality of life

6. Statistical considerations

The aim of the trial is the validation of the prediction of the 'in silico' model. The optimization of the model will be done according to a learning loop. The tumour volume will be the parameter for validation between the patient's real data and the 'in silico' model using the patient's data.

Correlation of the tumour volume reduction between the in vivo situation and the in silico model is given below. As long as there is no good correlation in all cases the model has to be optimized.

		In vivo	
		< 10 %	> 10 %
'in silico' model	< 10 %	good correlation / no preoperative chemotherapy is indicated	bad correlation / Oncosimulator has to be improved
	> 10 %	bad correlation / Oncosimulator has to be improved	good correlation / preoperative chemotherapy is indicated

This therapy optimization trial is single-center and prospective.

The accrual period of the trial is 2 years followed by an observation period of 2 years.

Defined variables will be checked with reference to their influence upon the survival variables by Cox regression.

Appendix 2: Prospective clinical trial for lung cancer

1. Organisation

1.1. Title of the study: Prediction of treatment response in lung cancer patients using 'In Silico' Modeling

1.2. Core Committee

Radiology:	Prof. Dr. Arno Bücker Klinik für diagnostische und interventionelle Radiologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 49 D-66421 Homburg/Saar
Thoracic surgery:	Prof. Dr. Hans Joachim Schäfers Klinik für Thorax und Herz-Gefäßchirurgie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 56, 57 D-66421 Homburg/Saar
Patholoy:	Prof. Dr. Rainer Bohle Institut für Allgemeine und Spezielle Pathologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 26 D-66421 Homburg/Saar
Chemotherapy protocol:	NN Klinik für Pneumologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude D-66421 Homburg/Saar
Radiotherapy protocol:	Prof. Dr. Christian Rübe Klinik für Strahlentherapie und Radioonkologie Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 49, 51 D-66421 Homburg/Saar
Biostatistics:	PD Dr. Stefan Gräber Institut für Medizinische Biometrie, Epidemiologie und Medizinische Informatik Universitätsklinikum des Saarlandes Kirrberger Straße, Gebäude 86 D-66421 Homburg/Saar

1.3. Primary study objectives

1.3.1. Offer a uniform, standardized concept for the treatment of patients affected by lung cancer.

1.3.1.1. Radiotherapy: Use of modern techniques for planning and treatment

1.3.1.2. Chemotherapy: Use of ….

1.3.2. Prediction of treatment response in lung cancer patients using 'In Silico' Modeling and optimization of the model by validation

2. Eligibility criteria

2.1. Age: all patients independent of age

2.2. Histology: lung cancer (ICD O-Code: M-8010/3, ICD-10 code: C34)

2.3. Primary tumour localization: lung, bronchi

2.4. Associated conditions: Patients are eligible for the trial regardless of smokers or non-smokers

2.5. Primary tumour diagnosis: The tumour should not be pretreated following surgery

2.6. Informed consent: The patient and/or his legal guardian (parents) have to have declared their written informed consent to the trial.

2.7. Randomization: There is no randomization within the trial.

3. Exclusion Criteria

3.1. Primary tumour localization: none

3.2. Dissemination: Patients with metastasis outside the lung will be excluded from the trial

3.3. Pretreatment: Patients treated with chemo- or radiotherapy prior to entering the study will not be included in the trial.

3.4. Preexisting impairments of health status, making the conduct of the study impossible or ethically unwise.

4. Overview of protocol treatment

4.1. Basic Protocol Scheme
All patients with lung cancer, eligible according to the above mentioned eligibility criteria (point 2.), should be entered into the current trial and follow the same general strategy concerning non-surgical therapy.

4.2. Treatment
The indication for non-surgical therapy in a patient with glioma following diagnosis is based upon the extent of surgical resection and the histological diagnosis.

4.3. Chemotherapy
… is the chemotherapy given in this protocol

4.3.1. Induction therapy
During the induction phase … is given in combination with radiotherapy

4.3.2. Consolidation therapy
During consolidation therapy … is given up to … year or to the time of progression.

4.4. Radiotherapy
Patients receiving radiotherapy shall be treated according to modern treatment planning and application recommendations concerning fields and doses (total and per fraction).

	Total dose	Dose per fraction	Treatment time
Tumour region			
Boost			

4.5. Specific study parameters

Sequencing of the EGF Receptor will be done.

Autoantibodies against lung cancer specific antigens will be determined.

Clinical, imaging, histological and molecular data will be used for the 'in silico' model.

5. Study end points.

All trial patients: Feasibility of validation of an 'in silico' model

Overall survival, progression free survival following diagnosis

Progression free survival, event free survival, overall survival

Response to non-surgical therapy

Long term sequelae, health status, quality of life

6. Statistical considerations

The aim of the trial is the validation of the prediction of the 'in silico' model. The optimization of the model will be done according to a learning loop. The tumour volume will be the parameter for validation between the patient's real data and the 'in silico' model using the patient's data.

Correlation of the tumour volume reduction between the in vivo situation and the in silico model is given below. As long as there is no good correlation in all cases the model has to be optimized.

		In vivo	
		< 10 %	> 10 %
'in silico' model	< 10 %	good correlation / no preoperative chemotherapy is indicated	bad correlation / Oncosimulator has to be improved
	> 10 %	bad correlation / Oncosimulator has to be improved	good correlation / preoperative chemotherapy is indicated

This therapy optimization trial is single-center and prospective.

The accrual period of the trial is 2 years followed by an observation period of 2 years.

Defined variables will be checked with reference to their influence upon the survival variables by Cox regression.

Appendix 3: Information Sheet

Clinically Oriented Translational Cancer Multilevel Modelling

PATIENT INFORMATION SHEET

Concerning the participation in a research project on the characterization of cancer

1. Invitation:

You are being invited to take part in a research project which aims at the molecular characterisation of different types of cancer. The project has been proposed to you because your cancer is of the type which will be analyzed in the context of this project.

The project is called ContraCancrum, which means against cancer. The full title is: “Clinically Oriented Translational Cancer Multilevel Modelling”.

ContraCancrum is sponsored by the European Union. The project is being conducted by members of the ContraCancrum research consortium which consists of about 8 research groups from different countries and disciplines. The tools and results produced by this research project will be publicly available.

We ask you to provide your data to this research project. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve.

This document describes the project in order to help you to decide whether or not you want to participate in it. Please feel free to discuss any issue regarding this research project with your medical doctor or medical staff.

Take time to decide whether or not you wish to take part. You must not feel obliged to participate in this research project. If you do decide to participate, you can withdraw your consent any time without any disadvantages.

Thank you for reading this.

2. Purpose of the project

ContraCancrum aims at significantly contributing to the understanding of hypercomplex biological phenomena through the multilevel modelling of cancer in the clinical setting. This will take cancer modelling research a step further by integrating molecular, cellular, tissue and higher level modelling concepts into a single technological entity that will simulate therapy outcome based on the individual patient information. This could serve as a powerful weapon to better understand and fight cancer.

The main objectives of the ContraCancrum project are:

- Develop a composite multilevel simulation model of malignant tumour growth and tumour and normal tissue response to therapeutic modalities and treatment schedules.
- Validate the models, by exploiting the outcome of pertinent clinical trials and designing and carrying out new dedicated clinicogenomic studies in (a) gliomas (b) lung cancer.
- Promote the development of individualised therapies and in silico therapy optimisation.
- Translate, in the mid and long term, validated multilevel cancer models into clinical practice.

The knowledge generated in the context of this project may eventually lead to a better, more individualized cancer therapy. However, it is important for you to understand that the project itself does not aim to develop a new treatment but to generate basic knowledge on cancer.

The research project will be completed around 2011.

3. Do you have to take part?

Your participation to this study is entirely voluntary: you are completely free to participate or not to this study. You are also free to withdraw at any time without giving any reason and this will not affect your medical care or the relationship with your medical doctor or medical staff.

By signing the informed consent form, you will confirm that you were properly informed about this project and that all your questions have been answered. A copy of the patient information sheet and of the consent form will be given to you to keep.

4. What will happen to you if you take part?

If you have decided to take part in ContraCancrum, your tumour and, your blood and your imaging studies will be analyzed with respect to different characteristics. On the genetic and pathological level cell types will be analyzed, proteins and patterns of gene expression (e.g. gene activity) will be characterized. Your imaging studies will be used to measure the tumour volume and correlate it to the histological findings.

The data generated by this analysis will be sent to ContraCancrum and used for simulating the natural course of your tumour and the response to treatment options in the computer (In Silico Oncology).

Since research is being done on blood, tumour samples and imaging studies which have already been collected from you before or during treatment, participation in this project does not imply extra visits to the hospital, nor extra examinations.

5. How are your data protected?

The data which are being transmitted to ContraCancrum are socio-demographic data (sex, age, marital status, number of children, profession, region, etc.), clinical data (type and stage of your cancer as well as other information related to your health and disease), imaging data (X-ray, CT-scans, MRI-scans, PET-scans, etc.), biological data (characteristics of cells and proteins, etc.), and genomic data (for example, data on the genes which are expressed in your cancer).

Personal information like your name or address is not being transmitted to ContraCancrum. Before data are being sent to ContraCancrum, any personal identifiers are removed by the hospital. The

procedure of disconnecting biological, clinical and other data from personal identifiers is called pseudonymization.

Your pseudonymized data will be stored until the project is completed, which will be around 2011.

If you decide to withdraw your consent to this project, no more data will be collected from your sample or file. The data which have already been sent to ACGT can only be used further on if they are anonymized. In that case, they can no longer be linked back to your person.

6. Will you be informed about the results of the project?

You will not be personally informed about the results of the research conducted in the context of ContraCancrum. Since this research aims at the generation of basic scientific knowledge, it is – at this stage – not relevant for the treatment of single patients.

This does not affect your right provided by law to access your processed data and ask for rectification of these data, if any inaccurate information is stored.

It may be possible, albeit not very likely, that research conducted in the context of ContraCancrum yields results which may be of direct relevance for your treatment or for the prevention of future ailment. If you consent to participating in this project, you may choose whether or not you want to be informed of such results by your doctor.

7. Risks and benefits of this project:

The data transmitted to ContraCancrum will be extracted from your patient file and the blood and tissue samples which have been collected by your medical doctor or medical staff before diagnosis and during treatment. Therefore, there will not be any extra procedure, examination or visit involving any risk needed.

Before your data can be used for research, they are disconnected from personal identifiers which confirms to the safety standards requested by law. Yet there is a residual, albeit extremely small risk that these data can be traced back to your person.

You will receive no direct benefit by agreeing to include your data in ACGT. However, the availability of such data for research use is important to the advancement of knowledge about cancer and the future development of new treatments.

8. Costs

There will not be any additional costs involved if you decide to participate in the ACGT-research project.

9. Contact

For any questions related to this research project please contact:

<Name/s>

<Institution/Hospital>

<Phone number/s>

Appendix 4: Template for 'Informed Consent'

Patient : _____

INFORMED CONSENT FORM

Title of Project: **Bio-Banking in ContraCancrum**

Name of Clinician: _____

Please initial box

_____

1. I confirm that I have read and understand the information sheet for the above research project and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐

3. If I withdraw from the research project, I consent to my doctor providing authorised researchers with basic clinical information that would routinely be collected and written in my medical records. ☐

4. I understand that I donate tissue of my tumour and other biological material (like blood or other fluids) freely to ContraCancrum solely for research. There are no extra procedures necessary for obtaining these biological samples. ☐

5. I understand that clinical, imaging and genetic data will be used for research purposes in ContraCancrum. ☐

6. I understand that sections of any of my medical notes may be looked at by responsible individuals from regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐

7. I understand that data, as described in the Patient Information Sheet will be passed to Members of the ContraCancrum Research Project. These members are: ☐

Participant No.	Participant Name	Country
1.	Foundation for Research and Technology Hellas (FORTH)	Greece
2.	Institute of Communication and Computer Systems (ICCS)	Greece
3.	Universität des Saarlandes (USAAR)	Germany
4.	University College London (UCL)	UK
5.	Universitat Bern (UBERN)	Switzerland
6.	University of Bedfordshire (BED)	UK
7.	Univerzita Karlova v Praze (CUNI)	Czech Republic
8.	Philips Technologie GmbH (PFL-H)	Germany

More information is provided on the following Web Page:
http://contracancrum.eu/?q=node/3

8. I understand that my name will be pseudonymized when I join the study, and that thereafter I will be identified by a unique trial number. Any information passed to regulatory authorities will not identify me as an individual. ☐

9. I understand the main objectives of the ContraCancrum project as explained in the information sheet. I did understand that the project itself does not aim to develop a new treatment for me but to generate basic knowledge on cancer. ☐

10. I want to be informed of research results by my doctor. ☐

11. I agree to take part in the above research project. I have no further questions. ☐

_____	_____	_____
Name of Patient	Date	Signature

I confirm that I have explained the nature, purposes and foreseeable effects of the trial to the subject whose name is printed above.

_____	_____	_____
Name of Person taking consent (if different from researcher)	Date	Signature

_____	_____	_____
Researcher	Date	Signature

1 sheet for patient; 1 for researcher; 1 to be kept with hospital notes

Abbreviations and acronyms

AA	Anaplastic Astrocytoma
ADC	Apparent Diffusion Coefficient
AE	Adverse Event
CT	Chemotherapy
DA	Diffuse Astrocytoma
DTI	Diffusion tensor imagaing
DWI	Diffusion weighted Imaging
ECM	Extracellular Matrix
EGFR	Epidermal Growth Factor Receptor
FLAIR	Fluid Attenuation Inversion Recovery
GBM	Glioblastoma
LOH	Loss of Heterozygosity
MO	Master Ontology
MPRAGE	Magnetization Prepared Rapid Gradient Echo
ObTiMA	Ontology based Trial Management Application
PA	Pilocytic Astrocytoma
RT	Radiotherapy
S	Surgery
SAE	Severe Adverse Event
SUSAR	Suspected Unexpected Severe Adverse Reaction
SWI	Susceptibility-weighted Imaging
URL	Uniform Resource Locator