Deliverable No. 5.1

Setting up of the data protection and data security framework

Grant Agreement No.:	270089
Deliverable No.:	D5.1
Deliverable Name:	Setting up of the data protection and data security framework
Contractual Submission Date:	31/01/2012
Actual Submission Date:	31/01/2012

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

SEVENTH FRAMEWORK PROGRAMME

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	***p-medicine***
Project Full Name:	From data sharing and integration via VPH models to personalized medicine
Deliverable No.:	D5.1
Document name:	Setting up of the data protection and data security framework
Nature (R, P, D, O)[1]	R
Dissemination Level (PU, PP, RE, CO)[2]	PU
Version:	2
Actual Submission Date:	31/01/2012
Editor: Institution: E-Mail:	Prof. Dr. Nikolaus Forgó Leibniz Universität Hannover, Institut für Rechtsinformatik forgo@iri.uni-hannover.de

ABSTRACT:

This deliverable contains an analysis of the relevant legal and ethical requirements for p-medicine and describes the data protection and data security framework for the project.

Within the framework-project on Advanced Clinico Genomic Trials on Cancer (ACGT) a data protection and data security framework for the use of patient data for scientific research has already been set up. This framework shall serve as a starting point for the development of the data protection and data security framework for p-medicine. The new framework for p-medicine will have to meet additional challenges, such as the creation of common data warehouses and the implementation of a patient identity management system.

This document provides a brief overview on the legal rules governing the use of personal data, in particular health related data, on a European level. This overview shall serve as a basis for the better understanding of the subsequent legal deliberations leading to the creation of the legal framework the processing of patient data within p-medicine. Furthermore, we briefly analyse the impact of Directive 2001/20/EC on clinical trials and of Directive 2001/83/EC on the Community code relating to medicinal products for human use on p-medicine.

Subsequently we provide an evaluation of the data protection approaches elaborated in comparable projects in the field of medical research. Starting with an in depth evaluation of the framework designed for ACGT, we then give an overview of selected other European medical research projects and legal and ethical guidelines developed for the use of patient data in medical research. By depicting these different approaches commonalities of research projects involving patient data will be identified, that could serve for the further elaboration and refinement of the p-medicine framework.

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

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

Taking into account the outcomes of this evaluation the data protection and data security framework for p-medicine will be designed. It is recommended to introduce a clear separation of the clinical area and the research area. Within the research area only de facto anonymous data shall be used. De facto anonymisation is not only the best way to ensure and protect the patients privacy, it also provides a high level of flexibility regarding their use for research, as the processing of de facto anonymous data does not fall under the processing restrictions of the of the Data Protection Directive 95/46/EC. Therefore, it is recommended to establish a data protection framework based on a double pseudonymisation procedure, the establishment of a Center for Data Protection that serves as a central data protection authority, the inclusion of a Trusted Third Party that serves as a trusted data custodian and the creation of a network of trust based on contractual agreements that shall ensure the compliance to the data protection rules set up for the framework. We analyse and describe the basic essentials to be ruled in the contracts explaining why the different clauses of the contracts have been designed the way they are. Furthermore we deliver the necessary contracts for the setting up of the framework, which are the “Data Transfer Agreement” (Annex A) and the “Contract on data protection and data security within p-medicine” (Annex B).

For the unlikely case that we will have in some situations personal data anyway, we are confident of still being in line with data protection regulations as we will of course have (also for ethical reasons) informed consents for the data processing from the patients. An explicit informed consent is a major possibility foreseen by the Directive to make the processing of sensitive personal data legal (see Art. 8 para. 2 lit a). Consent forms to be used are provided in Annex B of this deliverable and their usage will be a contractual obligation for partners bringing data into the p-medicine infrastructure.

KEYWORD LIST: data protection, informed consent, patient empowerment, genetic data, data security, anonymisation, pseudonymisation, network of trust access to information, transfer of data, good clinical practice.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 270089.

The author is solely responsible for its content, it does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of data appearing therein.

MODIFICATION CONTROL			
Version	Date	Status	Author
1.0	15/01/2012	Draft	Nikolaus Forgó, Hartwig Gerhartinger, Tina Krügel, Brecht Claerhout, Elias Neri
2.0	31/01/2012	Final Version	Nikolaus Forgó, Hartwig Gerhartinger, Tina Krügel, Brecht Claerhout, Elias Neri

List of contributors

- Nikolaus Forgó, LUH
- Hartwig Gerhartinger, LUH
- Tina Krügel, LUH
- Brecht Claerhout, Custodix
- Elias Neri, Custodix

Contents

1 Executive Summary

In this deliverable the data protection and data security framework for p-medicine is elaborated. This document contains an overview on the legal framework to be respected for the use of patient data for research purposes on a European level as well as an evaluation of the data protection regime set up for ACGT and of comparable research infrastructures. Based on this analysis the data protection and data security framework for p-medicine is designed.

The analysis of the legal requirements for lawfully processing patient data for the purposes of scientific research lays special emphasis on the issues of data protection and privacy. The starting point of the analysis is the European Data Protection Directive 95/46/EC, which introduces rules applicable to every processing of personal data and sensitive data on a European level. Since health data, in particular genetic data, contains very sensitive information not only about the patient but potentially also about his/her relatives and their possible diseases, etc., the processing of this kind of data is only possible under special requirements under Directive 95/46/EC. As every EU Member State has to implement the rules of this Directive into national law, for an EU-wide project like p-medicine, this Directive is the common legal basis for all participants residing within the EU or are processing the data over an infrastructure established in a EU Member State. Furthermore, the applicability of the Directives 2001/20/EC, known as Clinical Trials Directive, as well as Directive 2001/83/EC on the Community code relating to medicinal products for human use are analysed.

It is not the aim of p-medicine to elaborate a completely new infrastructure for medical scientific research from scratch. P-medicine shall rather use the work done in ACGT as a starting point. Hence, this document provides a sound evaluation of the framework of ACGT in order to identify the strengths of the ACGT framework as well as the potential for improvement p-medicine. Furthermore, it appears useful to analyse the data protection approaches developed for comparable research projects processing medical patient data. The analysis shows that ACGT already provided for a relatively high level of data protection, since data were only processed in (de facto) anonymous form. This approach is still state of the art, as the analysis of other projects and guidelines in the field of medical research shows.

The practical implementation of the framework, in particular the negotiations of the contracts with the hospitals delivering the data, was however rather time consuming in ACGT. Accordingly it is intended to simplify the contracts as much as possible within p-medicine. Furthermore there was no mechanism within ACGT needed and produced providing that the data set of the same patient coming different hospitals could be stored under the same pseudonym. The creation of synonyms and homonyms, however, can affect the quality of the data as well as the usability of the data for long-term prognosis in a negative way. Accordingly, it is a challenge for p-medicine to implement such a system. Therefore, a patient management system has to be implemented in the data protection and data security framework of p-medicine. The analysis of the data protection frameworks of comparable projects reveals that transparency is a very important factor for medical research projects. In order to ensure transparency a clear separation of responsibilities for the processing of the data is required.

The technical evaluation of the security framework in ACGT shows that the standard GRID middleware components covered a variety of security problems. Our analysis showed, however, that this solution also came with considerable disadvantages, for example: unused overhead functionality introduced maintenance of unnecessary components, a high complexity and performance degradation. Furthermore experience learned that some of the chosen technology (although standards based), i.e. X.509 client certificates, was not perceived as user friendly by the end users and by some considered hampering for their productivity.

Taking into consideration the outcomes of this analysis we recommend to build the data protection framework for p-medicine on a clear separation of the clinical domain and the research domain and to define clear responsibilities for each domain. Within the treatment domain data are collected in the hospital domain in the course of the patient's treatment. The processing of data within the treatment domain shall be under the responsibility of the hospitals, respectively the treating physician. The data processing within the research domain shall be under the control of a central data protection authority for p-medicine that concludes the necessary contracts for p-medicine and controls the compliance to these rules.

The law states that data shall only be used for the purposes of research in anonymous form. Anonymisation means that it is impossible to go back to the patient even if there may be findings that could be helpful for the patient's treatment. Given that the p-medicine infrastructure is supposed to offer that possibility to contact the patient in case of relevant findings, complete anonymisation is not an option. Hence p-medicine has to find a balance for these two competing aims.

In order to comply with current data protection legislation and to ensure the privacy rights of the patient, it is recommended in this deliverable that only de facto anonymous data shall enter the research domain. Pseudonymised data are regarded as de facto anonymous when the persons processing the data (and nobody else) are not able to legally establish the link to the patient with reasonable means. Within p-medicine de facto anonymisation will be achieved by a state of the art pseudonymisation of all data entering the network, the implementation of a Trusted Third Party, the work of a data protection authority within p-medicine and the signing of contracts by all participants ensuring the compliance with the p-medicine data protection and data security policies that will primarily exclude the use of matching tables for re-identification of patients.

Accordingly the data protection framework for p-medicine consists mainly of three parts:

First, a legal body will have to be implemented that serves as central data protection authority and is able to conduct contracts regarding data protection on behalf of p-medicine. This task will be fulfilled by the Center for Data protection, an association under Belgian law that had already successfully served as data protection authority within ACGT.

Second, a Trusted Third Party (TTP) is needed in this data protection framework, which is used for the pseudonymisation of the patient data. The TTP accepts a pseudonym or other identifying information and transforms it (e.g. by using HMAC) into a p-medicine pseudonym. The TTP will also act as a trustful custodian for the pseudonymisation key to re-identify the patient concerned upon request by the CDP if re-identification should be needed (which might be necessary if new knowledge relevant to the patient is discovered. A Patient Identity Management Service (PIMS) can be involved before calling the TTP. PIMS indexes the patient's identifying information (and only the identifying information, no additional data) to issue the same pseudonym for the same patient even if data comes from different centres. This pseudonym issued by PIMS is then passed through the TTP to ensure that there is no link between the pseudonym and the identifiers within the network of trust.

Third, contracts between all participating hospitals, researchers or other users of the patient data and p-medicine must be concluded in order to ensure confidentiality, data security and compliance with data protection legislation. Mainly three contracts are required to set up the framework. The first deals with the transfer of patient data to the p-medicine infrastructure (Data Transfer Agreement). This agreement is to be concluded between the CDP and the healthcare organisation/hospital delivering patient data. The second agreement concerns the data processing within the p-medicine framework (Contract on data protection and data security within p-medicine). This agreement will have to be concluded between the CDP and all end-users of p-medicine doing research on this data. This deliverable introduces the essential provisions of these two contracts drafted such as the question of data control within p-medicine, obligations concerning the network of trust, third beneficiary rights and the applicable law. The different topics are discussed and it is explained for what reasons it has been decided to design the different provisions as they are. The contracts are included in the Annex to this document. The third contract will have to be concluded between the CDP and the Trusted Third Party and shall regulate the independence of the TTP as well as conditions for the storage of the pseudonymisation links and the procedure for re-identifying the patient. This contract will have to be negotiated directly with body serving as TTP when these questions are solved.

De facto anonymisation of the patient data is not only the best way to ensure the patient's privacy rights while ensuring a possibility to address the patient when the research reveals findings that would be beneficial for the patient. It also provides for high flexibility for the researchers as de facto anonymous data do not fall into the scope of the Data Protection Directive and the national data protection laws and the processing of these data, thus, does not underlie the respective processing restrictions. Accordingly the concept of de facto anonymous data also provides for the possibility to build up data warehouses in different countries to which all researchers within the project can have access under the same conditions. These conditions are set up by the contractual agreements concluded by the CDP and the end users and equally guarantee for a high level of data protection and data security.

From a technical point of view the data security requirements as well as the end user needs for p-medicine shall be addressed with a lightweight dynamic security architecture based on commonly used standards and implementations such as SAML and XACML. The overall goal of p-medicine is to offer at least the same level of data protection compliance and security as within the ACGT project (meeting some additional requirements as explained in this document), while keeping a much stronger focus on the aspects which lead to sustainability (exploitation), such as: user-friendliness, standards compliance, integration with industry standard solutions, maintenance, etc.

For the unlikely case that we will have in some situations personal data anyway, we are confident of still being in line with data protection regulations as we will of course have (also for ethical reasons) informed consents for the data processing from the patients. An explicit informed consent is a major possibility foreseen by the Directive to make the processing of sensitive personal data legal (see Art. 8 para. 2 lit a).

2 Introduction

Within p-medicine a transfer of patient data will take place among the partners of the project. The transfer of personal data is governed by privacy regulation, in particular the regime of Directive 95/46/EC (Data Protection Directive), so that the rules and mechanisms of the Directive have to be respected. The processing of personal data may only be effected in compliance with the applicable (national) data protection rules that are transpositions of the Directive. Compliance with the relevant laws is therefore a primary task of the legal/ethical WP in p-medicine. Compliance does not come on its own but has to be negotiated, agreed and controlled. A proper tool for those tasks are contracts concluded between project partners in addition to the existing Consortium agreement. To ensure the compliance to the standards set up by these laws the transfer of data may only be effected on the basis of contracts developed by WP 5 that are by far more specific than the Consortium agreement.

Within the framework-project on Advanced Clinico Genomic Trials on Cancer (ACGT) a data protection and data security framework has already been set up. This framework shall be used as a starting point for p-medicine. It will be evaluated and extended according to the needs of p-medicine.

The additional legal and ethical questions related to p-medicine regarding the data warehouse and data mining as well as the access to biobanks or the legal and ethical issues regarding patient empowerment and international clinical trials are subject to upcoming deliverables and therefore will be only briefly be addressed where necessary.

3 Structure

The document is divided in four main parts.

The first part (chapter 4) shall give a brief overview about the legal and ethical requirements for the use of patient data for the purposes of scientific research. Furthermore the possible implications of the Directive 2001/20/EC on clinical trials and of Directive 2001/83/EC on the Community code relating to medicinal products for human use as far as relevant for p-medicine shall systematically be analysed.

On the basis of this overview we will choose the framework used in ACGT as a starting point and evaluate its strengths and weaknesses (chapter 5). This evaluation aims to clarify the strengths and the potential for improvement of the existing framework. In order to facilitate this evaluation the legal framework (subchapter 5.3) and the data security framework (subchapter 5.4) will be analysed separately. This chapter aims to identify the challenges for the p-medicine data protection and data security framework.

In the following part (chapter 6) we will analyse comparable data protection frameworks of respective medical research projects in order to examine possible improvements or simplifications for the legal framework to be set up in the course of p-medicine.

In the fourth part (chapter 7) the data protection and data security framework for p-medicine will be elaborated. After a short introduction the data flows within the p-medicine infrastructure will be analysed. Subsequently the data protection framework (subchapter 7.3) and the data security framework (subchapter 7.4) for p-medicine will be set up. The legal data protection framework requires the conclusion of contractual agreements building up a network of trust in order to ensure the patients´ privacy rights. This deliverable introduces the essential provisions of the contracts drafted (such as the question of data control within p-medicine, obligations concerning the network of trust, third beneficiary rights and the applicable law) and explains for what reasons it has been decided to design the different provisions as they are. The contractual agreements proposed are attached in the Annex to this document. They will be presented to the consortium and will be negotiated within WP 5.

The final chapter 8 contains the legal and data security conclusions.

4 Overview of the legal requirements

4.1 Introduction

This chapter shall give an overview of the legal (and ethical) requirements for the transfer and processing of medical data, focusing on the transfer and processing of genetic data for the purposes of scientific research.

In a first step we provide an analysis of the legal requirements established on a European level in order to give an overview of the legal standards to be respected to lawfully establish the p-medicine framework. Within EU legislation we will focus on the general rules and principles for processing of personal data stated by the Directive 95/46/EC on the protection of individuals with regard to the processing personal data and on the free movement of such data[3] (Directive 95/46/EC, Data Protection Directive). The Data Protection Directive sets out the rights of the data subject and control mechanisms, establishes general rules on the lawfulness of the processing of personal data, and regulates the transfer of personal data into third countries. Directive 95/46/EC, thus, introduces the rules applicable to every processing of personal data throughout the EU. As it only covers the processing of personal data, whereas the processing of anonymous data does not fall into its scope, special attention shall be laid on the definition of these terms.

However, there may be more specific rules governing the use of patient data under specific circumstances. This might be the case, when data are used in the context of clinical trials. Therefore, Directive 2001/20/EC on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use[4] (Directive 2001/20/EC, Clinical Trials Directive) that seeks to simplify and harmonize the administrative provisions governing clinical trials by establishing a clear, transparent procedure as well as the Directive 2001/83/EC on the Community code relating to medicinal products for human use. These two Directives, therefore, will also have to be analysed in order to identify possible implications for the data protection and data security framework for p-medicine.

Ethical requirements will additionally be taken into consideration. These ethical standards are not always written in a legal reference document and may go beyond the legal rules. The main ethical issues to be discussed are related to the notion of "informed consent" and the possibility to access personal information ("right to know"), the "duty to inform" the patient of research results and the "quality of feedback" given by the researchers.[5]

4.2 Legal requirements for the use of data within p-medicine

4.2.1 Data protection Directive (Directive 95/46 EC)

Directive 95/46/EC has two main purposes:

(1) to allow for the free flow of data within the EU in order to prevent the Member States from blocking cross-border data flows on grounds of data protection within the EU and

(2) to establish a minimum level of data protection throughout all Member States.[6]

[3] Published in Official Journal L 281, 23/11/1995, p. 31.

[4] Published in Official Journal L 121, 01/05/2001, p. 34.

[5] See chapter 4.3 below.

[6] Kuner, European Data Protection Law, rec. 1.40.

It had to be transposed to national law by all EU Member States, so that all national laws within the EU reflect the basic rules set by this Directive.[7] In the respective landmark ruling of the European Court of Justice, C-101/01, Lindquist, 06/11/2003,[8] the ECJ clarified that Directive 95/46/EC envisages complete harmonisation of the data protection regime within its scope. Accordingly Member States in principle have to adopt national legislation conforming to regime of the Directive. However, certain provisions of the Directive can explicitly authorise the Member States to adopt more constraining data protection rules. In doing so they, however, have to maintain a balance between free movement of personal data and protection of private life. Accordingly the Member States are allowed to set higher standards under specific circumstances, so that the data protection law is not completely harmonised within the EU. With regard to areas excluded from scope of application of Directive 95/46/EC Member States are free to regulate these areas in their own way, whenever there is no other rule of Community law providing otherwise.

In the following we will depict the most important legal concepts given in Directive 95/46/EC, including the different data categories as well as the requirements of a fair and lawful processing of data.

4.2.1.1 Scope of the Directive and categories of data

Directive 95/46/EC distinguishes between several categories of data: “personal data”, “pseudonymous data”, “sensitive data” and “anonymous data”. The main distinction is made between personal data and anonymous data, since according to Art. 3 para. 1 Directive 95/46/EC, the Directive is applicable only to the processing of “personal data”, whereas “anonymous data” is not subject to the processing-restrictions of the Directive.[9] Pseudonymous data and sensitive data define special categories of personal data, so that these categories of data in general underlie the scope of the Directive.

4.2.1.1.1 Personal data

Art. 2 lit. a Directive 95/45/EC defines the term “personal data” as “any information relating to an identified or identifiable natural person ('data subject')“. An identifiable person according to this provision is a person that “can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity”.[10] Accordingly the main criterion appearing in these definitions is that of identifiability, meaning the potential of information to enable identification of an individual.[11]

Identification can be achieved through “identifiers”, meaning pieces of information, which hold a particularly privileged and close relationship with the particular individual concerned. For direct identification the name of a person is the most common identifier. However, a unique identifier, e.g. a number, is often assigned to the persons registered in a file in order to avoid confusion between two persons in the file. This unique identifier, i.e. the number, then refers to an identified natural person. Indirect identification usually requires several identifiers, leading to the phenomenon of unique

[7] See Art. 32 Directive 95/46/EC.

[8] European Court of Justice, C-101/01, *Lindquist*, 06/11/2003, ECR 2003, I-12971.

[9] See also Recital 26 of Directive 95/46/EC: „Whereas the principles of protection must apply to any information concerning an identified or identifiable person; … whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable; …”.

[10] Art. 2 lit. a) Directive 95/46/EC.

[11] See also Kuner, European Data Protection Law, rec. 2.74.

combination that allows a certain person to be singled out, e.g. date of birth, gender, address.

The question whether data has to be regarded as identifiable poses severe problems in practice. According to recital 26 of Directive 95/46/EC in answering this question "account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify said person." Thus, if the possibility of "all the means likely reasonably to be used" does not exist or is negligible, the person is not identifiable and the information not qualified as personal data.[12] In this context the Article 29 Data Protection Working Party[13] states in its Opinion 4/2007 that the relevant factors for assessing the question of means likely reasonably to be used were cost, intended purpose, the way the processing is structured, the advantage expected by the controller, the interests at stake for the individuals as well as the risk of organisational dysfunctions (e.g. breaches of confidentiality) and technical failures.[14] It is pointed out that the assessment of these factors is likely to change over time and therefore the system should be able to adapt to new developments as they happen by incorporating the appropriate technical and organisational measures.[15]

In the mentioned opinion the Working Party gives an example of pharmaceutical research data where identifiability of the data subject is unlikely:

"Hospitals or individual physicians transfer data from medical records of their patients to a company for the purpose of medical research. No names of the patients are used but only serial numbers attributed randomly to each clinical case, in order to ensure coherence and to avoid confusion with information on different patients. The names of patients stay exclusively in possession of the respective doctors bound by medical secrecy. The data do not contain any additional information, which make identification of the patients possible by combining it. In addition, all other measures have been taken to prevent the data subjects from being identified or becoming identifiable, be it legal, technical or organisational. Under these circumstances, a Data Protection Authority may consider that no means are present in the processing performed by the pharmaceutical company, which make it likely reasonably to be used to identify the data subjects."[16] Therefore, in this example, data protection legislation would not apply to the processing of the data sets for medical research at the company.

However, when referring to the purpose of the processing of personal data, the Working Party points out that: "where the purpose of the processing implies the identification of individuals, it can be assumed that the data controller or any other person involved have or will have the means likely reasonably to be used to identify the data subject. In fact, to argue that individuals are not identifiable, where the purpose of the processing is precisely to identify them, would be a sheer contradiction in terms."[17] In this case, the Working Party concludes, the processing is subject to data protection rules.

If, however, the purpose of the processing is not the identification of the person concerned, technical and organisational measures taken to prevent identification play

[12] See also Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, p. 15, available at http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

[13] The Working Party has been established by Article 29 of Directive 95/46/EC. It is the independent EU Advisory Body on Data Protection and Privacy. For further information please see http://ec.europa.eu/justice/policies/privacy/workinggroup/index_en.htm.

[14] Article 29 Data Protection Working Party, Opinion 4/2007, p. 15.

[15] Article 29 Data Protection Working Party, Opinion 4/2007, p. 15.

[16] Article 29 Data Protection Working Party, Opinion 4/2007, pp. 15-16.

[17] Article 29 Data Protection Working Party, Opinion 4/2007, p. 16.

an important role in assessing whether the data are personal or not. If state of the art technical and organisational measures are taken to protect the data against identification, the persons will not be identifiable taking account of all the means likely reasonably to be used by the controller or by any other person to identify the individuals.[18]

4.2.1.1.2 Anonymous data

Directive 95/46/EC considers data as anonymous only, if the data subjects not identifiable. This is the case if the link that refers to the data subject has been irrecoverably erased or has never existed.

Anyhow the German legislation for instance seized the suggestion of the proposal and, unlike the European legislation, implemented the "excessive effort" in its definition of anonymous data (section 3 para. 6 of the Federal Data Protection Act (BDSG)[19]).

Meanwhile the "excessive effort" has been introduced to EU law.[20] Therefore information concerning personal or material circumstances that can only with a disproportionate amount of time, expense and labour be attributed to an identified or identifiable individual is anonymous data.

In this context it has to be pointed out that anonymisation of personal data is processing of personal data so that the principles of protection apply to personal data before it is rendered anonymous.

4.2.1.1.3 Pseudonymous data

In contrast to some national data protection regulations, Directive 95/46/EC does not know the term "pseudonymous data". Pseudonymous data are therefore to be seen as personal data.

The German Federal Data Protection Act (BDSG) e.g. defines in section 3 para. 6a pseudonymising as "replacing a person's name and other identifying characteristics with a label, in order to preclude identification of the data subject or to render such identification substantially difficult". Especially in a medical research project, the use of pseudonymous data often is mandatory and beneficial for the patient, because it is possible to re-identify the patient in case of newly developed treatments.

4.2.1.2 Territorial application

The Data Protection Directive is based on the territoriality principle. Art. 4 para. 1 lit. a Directive 95/46/EC provides that the Member State shall apply the national provisions it adopts pursuant to this Directive to the processing of personal data where the processing is carried out in the context of the activities of an establishment of the controller on the territory of the Member State. Accordingly the processing of personal data underlies the national law of the country in which the controller is established, regardless of where the actual processing takes place. The EU data protection laws already apply when the data processing takes place "in the context of the activities" of an establishment. Hence, it is not required that the processing is carried out by the establishment in a Member State. When the same controller is established on the territory of several Member States, he/she must take the necessary measures to

[18] Article 29 Data Protection Working Party, Opinion 4/2007, p. 17.

[19] Sect. 3 para. 6 BDSG: „Anonymising means altering personal data so that the individual information about personal or factual relations cannot be attributed to an identified or identifiable natural person, or this can only be done with an unreasonably large expenditure of time, cost and effort.

[20] See Art. 2 k of the Council Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters.

ensure that each of these establishments complies with the obligations laid down by the applicable national law.[21] Moreover the national law of the Member State shall apply when the data controller is not established on the Member State's territory, but in a place where its national law applies by virtue of international public law.[22]

The key provision for the applicability of the EU data protection laws to data controllers not residing within the territory of the EU is Art. 4 para. 1 lit. c of Directive 95/46/EC. Accordingly the EU data protection laws also apply when the "controller is not established on Community territory and, for purposes of processing personal data makes use of equipment, automated or otherwise, situated on the territory of the said member state, unless such equipment is used only for purposes of transit through the territory of the Community." By connecting the applicable law to the location of the equipment used for the processing, the Directive still applies the territoriality principle. In order to ensure that the data subjects can effectively exercise their data protection rights against a controller not residing in the EU Directive 95/46/EC provides in Art. 4 para. 2 that a non-EU controller that uses equipment on Community territory must designate a representative established on the territory of the relevant Member State. Accordingly the data protection provisions of the Directive are to be respected by project partners residing outside the EU (like Japan) whenever the data are processed over equipment located within the EU.

The processing of data outside the EU by a data controller not residing in the EU, however, does not fall in the scope of the Directive. In this case the national data protection law of the non-EU country will apply. In order to ensure that the data protection regime of Directive 95/46/EC cannot be circumvented by a transfer of the data to a third country, the Directive provides in its Art. 25 and 26 Directive 95/46/EC that such a transfer is only licit when an equivalent level of protection is provided for by the applicable national law or contractual agreements between the data exporter and the recipient of the data.[23]

4.2.1.3 Requirements for the fair and lawful processing of data

The term "processing of personal data" (processing) is extraordinarily broad, covering "any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction". Hence, this definition includes virtually any operation performed on personal data. Accordingly, also the anonymisation of personal data is to be regarded as processing of personal data.

Directive 95/46/EC establishes several principles regarding the fair and lawful processing of data (Art. 6), sets out criteria for the legitimate processing of personal data (Art. 7 and 8) and provides for technical and organisational measures to ensure the security and safety of personal data (Art. 17).

4.2.1.3.1 Principles set up for the processing of personal data

Art. 6 Directive 95/46/EC establishes several principles for the processing of data. These are the principles of legitimacy, purpose limitation, transparency, proportionality, security and control. In the following these principles shall briefly be outlined.

[21] Art. 4 para. 1 lit. a Directive 95/46/EC.

[22] Art. 4 para. 1 lit. b Directive 95/46/EC.

[23] For further information see chapter 7.3.10 of this deliverable.

The principle of legitimacy (Art. 6 para. 1 lit. a) generally requires first of all that data processing may only be conducted with a legal basis and in compliance with all legal requirements, and second, that the rights to data protection must be balanced against the interests of others in processing the data.

The purpose limitation principle derives from Art. 6 para. 1 lit. b of the Directive, which provides that personal data "must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes." The principle has two components:

(1) the data controller must specifically inform the data subject of the purposes for which data are being collected, and

(2) once they have been properly collected, the data must not be used for further purposes incompatible with the original purposes.

In general the further processing of data will only be found to be compatible if it is closely connected to the original purpose.[24] This strict limitation of purpose might pose a problem to research projects in which data are made accessible to other partners that may use the data also for other purposes than those the data had been collected for. This would, for instance, be the case if data that had been collected for the purposes of a specific research in the field of cancer shall be used for research in another field of medical research. However, the Directive provides that the further processing of data for historical, statistical or scientific purposes shall not be considered as incompatible provided that Member States provide appropriate safeguards.

In addition, personal data must be adequate, relevant and not excessive in relation to the purposes for which they are collected and further processed (proportionality principle). This requires the evaluation of proportionality, taking into account the risks for the protection of fundamental rights and freedoms of individuals and notably whether or not the intended purpose could be achieved in a less intrusive way.

Finally, personal data has to be kept "in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed."[25] Here again the Directive provides for the possibility to store data for longer periods for historical, statistical or scientific use. The Member States are obligated to lay down appropriate safeguards.

4.2.1.3.2 Criteria for legitimate data processing

As a general rule, the processing of personal data is prohibited, unless one of the exemptions listed in this provision applies. Summarising these exemptions, it can be said that, according to Art. 7, the processing of personal data is permitted, if the data subject has given his/her consent, or if the processing occurs in his/her interest or in the public interest. However, the processing of personal data is limited by the fundamental rights and freedoms of the data subject. This also is reflected in the basic principle of purpose specification.[26]

Furthermore, Art. 8 of Directive 95/46/EC grants special protections to "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life." These data are also referred to as „sensitive data". As the data processed within p-

[24] Kuner, European Data Protection Law, rec. 2.90.

[25] Art. 6 para. 1 lit. e Directive 95/46/EC.

[26] The purposes for processing of personal data must be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed. Personal data must not be further processed in a way incompatible with those purposes.

medicine will mostly contain genetic data and medical patient data, that are to be considered as health data in the sense of the Directive, this provision has to be considered carefully.[27] According to Art. 8 para. 1 such "sensitive data" may not be processed, except under certain clearly defined circumstances provided by para. 2 to 5.

According to para. 2 "sensitive data" shall only be processed, if

a) the data subject has given his explicit consent to the processing of those data or

b) the processing is necessary for the purposes in the field of employment law or

c) it is necessary to protect the vital interests of the data subject or

d) processing is carried out in the course of its legitimate activities with appropriate guarantees by a foundation, association or any other non-profit-seeking body with a political, philosophical, religious or trade-union aim or

e) the processing relates to data which are manifestly made public by the data subject or is necessary for the establishment, exercise or defence of legal claims."

According to Art. 8 para. 2 the processing of health data would be legitimate, if the data subject has given his/her consent. Furthermore, in many cases the collection of personal patient data will be covered by the vital interest of the patient, as it is needed for the purposes of medical treatment. In this case the collection of data in the context of medical treatment is covered by the exemption provided by para. 2 lit. c of Art. 8. The research to be carried out within p-medicine will help clinicians to find possible ways of treatment for patients. The processing of data in this context, however, will not be in the vital interest of the patients who´s data are processed, but in that of the concrete patient. Therefore, this exemption does not apply to the processing of data within the p-medicine infrastructure. As a consequence, with respect to the scientific research that is planed to be carried out in the course of p-medicine only the first exemption is applicable.

Art. 8 para. 3 provides exemptions in case of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services and where those data are processed by a health professional subject under national law. The collection and processing of data may be effected under this exemption. However, it is required for a concrete treatment of a concrete patient (data subject). Hence Art. 8 para. 3 does not provide an exemption to the general prohibition of the processing of sensitive data. Therefore the processing of data could only be based on this exemption if it is carried out for the purpose of the medical treatment of the patient. The processing of data of a specific patient for the purpose of medical treatment of other patients as it is envisaged for p-medicine is however not covered by this exemption.

However, Art. 8 para. 4 of the Directive provides that Member States may lay down additional exemptions by national law or by decision of the supervisory authority for reasons of substantial public interest, subject to the provision of suitable safeguards. Art. 8 para. 4 therefore does not provide an exemption per se, but empowers Member States to introduce national exemptions for reasons of substantial public interest and subject to the provision of suitable safeguards. Examples for such a substantial public interest are introduced by Recital (34) of the Directive:

"Whereas Member States must also be authorized, when justified by grounds of important public interest, to derogate from the prohibition on processing sensitive

[27] See Working Document of the Art. 29 Data Protection Working Party: Working Document on Genetic Data, p. 5, available at: http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2004/wp91_en.pdf.

categories of data where important reasons of public interest so justify in areas such as public health and social protection - especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system - scientific research and government statistics; whereas it is incumbent on them, however, to provide specific and suitable safeguards so as to protect the fundamental rights and the privacy of individuals;"

According to Recital (34) scientific research as it will be undertaken in the course of p-medicine is a possible example for an important public interest in the meaning of Directive 95/46/EC. Member States can therefore introduce regulation permitting the processing of sensitive personal data for scientific research purposes under the condition to provide specific and suitable safeguards so as to protect the fundamental rights and the privacy of individuals. The exemptions can be introduced either by national law or by decision of the supervisory authority. However, the disadvantage of para. 4 for European scientific research projects is that it is the free choice of each Member State whether it decides to introduce such exemptions in his/her national law at all and if so, under which preconditions they are introduced. Therefore, the conditions under which the data may be used in within the p-medicine infrastructure under this exemption may vary within the countries of the EU and may also be subject to changes. As a consequence the respective national legislation of all the countries involved will have to be constantly revised, in order to react to possible changes of the national laws.

It has to be noted that these exceptions are only hypotheses where the legitimacy of the data processing is formally assumed. The legitimacy of the processing of sensitive data is not given when only formally fitting into one of these exemptions. In fact the balance of the interests deriving from the processing of sensitive data has to be assessed in every concrete case.

4.2.1.3.3 Technical and organisational measures and Security policy

Directive 95/46/EC states in its Art. 17 that Member States shall provide that the controller/processor must implement appropriate technical and organisational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing.

Technical measures deal with the practical use of the methods implemented to secure the data being processed (including prevention of physical access to the hardware, such as secure premises and access control).[28] They include the use of encryption, secure connections, firewalls or access by biometric identification or similar methods. Organisational measures refer to a set of rules to enable data security by regulating authorization and authentication procedures, i.e. access policies and identity management for the IT system processing the data.[29]

Which security measures have to be provided in a specific case depends on various factors, one being the "state of the art". Directive 95/46/EC does not regulate what

[28] Terstegge, Directive 95/46/EC, Art. 17 no. 1, in: Büllesbach/Poullet/Prins, Concise European IT Law, Alphen aan den Rijn 2006.

[29] The Institute for Legal Informatics of the Leibniz University Hanover provides a sound examination of data security measures in EU FP7 framework project OPTIMIS dealing with appropriate technical and organisational measures to protect personal data to be taken in the field of cloud computing. See OPTIMIS, D7.2.1.2 – Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, pp. 17 ss., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

precisely „state of the art“ technology is. Art. 17 of the Directive does neither provide for a specific security system nor does it refer to existing security standards. This has various reasons, the most important being the fact that data security is (to a large extent) a technical matter. As the available data security technology constantly changes due to the technical development in the field, it was necessary to choose a technology neutral approach. Accordingly the Directive does not provide for specific security measures but rather contains quite general data security requirements. Furthermore the data security measures to be applied shall reflect the needs of a large variety of different data controllers/processors. Hence the Directive acknowledges that the data security requirements can vary largely depending on the organisation involved by stating that an appropriate level of data protection has to be established. Accordingly a certain standard appropriate for a large organisation may not be suitable for a small organisation, and vice versa.[30]

For guidance on specific implementation of data security concepts, Directive 95/46/EC relies on the domains of computer science and information security. In principal, it is therefore up to the developers to decide which security measures are to be considered "state of the art".[31] In the absence of detailed security requirements Directive 95/46/EC, one generally accepted information security standard is "ISO/IEC 27001:2005 Information technology – Security techniques – Information security management systems".[32] This standard specifies requirements for the establishment, implementation, monitoring and review, maintenance and improvement of a management system for managing an organisation's information security risks.[33] The standard is part of the ISO 27000 series, which contains a different number of standards, which can be can be categorised in three groups: The ISO/IEC 27000 contains the fundamentals and vocabulary, providing an overview of the ISO 27000 series. ISO/IEC 27001 sets up normative requirements for the development and operation of information security management systems, providing a set of (customisable) security controls and mitigation of the risks associated with the information, which the organisation seeks to protect. The remaining standards (ISO/IEC 27002 to ISO/IEC 27007) contain guidance standards or guidelines, good practice and methodologies. ISO/IEC 27002 provides specific implementation advice and guidance on best practice in support of the security controls specified in ISO/IEC 27001.[34]

The data controller, however, is in general not obliged to provide the highest security standards according to the state of the art. Security measures rather have to be appropriate with regard to the anticipated risks inherent in the data processing, as well

[30] C.f. OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, p. 61 s., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

[31] See OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, pp. 27 ss., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

[32] Available at: http://www.iso.org/iso/catalogue_detail?csnumber=42103.

[33] http://www.iso.org/iso/catalogue_detail?csnumber=42103. For further information to this and to other standards please see OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, pp. 61 ss., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

[34] See also OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, pp. 61 s., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

as with regard to the nature of data and the costs of their implementation.[35] When assessing appropriate security measures for p-medicine, the potential risks for the patients related to the collection and processing of their (genetic) data as well as the technical and organisational measures to prevent these risks have to be thoroughly evaluated. As mostly sensitive data in the sense Art. 8 para. 1 Directive 95/46/EC are processed in p-medicine a high data security level will be required. Due to the technical development the state of the art can change over the time, so that the security measures undertaken shall be reviewed periodically by the data controller by re-assessing the risks as well as the technical possibilities available to prevent these risks.

Further guidance with regard to appropriate security measures can be found in the Recommendation R (97) 5 of the Committee of Ministers of the Council of Europe on the Protection of Medical Data that also calls for appropriate technical and organisational measures to be taken to protect personal data in its section 9.[36] The Recommendation gives examples for appropriate technical and organisational measures that ensure the confidentiality, integrity and accuracy of processed data. These include, among others, access, transmission, input, job and availability control.

Access control is used to prevent unauthorised persons from gaining access to data processing systems with which personal data are processed or used, to prevent data processing systems from being used without authorisation, to ensure that persons entitled to use a data processing system have access only to the data to which they have a right of access and that personal data cannot be read, copied, modified or removed without authorisation in the course of processing or use and after storage.[37] Transmission control ensures that personal data cannot be read, copied, modified or removed without authorisation during electronic transmission or transport and that it is possible to check and establish to which bodies the transfer of personal data by means of data transmission facilities is envisaged.[38] Input control ensures that it is possible to check and establish whether and by whom personal data have been submitted to data processing systems, modified or removed. Job control ensures that, in the case of commissioned processing of personal data, the data are processed strictly in accordance with the instructions of the principal. Availability control ensures that personal data are protected from accidental destruction or loss. Lastly, it is necessary to ensure that data collected for different purposes can be processed separately.[39]

4.2.1.3.4 Data controller

Directive distinguishes between the controller and the processor of data. The term „controller“ is defined in Art. 2 lit. d Directive 95/46/EC as “the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the

[35] Terstegge, Directive 95/46/EC, Art. 17 no. 2, in: Büllesbach/Poullet/Prins, Concise European IT Law, Alphen aan den Rijn 2006.

[36] http://www.coe.int/t/e/legal_affairs/legal_co-operation/data_protection/documents/international%20legal%20instruments/1Rec(97)5_EN.pdf.

[37] See also OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, pp. 61 s., available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

[38] See OPTIMIS, D7.2.1.2, Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation, p. 35, available at: http://www.optimis-project.eu/sites/default/files/content-files/document/d7212-optimis-cloud-legal-guidelines.pdf.

[39] See also ACGT, D10.2, The ACGT ethical and legal requirements, p. 74, available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf.

controller or the specific criteria for his/her nomination may be designated by national or Community law.“ The term „processor“ defined in Art. 2 lit. e Directive 95/46/EG means „a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller“.

The classification of an actor as a „controller“ or „processor“ has important consequences as most of the data protection obligations under the Directive have to met by the data controller, who is moreover liable for data protection violations. The data processors, however, have a severely reduced role, as they are supposed only to process data as directed by the controller.

According to Art. 6 para. 2 Directive 95/46/EC the data controller has the duty to ensure that personal data are processed fairly and lawfully. Consequently, the controller has to safeguard that personal data are only collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. Moreover, the data controller has to warrant that the data are not excessive in relation to the purposes for which they are collected and/or further processed. Moreover, every reasonable step must be taken by the data controller to ensure that data, which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified. Likewise, the data controller has to ensure that the data are kept in a form, which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. Further, the data controller has to implement appropriate security measures in the sense of Art. 17 Directive 95/46/EC.[40]

Since it is the data controller who is liable for the legality of data processing and the fulfilment of the obligations towards the national data protection authority and the data subjects, it is essential that the data controller is always identifiable. Thus, Articles 10 and 11 of Directive 95/46/EC state that the data controller must provide a data subject from whom data are collected the identity of the controller and of his/her representative, the purpose of the processing as well as any recipients of the data. Furthermore, Article 12 establishes that the data controller has to guarantee every data subject the right to obtain information about the processing of his/her data.

In case the data controller fails to fulfil his/her duties in accordance with the Directive and thus fails to respect the rights of data subjects, Article 23 of Directive 95/46/EC states that “*any person who has suffered damage as a result of an unlawful processing operation or of any act incompatible with the national provisions adopted pursuant to the Directive is entitled to receive compensation from the data controller for the damage suffered*”. However, the data controller may be exempted from this liability, in whole or in part, if he/she proves that he/she is not responsible for the event giving rise to the damage.

4.2.1.3.5 Trusted Third Party

In the broadest sense of the term a Trusted Third Party (TTP) is a party who two other parties trust. The concept of using a TTP for the delivery of security services is well established and is already being implemented widely for delivering Public Key Infrastructure (PKI) services. Through this type of service, key-pairs for digital signatures, authentication or encryption are already being managed.[41] However, the activities of Trusted Third Parties are quite diverse and not limited to PKIs. TTPs may

[40] For detailled information to Art. 17 Directive 95/46/EC see chapter 4.2.1.3.3 above.

[41] See e.g. Directive 99/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, published in Official Journal L 13, 19/01/2000, p. 12.

also serve as registration authorities, certification authorities, validation or time-stamping authorities are some examples of other TTP services.[42] All these different types of TTPs have in common that they provide as independent intermediaries "trust services" to other parties.

In the context of data protection a Trusted Third Party (TTP) is regarded as a trustful custodian for personal data or the links to identify the concerned data subject, which shall ensure the privacy of the concerned data subject. When the security solution is based on pseudonymisation, the trustee is a pseudonymisation TTP. Pseudonymisation, meaning the withdrawal of the true identity of a person and their replacement with a pseudonym, is often used in research projects in order to protect the privacy of the persons concerned. Contrary to simple anonymisation, pseudonymisation still enables the linkage of data associated to the pseudonym back to the person. Secure pseudonymisation can only be performed when this linking information is stored at a trust service provider. In this context a TTP is regarded as a trustful custodian for personal data or the links to identify the concerned data subject, which shall ensure the privacy of the concerned data subject.

The main features of a TTP are its strict independence as an organisation, the trustworthiness of its methods, its adherence to the principles of open- ness and transparency regarding its methods, and the provision of professional expertise related to the domain of relevance. Further the TTP has to provide project-specific security policies, secure modules, platforms and infrastructure as well as monitoring and quality assessment activities, documentation, operating reporting and auditing systems and strict code of conducts, trust practice statement and secrecy agreement policy.[43]

The TTP must be bound to professional discretion to protect the links of the participating patients sufficiently. The links have to be protected by the TTP against (unlawful) access and also against seizure. It is the duty of the TTP as an independent data controller to provide adequate technical and organisational measures.[44]

4.2.2 Directive 2001/20/EC

Directive 2001/20/EC[45], also known as Clinical Trials Directive, states the requirements for the conduct of clinical trials in the EU. It is concretised by the Commission Directive 2005/28/EC[46] laying down principles and detailed guidelines for good clinical practice as

[42] C.f. De Moor/Claerhout/De Meyer, Privacy Enhancing Techniques, The Key to Secure Communication and Management of Clinical and Genomic Data, Methods Inf Med 2/2003, p. 150, available at: http://www.schattauer.de/en/magazine/subject-areas/journals-a-z/methods/contents/archive/issue/694/manuscript/271/download.html.

[43] De Moor/Claerhout/De Meyer, Privacy Enhancing Techniques, The Key to Secure Communication and Management of Clinical and Genomic Data, Methods Inf Med 2/2003, p. 149, available at: http://www.schattauer.de/en/magazine/subject-areas/journals-a-z/methods/contents/archive/issue/694/manuscript/271/download.html.

[44] See also chapter 7.3.3 below.

[45] Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use, published in Official Journal L 121, 1.5.2001, p. 34.

[46] Commission Directive 2005/28/EC of 8 April 2005 laying down principles and detailed guidelines for good clinical practice as regards investigational medicinal products for human use, as well as the requirements for authorisation of the manufacturing or importation of such products (Text with EEA relevance), published in Official Journal L 91, 9.4.2005, p. 13.

regards investigational medicinal products for human use, as well as the requirements for authorisation of the manufacturing or importation of such products (Good Clinical Practice Directive).

By adopting this Directive the European legislator aimed to facilitate the performance of multi-national clinical trials in Europe through the harmonisation of the regulatory procedures. It, however, only provides for minimum standards the national rules transposing the Directive in the Member States can vary considerably.

This subchapter shall outline the scope and the content of this Directive as far as relevant for p-medicine.

4.2.2.1 Scope of the Directive

The Directive 2001/20/EC „establishes specific provisions regarding the conduct of clinical trials, including multi-centre trials, on human subjects involving medicinal products ..., in particular relating to the implementation of good clinical practice“.[47] It is applicable to clinical trials performed in the European Union.

The term "clinical trial" is defined by Art. 2 lit. a of the Directive. It covers „any investigation in human subjects intended to discover or verify the clinical, pharmacological and/or other pharmacodynamic effects of one or more investigational medicinal product(s), and/or to identify any adverse reactions to one or more investigational medicinal product(s) and/or to study absorption, distribution, metabolism and excretion of one or more investigational medicinal product(s) with the object of ascertaining its (their) safety and/or efficacy“.[48]

Directive 2001/20/EC, however, does not apply to non-interventional trials,[49] meaning studies „where the medicinal product(s) is (are) prescribed in the usual manner in accordance with the terms of the marketing authorisation. The assignment of the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice and the prescription of the medicine is clearly separated from the decision to include the patient in the study. No additional diagnostic or monitoring procedures shall be applied to the patients and epidemiological methods shall be used for the analysis of collected data“.[50]

Thus Directive 2001/20/EC applies to trials on medicinal products, covering trials on pharmaceuticals, which are in development, and "investigational medicinal products". The Directive does not apply to other health related research, such as physiological research, research into medical devices, observational studies, research into tissue, organs or blood, or embryo research.[51]

When evaluating the applicability of this Directive for clinical trials carried out in p-medicine it has to be clarified whether the assignment of the patient to a particular therapeutic strategy falls within the current practice and the prescription of the medicine is clearly separated from the decision to include the patient in the study, which would exclude the applicability of Directive 2001/20/EC. The Directive is however applicable,

[47] Art. 1 para. 1 Directive 2001/20/EC.

[48] An "investigational medicinal product" according to Art. 2 lit. d Directive 2001/20/EC is „a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical trial, including products already with a marketing authorisation but used or assembled (formulated or packaged) in a way different from the authorised form, or when used for an unauthorised indication, or when used to gain further information about the authorised form“.

[49] Art. 1 para. 1 Directive 2001/20/EC.

[50] Art. 2 lit. c Directive 2001/20/EC.

[51] Hervey/McHale, Health law and the European Union, p. 251.

if the patient is included in a certain study to test the effects of a certain therapy or medical product.

Directive 2001/20/EC does not affect the data protection regime provided by Directive 95/46/EC.[52]

4.2.2.2 Protection of clinical trial subjects

According to the minimum standards established by Directive 2001/20/EC a clinical trial may only be undertaken, if the foreseeable risks and inconveniences have been weighed against the anticipated benefit for the individual patient concerned and other present and future patients. It may be initiated only if the Ethics Committee and/or the competent authority comes to the conclusion that the anticipated therapeutic and public health benefits justify the risks and may be continued only if compliance with this requirement is permanently monitored (lit. a).

Furthermore the inclusion of a patient in a clinical trial shall generally be based on the patient´s informed consent. Patients incapable of giving legal consent to clinical trials shall however not be excluded from clinical trials, but should rather be given special protection,[53] as the Directive recognises that there is a need for clinical trials involving children to improve the treatment available to them. Accordingly the clinical trial on children shall only be effected if the medicinal product would be of direct benefit to the patient, thereby outweighing the risks. When the patient is not able to give informed consent, his/her legal representative has had the opportunity, in a prior interview with the investigator or a member of the investigating team, to understand the objectives, risks and inconveniences of the trial, and the conditions under which it is to be conducted and has also been informed of his/her right to withdraw from the trial at any time.[54]

In addition that Directive provides that Member States shall regulate that the patient may without any resulting detriment withdraw from the clinical trial at any time by revoking his/her informed consent. Further, the rights of the patient to physical and mental integrity, to privacy and to the protection of the data concerning him/her in accordance with Directive 95/46/EC shall be safeguarded. Finally the Directive regulates that the patient shall be provided with a contact point where he/she may obtain further information.[55]

4.2.2.3 Further Guidelines for Clinical Trials

Certain aspects of the Directive 2001/20/EC, such as the requirements regarding Good Clinical Practice, including the documentation, of the clinical trials, the information to be submitted to the competent authorities and to the ethics committees, the requirements on safety monitoring and the reporting of adverse reactions or the inspections of competent authorities and the applicable procedures, are specified by several guidelines.

[52] Recital 17 Directive 2001/20/EC.

[53] Recital 3 Directive 2001/20/EC.

[54] The current situation concerning the legal, ethical, technical and clinical handling of consent, mainly in European projects dealing with vulnerable patient groups is currently being analysed in the EU project CONTRACT, Consent in a trial and care environment. The purpose is to identify existing practices and problems encountered in translational research throughout Europe. The Institute for Legal Informatics of the Leibniz University of Hanover is elaborating solutions to tackle the obstacles hindering the legal and ethically correct exchange of personal data based on informed consent in running and planned European projects. Further information can be found on the CONTRACT website: http://www.contract-fp7.eu as well as at: http://www.iri.uni-hannover.de/contract-1634.html.

[55] Art. 3 para. 4 Directive 2001/20/EC.

The European Commission issues further guidelines in Volume 10 of Eudralex.[56] The European Medicines Agency (EMA) has also published guidelines of a similar nature. To summarise, these guidelines mainly cover the area of inspection procedures and guidance for Good Clinical Practice inspections carried out in the context of the centralised procedure requirements relating to the quality, safety and efficacy of products, as well as specific types of products.

4.2.3 Directive 2001/83/EC

Furthermore the Directive 2001/83/EC on the Community code relating to medicinal products for human use has to be taken into account. This Directive deals with the disparities between certain national provisions, in particular between provisions relating to medicinal products. A medicinal product is defined as any substance or combination of substances presented for treating or preventing disease in human beings or any substance or combination of substances which may be used in or administered to human beings either with a view to restoring, correcting or modifying physiological functions by exerting a pharmacological, immunological or metabolic action, or to making a medical diagnosis.[57] The Directive provides inter alia rules regarding the placing of medicinal products on the market, such as marketing authorisation, specific provisions applicable to homeopathic medicinal products, procedures relevant to the marketing authorisation. Furthermore the Directive regulates the manufacture and importation of medicinal products, their labelling and the package leaflet as well as the wholesale distribution of and advertisement for medicinal products.

Directive 2001/83/EC however only applies to industrially produced medicinal products for human use intended to be placed on the market in Member States.[58] Art. 3 of the Directive provides several limitations of the scope. Accordingly the Directive does inter alia not cover medicinal products intended for research and development trials. Furthermore the Directive does not apply to whole blood, plasma or blood cells of human origin. Hence this Directive does not apply to the use of medicinal products within p-medicine if these products are merely used for research and development trials. In the event that industrially produced medicinal products for human use intended to be placed on the market shall be tested the Directive provides inter alia for a clinical documentation.[59] This documentation will, however, have to be effected on a clinical level and not within the research infrastructure provided by p-medicine.

4.3 Ethical requirements

This subchapter shall highlight some ethical requirements for p-medicine with a special emphasis on informed consent and patient empowerment.

Informed consent and patient empowerment both derive from the ethical understanding that in the context of medical treatment and/or research involving patients, their rights of autonomy and self-determination[60] are to be guaranteed and highly protected.

[56] For further information please see: http://ec.europa.eu/health/documents/eudralex/vol-10.

[57] Art. 1 para. 2 Directive 2001/83/EC as amended by Directive 2004/27/EC.

[58] Art. 2 Directive 2001/83/EC.

[59] See in particular Part 4 of Annex 1 to Directive 2011/83/EC.

[60] McLean, Autonomy, consent and the law, in: Taylor/Francis, 2009, 40; Leino-Kilpi, Patient's autonomy, privacy, and informed consent, p.55.

Autonomy is thought to be part of the liberal western tradition granting special importance to individual freedom and choice, both for political life and for personal development.[61] This autonomy allows humans to take their lives in own hands and make decisions on their own account. The "principal legal mechanism through which the right of autonomy has been delivered is informed consent".[62]

In other words this means that no medical intervention, be it for purposes of prevention, diagnosis, therapy or research, is to be undertaken without patient's prior information and consent.

4.3.1 Informed consent

Similar to the legal concept of informed consent, the ethical approach provides specific requirements or preconditions in order to classify an informed consent as valid.[63]

The three fundamental preconditions are (1) informed, (2) voluntarily and (3) capable to take decisions.

With respect to the requirement "informed" the Declaration of Helsinki determines that participants in research projects should be provided with information on "the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail. [In addition t]he subject should be informed of the right to abstain from participation in the study or to withdraw consent to participate at any time without reprisal."[64] What information has to be provided specifically, in the meaning of how much information is needed, depends on different criteria of which the most important one is the capability of the respective patient. The information should therefore be not only provided, but also provided in a way that the respective patient or participant can *understand* and is not overwhelmed by the information – that is why the relevant information necessary for a valid consent will differ in each particular case.

Consent likewise has to be given "voluntarily" or "freely" to be valid. This means "that the individual is free from external pressure to make a particular decision"[65]. External pressure is to be assumed where consent is given under coercion or duress, under pressure, or under manipulation or undue influence[66].

Last but not least the patient must be – in general - capable to take decisions. Decisional capacity is discussed in both law, as well as in ethics. Two levels of discussion can be recognised – the theoretical legal framework and the practical assessment of capacity. The practical assessment however covers a not fixed set of abilities[67], which prove decisional capacity or their lack. Those agreed on however are: understanding given information (can understand the physician's advice, the treatment and research options

[61] Faden/Beauchamp, The History and Theory of Informed Consent, p. 7.

[62] Faden/Beauchamp, The History and Theory of Informed Consent, p. 7.

[63] This issue is also discussed in the European FP 7 Project CONTRACT, http://www.contract-fp7.eu/site.

[64] Art 13 of the World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects, available online on the WMA webpage: http://www.wma.net/en/30publications/10policies/b3/index.html .

[65] McLean, Autonomy, consent and the law, p. 51.

[66] Brock, Life and death, p. 43.

[67] Appelbaum/Grisso, Assessing patients' capacities to consent to treatment, The New England Journal of Medicine 319, no. 25, p. 1635-1638.

and alternatives), appreciation of the situation (ability to judge with a set of important values, how a particular decision will help in achieving what the patient believes is good for him/her), reasoning (understand causal relations, probability and percentages, as well as logical reasoning) and finally making and communicating the decision taken.[68]

4.3.2 Feed back

Besides such ethical requirements with respect to the informed consent of participating patients, in p-medicine the feedback to the patient in case of individually important findings in the course of the project is a crucial ethical requirement and challenge.

The first question to decide on is the type of data to feed back. The relevance of research results is not easy to define as such results are usually characterised by a lack of established common interpretation and independent validation or may change as data become more reliable. As such information could be helpful as well as harmful for patients, they could be harmed by being excluded from individual information as well as by being provided with it.[69] Therefore, taking into account patients' right to know or not to know, from an ethical point of view it is recommended to give patients the option, to decide whether they would like to have feedback.

Another issue to consider is who the addressee of such feedback should be. Especially when it comes to genetic research or research on tissue, as in p-medicine, such research may reveal data of diseases that are of predictive nature for future diseases and relevant not only for the patient concerned, but also for his/her blood relatives. There is a controversial discussion whether such genetic information is the most private information of all,[70] with the consequence that only the respective patient is to inform or whether such information is a family affair, with respective information requirements[71].

However, as far as genetic information on cancer is concerned, individuals' right generally overweight interests of family members. Only "where there is a high risk of having or transmitting a serious disorder and prevention or treatment is available, immediate relatives should have access to stored DNA for the purpose of learning their own status".[72] Since genetic research on cancer usually provides only moderate predictive results, it seems more likely to increase emotional and psychological distress by healthy family members providing them with research findings than by not disclosing them.[73] Feedback in p-medicine therefore will only be given to the patient concerned and only, if he/she agreed to such feedback.

[68] Appelbaum, Assessment of Patients' Competence to Consent to Treatment, New England Journal of Medicine 357, no. 18, p. 1836.

[69] See also ACGT, D10.2, The ACGT ethical and legal requirements, pp. 41 ss., available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf.

[70] Clarke et al., Genetic professionals' reports of nondisclosure of genetic risk information within families, European Journal of Human Genetics, 13 (5), p. 561.

[71] Miyata et al., Disclosure of cancer diagnosis and prognosis: a survey of the general public's attitudes towards doctors and family holding discretionary powers, p. 2, available at www.biomedcentral.com/1472-6939/5/7.

[72] HUGO Ethics Committee, Statement on DNA sampling: control and access, February 1998, www.hugo-international.org/Statement_on_DNA_Sampling.htm, p. 2.

[73] See also ACGT, D10.2, The ACGT ethical and legal requirements, p. 49, available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf.

5 Evaluation of the data protection and data security framework set up in ACGT

5.1 Introduction

In this chapter the data protection and data security framework that had been set up in ACGT shall be evaluated. This analysis shall identify the strengths and the weakness of the existing framework and thus serve as a basis for the further development of the new framework that has to be set up according to the needs and requirements of p-medicine.

We will first give a rough outline about the legal framework set up for ACGT describing the concept of de-facto anonymous data, the involvement of a Trusted Third Party (TTP), the role of the Center for Data Protection (CDP) and the contractual agreements that have been designed in order to guarantee compliance of the project partners to the legal data protection and data security requirements. The evaluation takes into account the level of data protection achieved by the data protection and the data security framework, the practical implementation of the framework (establishment of the CDP, conclusion of the contracts, technical implementation), compliance with the legal framework, compliance with technical standards, patient involvement as well as ethical considerations.

In a second step the data security framework set up in ACGT will be evaluated.

5.2 Overview of the legal framework within ACGT

Similar to p-medicine, in ACGT the overall goal was to improve medical research by providing an IT-infrastructure that simplifies access to, exchange of and research on patient data throughout Europe. To achieve this goal while complying with data protection provisions and safeguarding patients' rights the following preconditions were defined:

The basic assumption of the whole data protection framework was that the best way to safeguard patients' rights would be achieved, if only anonymous data were processed in the project. This however led to the difficulty, that anonymisation of genetic data in the strict meaning (nobody can re-identify the respective patient) is hardly possible.[74] Besides strictly anonymous data did not meet the needs of the project, as a feedback procedure to the patient was required in order to let patients participate in possible medical findings of the project.

Hence the ACGT data protection and ethics framework was designed as a 'safety net' consisting of three pillars. The first pillar was the design of the infrastructure, ensuring that the patient data processed would be estimated as de-facto anonymous data. The second pillar ensured patient involvement. This was achieved by obligatory patient information and the requirement of informed consent as well as the establishment of a central contact point for all participation patients in the project. Thirdly as last pillar national legislation was analysed in order to identify provisions that allow the processing of personal data for research purposes.

Huge effort was invested in the first pillar to build up a "Network of Trust" within the project that provided a "context of anonymity". Such 'context of anonymity' was achieved by the establishment of a legal body taking over the responsibility of the data controller within the

[74] ACGT, D10.2, The ACGT ethical and legal requirements, .pp.107 ss., available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf.

project, the involvement of a Trusted Third Party, the conclusion of legally binding contracts and a security infrastructure.[75]

5.2.1 Data flow within ACGT

In order to be able to set up an innovative, useful and advanced IT-infrastructure for medical research already during the design it was crucial to work on real patient data as soon as possible. Hence when the framework was set up participating hospitals (data exporters) provided patient data to the consortium. Such patient data were submitted from the hospital database to a separate database located inside the respective hospitals. During submission such data were state-of-the-art pseudonymised by a toolkit named 'CAT', which was developed in the course of ACGT.[76] Thereafter such data were available through the ACGT infrastructure for participating partners.

Accompanied by contracts between the data exporter and the central data controller of the project (the Center for Data Protection 'CDP') the separated databases were under the control of the CDP. The keys, containing the link back to the patient, which were created during the pseudonymisation procedure, were stored at a Trusted Third Party 'TTP'. The TTP saved such keys for the sole purpose to enable the project to give feed back to a respective patient in case of new finding relevant for him/her.

Beside this all data processing inside of ACGT was conducted on so called de-facto anonymous patient data. The following figure illustrates the data flow (simplified, as it only shows one hospital providing data).

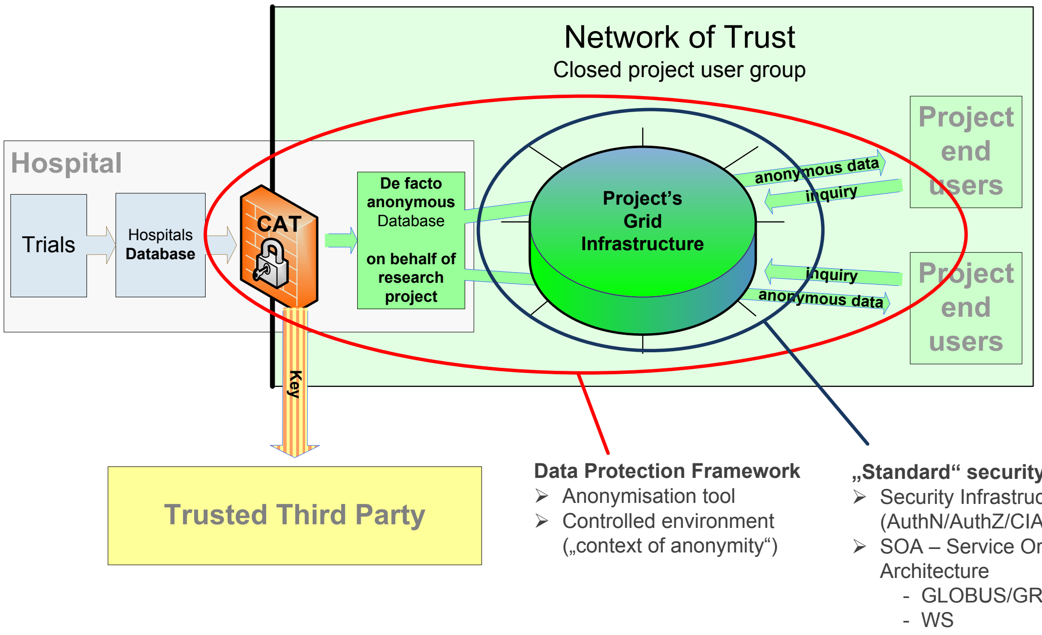
Network of Trust
Closed project user group
Hospital
Trials
Hospitals Database
CAT
De facto anonymous Database
on behalf of research project
Project's Grid Infrastructure
anonymous data
inquiry
inquiry
anonymous data
Project end users
Project end users
Key
Trusted Third Party
Data Protection Framework
➢ Anonymisation tool
➢ Controlled environment („context of anonymity")
„Standard" security
➢ Security Infrastructure (AuthN/AuthZ/CIA)
➢ SOA – Service Oriented Architecture
- GLOBUS/GRIDGE
- WS

Figure 1: ACGT Network of Trust

[75] An intense description of the whole framework can be found in: Forgó/Arning/Krügel/Kollek/Petersen, Ethical and Legal Requirements for Transnational Genetic Research, 2010.

[76] For further information on CAT see chapter 5.3.1.5 below.

5.2.2 Legal framework

5.2.2.1 Context of anonymity

According to Directive 95/46/EC the use data for the purpose of scientific research shall be carried out in an anonymous way, as soon as the research purpose allows it. The analysis of the data flows within ACGT showed, however, that the data to be transferred could not be anonymised. In order to give patients the best therapy the re-identification of each patient had to be guaranteed, so that the data had to keep an identifier linking to the patient (pseudonymous data). Furthermore the uniqueness of genetic data had to be taken into consideration. Every reference dataset makes it possible for third persons to match data and identify the patient. Thus, it is not possible to render genetic data completely anonymous, but rather *de facto* anonymous.

Considering this, a variety of technical and legal measures were taken in order to guarantee the processing of *de facto* anonymised data only within ACGT.

It was decided to conclude contracts between the project participants and the project consortium, providing amongst others data protection policies, clauses on liability as well as a contractual penalty clause, if patient data was unlawfully matched or disclosed. As a project consortium has no legal personality, the Center for Data Protection 'CDP', a Belgian non-profit organisation, was founded to act on behalf of the consortium. The CDP concluded the mentioned contracts with all participants providing and/or having access to patient data. Besides this a security infrastructure was set up (see in detail below under 5.3), including a pseudonymisation software and authentification services. When finalised, the 'context of anonymity' hereby established, provided an environment, in which the re-identification of participating patients from inside or outside the project was not possible with means likely reasonable to be used with respect to of time, expense and labour. From a legal point of view therefore only anonymous data, which is no longer subject to the data protection regime, were processed.[77]

5.2.2.2 Fall back scenarios

5.2.2.2.1 Consent

The major fall-back scenario for the unlikely case of a failure of that system of de facto anonymity was the second pillar of the framework, which called for obligatory patient information and a qualified informed consent of each patient to the processing of his/her data within the system. Hence for the unlikely event, that patient data were despite the measures taken not de-facto anonymous and therefore subject to the data protection legislation, the informed consent provided besides the crucial involvement of the patients a legal basis for the processing of such data in ACGT.

The respective consent modes derived from an ethical approach and were specified for each problematic case to make them legally applicable in a contractual fashion.

5.2.2.2.2 National exceptions

The third pillar, for the very unlikely case that both the other two pillars would have failed to adequately protect the patients, built on the various exemptions that have been introduced to national laws of the Member States of the European Union, according to Art. 8 of Directive 95/46/EC.

[77] See chapter 4.2.1.1.2 above.

5.2.3 Evaluation

This subsection will provide an evaluation as to in how far the framework set up for ACGT was successful and whether it fulfils the legal and ethical requirements that arise in the course of p-medicine (see 5.4). As mentioned above the evaluation takes into account the level of data protection achieved by the introduced framework, the practical implementation of the framework (establishment of the CDP, conclusion of the contracts, assignment of TTP, etc.), patient involvement and ethical considerations.

5.2.3.1 Data protection level

As the level of data protection needed in one or the other situation strongly depends on the risks that arise for the data subject in the course of the respective data processing, there is no common scale to estimate the level of data protection. It rather is a case by case decision taking into account the overall goal of the processing, the efforts in terms of expense, time and labour that data protection measures might cause as well as the data subjects interests.

Research in ACGT- as in p-medicine – was conducted on probably the most sensitive data possible, on genetic data. This alone emphasises that the framework had to serve the highest standards possible. Still, medical research, in particular research on widespread diseases as cancer, can only be carried out on genetic data. The challenge to cope with therefore was to find the right balance in order to not hinder such research while safeguarding patients' rights at best.

The framework therefore stepped in at a very early stage. While usually medical research projects address data protection issues by intense information and informed consents of the data subjects on the one and a secure architecture to the outside on the other side, by contrast the ACGT framework invested huge efforts in order to avoid the processing of personal data at all, while – only as second step – relay on information, consent and a secure architecture.

To manage this, the framework aimed for a so-called 'context of anonymity' inside the project. This was achieved by the conclusion of contracts and the involvement of a Trusted Third Party on the one hand and the development of a state-of-the-art pseudonymisation tool on the other hand. Such contracts set very restricted data protection policies whose compliance was flanked by contractual penalties.

Hence predominantly all data processing within ACGT was de-facto anonymous, which is the best way to protect patients' rights.

Anyhow informed consents and patient information were not neglected - although in general from a legal point of view not needed for the processing of anonymous data. Last but not least the CDP served as a central contact point with regard to data protection related issues/infringements for all participating patients. Patients therefore were provided with only one responsible body to talk to, preventing the burden of investigation who or which organisation processed the patient's data.

In summary the level of data protection within ACGT was on an advanced level both from a general point of view but also and especially in comparison with other medical research projects.

5.2.3.2 Practical implementation

The practical implementation however was much more complex than originally foreseen as different contracts had to be designed and concluded, the CDP had to be established and a TTP needed to be involved.

5.2.3.2.1 Center for Data Protection

The Center for Data Protection was founded as non-profit-organisation under Belgian law. This legal form was chosen as it didn't need any significant monetary resources, is not expensive with respect to its functioning but provides a legal personality. Still the organisation needed to be registered, which took time, and natural people inside of the consortium needed to be found, who agreed to become founder members and take the liability for the organisation personally, as insurance was not foreseen in the framework.

5.2.3.2.2 Contracts

For the framework three contracts (data exporter agreement, agreement with users, agreement with the TTP), a patient information sheet, General Terms and conditions and several informed consents were designed.

Especially the negotiation of the data exporter agreement but also of the agreement with the users of the platform turned out to be very time consuming. This had different reasons: First of all the contracting party on the data exporter side and sometimes even in the background of the user side is the legal department of the respective hospital. As the legal departments are not involved in the project, it needed joint effort by the legal partners of ACGT and the physicians of the respective hospital to outline the idea of the project and the necessity to sign such contracts.

Besides this given difficulty the contracts in addition were of great complexity. This was to some extent due to the multitude of mutual obligations of the contracting parties that needed to be codified. Besides various provisions that are obligatory in terms of data protection, such multitude was also caused by the data flow chosen in ACGT. Especially the location of the servers, containing de-facto anonymous data for the ACGT infrastructure inside the hospitals under the control of the CDP, caused a lot of legal as well as technical ‘extra obligation’ for the data exporter (hospital).

Finally but probably most important the contractual penalty clauses led to difficult negotiations, as the legal departments of the participating hospitals needed intensive convincing to sign a clause that initiates a contractual penalty of the hospital in case the respective physician infringes the contractual obligations. Negotiations therefore needed over a year of time in some cases.

The designed informed consent sheets, the patient information sheets and the general terms and conditions however didn’t lead to any difficulties with respect to signature. Anyhow, it showed that these documents in some parts were not easy to understand and very complex.

5.2.3.2.3 Patient Identity Management

In ACGT it appeared that the infrastructure did not provide any feature able to manage pseudonyms in order to prevent having multiple data sets of the same patient from different sources separated from each other in the database. This might lead to inconsistency of the database likely to cause incorrect research results on the one hand and difficulties with respect to the feedback procedure on the other hand.

5.2.3.2.4 Trusted Third Party

To involve a Trusted Third Party (TTP) didn’t show any difficulties. Contracts that need to be concluded with such data custodians are common and similar to each other. Hindering might have been that within the course of ACGT it was stated that the TTP necessarily needed to be somebody from outside the consortium to meet the condition ‘third’ party; this excluded technical partners from inside the consortium.

5.2.3.3 Compliance with the legal framework

Compliance with the legal framework was to be achieved by a broad audit right of the Center for Data protection as well as by including penalty clauses in the contracts to be concluded. However, legal compliance of such a large international research project remained a demanding necessity.

There were three major possible threats identified:

1.) a partner's violation of contract clauses could lead to a failure of the pseudonymisation procedure,
2.) informed consents of patients were missing or invalid (e.g. because the patient wasn't informed properly) or
3.) an ACGT end user could violate the concluded contracts by matching or disclosing patient data.

The risk of the first threat was considered to be very low as it ran fully automatically. The same, meaning low risk, was true for the second and third threat. Consents were invalid in case they were not given voluntarily or patients had not been informed properly. The same could happen if consents were too broad or unclear. However there were special consent forms created for ACGT and participating hospitals were bound by contract – and penalty - to use them. Matching or disclosure of data finally was prohibited by contract combined with a fairly high penalty and audit routines as well.

Compliance with the legal framework therefore was well achieved. Auditing however was crucial with respect to the complex framework.

5.2.3.4 Patient involvement

Besides the contractual framework that was achieved, the establishment of the CDP brought great advantages for the participating patients. As mentioned above, consortia of transnational European research projects don't have a legal personality. Data processing in such projects therefore brings up the problem that data transfer in the course of the project has to be controlled by the respective partners who are involved in such transfer. This makes it very difficult for patients to exercise their rights of information, access, rectification or erasure.

As well planed side affect, the establishment of a central data controller for all data processing within the project therefore was that patients needed to contact only one single entity, receiving all project-related information available.

5.2.3.5 Ethical considerations

Medical research also raises a lot of ethical issues. Patient autonomy is of high relevance and can only be achieved through informed consent.

In ACGT ethical requirement with respect to the information of the patient, the informed consent, the feedback procedure to the patient in case of medical findings and patients' right to know and not to know were analysed carefully[78] and implemented in the procedures and policies.

To proof the results found and understand patients' needs in practice an international and national empirical survey on patients' and parents' of minor patients perspectives and needs regarding informed consent and data protection in clinical and clinico-genomic trials was carried out using questionnaires and interviews.[79]

5.2.4 Strengths and weaknesses of the ACGT legal Framework

The evaluation shows that the framework set up for ACGT provided a high level of data protection while breaking new ground. The establishment of the CDP as a central data controller for the project infrastructure as well as the "context of anonymity" set up by strong contractual agreements and the concept of de-facto-anonymisation can be seen as

[78] ACGT, D10.2, The ACGT ethical and legal requirements, pp. 11 ss., available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf.

[79] http://eu-acgt.org/uploads/media/D10_6_2_final_01.pdf.

main strengths of the framework and were considered to have the potential to be well serviceable in practice[80] or even to be trend setting.[81]

Experience showed however that with respect to the contract negotiations the practical implementation from a legal point of view was rather complicated and time consuming. This is something to be addressed in the course of p-medicine. The aim should be to adopt the data flow in order to be able to simplify the contracts to be concluded as much as possible. Anyhow, to some extent the complexity – and, with respect to contractual penalties, controversy - of such contracts will and ought to remain, if they shall serve as a backbone of a secure infrastructure that provides an environment where de-facto anonymous data only is processed.

The infrastructure was lacking of a patient identity management system. This was nothing that was identified to limit the usability of the framework. However, having in mind future international research projects data quality and data reliability will be increasingly important. A patient identification management system therefore is from a research point of view reasonable. From a data protection point of view, however, it is complex to use, as identification is the main goal and identification is exactly what should be avoided wherever possible from privacy perspective. Implementation should therefore only be pushed if the same data protection level can be achieved.

Besides this, in practice it appeared somewhat artificial to set up such a framework, including physicians, computer scientists, security specialists, ethicists, lawyers and members of many more disciplines, while at the same time there was the need of looking for a Trusted Third Party outside the project instead of nominating one of the (technical) partners of the project. Therefore it would be helpful to analyse if such strict interpretation (third party vs. project partner) is required in the context of interdisciplinary projects, and if not required under which conditions a partner of the consortium could serve as TTP.

Finally, already ACGT has shown that projects as such are not only subject to partners based in the European Union, as one of the ACGT- partners was from Japan. Still the set up framework did not provide a solution for a data transfer to this ‘third’ country.

Facing more and more worldwide research projects it would be reasonable to explore from a data protection point of view how such countries could be included in the data flow, without infringing European data protection law or jeopardising the ‘context of anonymity’.

[80] Review by Roger/Brownsword, http://www.uni-hamburg.de/fachbereiche-einrichtungen/fg_ta_med/rev_brownsword_lit_2011.pdf.

[81] Review by Wilms, http://ejil.oxfordjournals.org/content/22/2/614.short; Review by Seitz, http://beck-online.beck.de/default.aspx?typ=reference&y=300&z=NJW&b=2011&s=906&n=2.

5.3 Overview of the security framework within ACGT

5.3.1 Evaluation

5.3.1.1 Grid Security Infrastructure

The ACGT infrastructure was a Grid infrastructure based on two grid middleware solutions: the Globus Toolkit 4 (GT4)[82] and GRIDGE[83]. GRIDGE can be considered as a layer of basic services built on top of the low-level grid functionality provided by GT4.

As such the **Grid Security Infrastructure** (GSI), which is part of the Globus Toolkit, was the foundation of the ACGT security infrastructure. GSI is a widely-used specification for secret, tamper-proof, delegatable communication between components in Grid computing environments.

GSI supports both message-level and transport-level security. However, due to implementation and partly specification issues, message-level security has a relatively poor performance.[84]. Therefore transport-level security was used by default.

GSI uses X.509 end entity certificates (EECs) to identify persistent entities such as users and services. Authentication through plain username and passwords is also supported. When using username and password though, opposed to X.509 credentials, GSI only provides authentication and not advanced security features such as delegation, confidentiality, integrity, and replay prevention.

GSI also supports delegation and single sign-on through the use of standard X.509 Proxy Certificates. Proxy certificates allow bearers of X.509 EECs to delegate their privileges temporarily to another entity. For the purposes of authentication and authorisation, GSI treats EECs and Proxy Certificates equivalently.

Authorisation in GSI can be done using the Security Assertion Markup Language[85] (SAML)[86] standard from OASIS[87]. SAML defines security assertions and a protocol for retrieving them. Through SAML a third party authorisation decision service, such as the Gridge Authorisation Service (GAS, for details see chapter 5.3.1.2 below), can be used for access control requests to GT4-based services.

Although standard solutions to complex security problems were provided through the use of existing middleware (GT4, GRIDGE), it is to be doubted in hindsight if the burden introduced by it outweighs this advantage. For one, a considerable part of functionality implemented by the middleware was not strictly needed within ACGT (i.e. introducing maintenance of unnecessary components, complexity and performance degradation).

5.3.1.2 Authorisation

The Gridge Authorisation Service (GAS) is an authorisation system that acts as Policy Decision Point (PDP) for the components in a grid-based system. Its main goal is to fulfil the authorisation requirements for these components. It holds security policies (for all the connected grid components), which are used to make authorisation

[82] http://www.globus.org.

[83] http://www.gridge.org.

[84] C.f. Shirasuna et al., Performance Comparison of Security Mechanisms for Grid Services, available at: http://www.extreme.indiana.edu/xgws/papers/sec-perf-short.pdf.

[85] The Security Assertion Markup Language (SAML) defines an XML-based protocol, making it possible to exchange authorisation and authentication data between one or more security domains.

[86] p-medicine; 2011; D3.1 State-of-the-Art report on Standards, Chapter 8.1.1 SAML.

[87] The Organization for the Advancement of Structured Information Standards (OASIS) is a global consortium that drives the development, convergence, and adoption of e-business and web service standards: www.oasis-open.org.

decisions upon client requests. GAS has been designed in such a way that it can be integrated with external components easily and can manage security policies for complex systems. GAS has a flexible modular structure, making it possible to extend the GAS functionality with new communication modules.

GAS has a client-server architecture (Figure 2) in which the GAS server is responsible for all authorisation and management actions. Clients interact with the GAS server to query for authorisation responses and to configure and manage GAS.

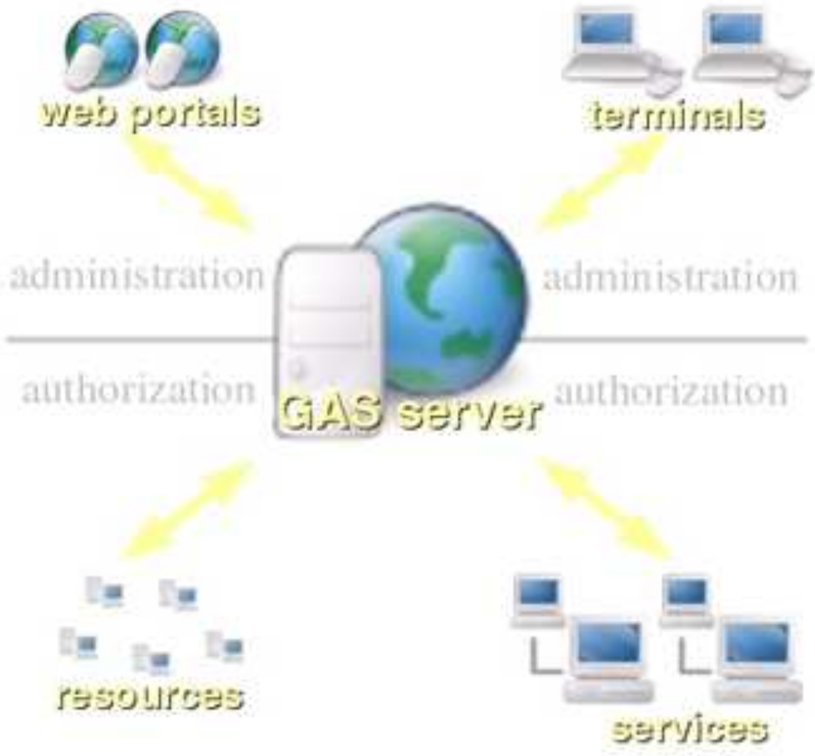

Figure 2: GAS client-server architecture

A GAS Server typically consists of the following components (Figure 3):

- a GAS Server which is the heart of the system. All security policies are stored here. GAS responds to a user request and returns an authorisation decision or security policy rules
- a GAS Client which can be used to communicate with the GAS server to get an authorisation decision or security policy, or to administrate the GAS server.
- a GAS Plugin is a component on the low (Globus) level, which can ask the GAS Server for an authorisation decision.
- a GAS Portlet is used to administrate the GAS Server. It is the easiest graphical way of managing the GAS Server.

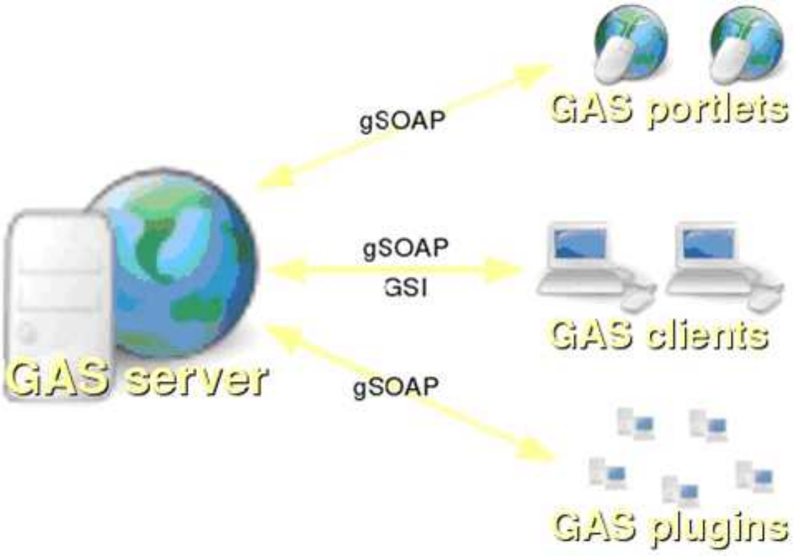

Figure 3: GAS Components

Before anybody can use any resource the security policy stored in the Gridge Authorization Service (GAS) must be checked (authorisation). This is what the **GAS plugins** are responsible for: they are calling GAS and ask if a given person is allowed to submit a task or transfer data to the resource. The GAS plugins support the “virtual account” mechanism, which is a convenient way of accessing resources in a Virtual Organisation. The idea is based on policies stored in a central authorization service. Based on that policy there is a pool of accounts created and managed by a service in a dynamic way. When a person accesses a resource, and is allowed to do so, he/she is (temporarily) mapped to one of virtual accounts on a local machine.

5.3.1.3 Public key Infrastructure

A dedicated **Public Key Infrastructure** (PKI) was set up for ACGT as authentication of services and end-users in ACGT was based on Public Key certificates (X.509). The ACGT PKI was a commercial grade PKI implementation (with respect to the followed security practices), which followed the X.509 standards and was composed of several interdependent modules. The service was not specific for GRID infrastructures, but rather supportive to the Common Grid Infrastructure.

The Certificate Authority (CA) module was the central component that issued and signed certificates for end-users and services. It was not directly accessible by end-users, but was used by the end-user administration site (Registration Authority front-end discussed earlier) and other PKI services.

The OCSP (Online Certificate Status Protocol) service and CRL (Certificate Revocation List) distribution point are services that allow checking of the revocation status of certificates. This information was essential in restricting access on ACGT resources to authorized users (and services) only.

A separate management site for supporting enrolment and de-activation of users and services through the ACGT PKI was provided within the project. Through this administration site, prospective users or new services were able to register themselves (and obtain their main ACGT credentials), and revoke and renew credentials. For the end-users, the registration process (i.e. credential generation) was made as transparent as possible through the site. A Java applet handled the key generation (of the private key) and the correct installation of the certificate generated by the ACGT Certification Authority on the end-user’s computer. Equally credential management for login to the portal (i.e. MyProxy bootstrapping with a delegated credential) was made transparent for the end-user through another Java applet that was included in the portal site. The tool greatly improved the user experience with small features, e.g. automatically looking for installed credentials on a user hard disk and on plugged in USB sticks.

Although ACGT tried to make the end-user experience as smooth as possible, the use of X.509 certificates for end-user authentication and proxy certificates for delegation was still bothersome to the end-user. Trusted certificates and the user's private key needed to be installed on the end-user's local machine. Due to a wide variety of browsers and operating systems this was not an easy task to accomplish. End users were required to use a USB-stick, on which they had to install their private key, to be able to authenticate from different client machines with the same credentials.

5.3.1.4 Delegation

Credential Delegation to software agents is required for complex and long running tasks. These agents can then act on behalf of the end user while he/she is offline. Delegation means that a user “transfers” his/her access rights (typically for a restricted period of time, scope or functionality) to another actor (service).

Delegation in ACGT was mainly provided through X.509 proxy credentials. Proxy credentials are basically certificates signed by the end-user’s certificate instead of a dedicated Certificate Authority. By issuing such a certificate the end-user delegates his/her rights to a specific service or person.

MyProxy is an online credential repository supporting this form of delegation. A dedicated ACGT MyProxy service was deployed and configured to allow only certificates generated by ACGT approved CAs. Although delegation in ACGT was not restricted to MyProxy based delegation, the delegation was always bootstrapped by the MyProxy service at the portal level.

When doing delegation through X.509 proxy certificates the end user signs the delegation credential, which is limited in time, scope and functionality, with his/her private key. In this way delegation is cryptographically enforced. As X.509 credentials are used, the same drawbacks apply as explained earlier for the PKI.[88]

5.3.1.5 De-identification

The de-identifying in the Data Protection Framework was handled by the **Custodix Anonymisation Tool** (CAT). CAT aimed to simplify the process of de-identifying and exporting of personal data. The tool was innovative in a sense that it offered a generic solution regardless of the type of data to be treated or of de-identification requirements.

CAT consisted of a “workbench” and a “wizard”. The “workbench” served at defining the mechanics (data protection profile) through which data was exported for sharing, the “wizard” allowed to easily apply those profiles on various data sources. The workbench allowed domain experts and privacy professionals to:

- create a mapping from a specific data format such as flat files (e.g. CSV), imaging data (e.g. DICOM), microarray data, structured data (e.g. XML, databases) to a generic data model;
- define the set of actions that should be performed on the generic data model in order to de-identify data (i.e. the data protection profile).

Once that a data mapping and a data protection profile were created in the “workbench”, end users (i.e. physicians) were able to easily export several data sources at once by using the wizard. A command line version was made available so that it was possible to automate this operation.

[88] See chapter 5.3.1.3 above.

An advantage of this approach is that a uniform and auditable data protection strategy can be followed in the different sites partnering in ACGT. CAT and its command line version were not really services though but rather standalone tools which were hard to integrate within the data workflow. ACGT required manual management of the CAT secrets, which does not scale in practice.

5.3.2 Strengths and weaknesses of the ACGT security Framework

From the evaluation we can conclude that the middleware used in ACGT (Globus Toolkit 4, GRIDGE) had a broad scope. It provided standard solutions for a lot of security problems. It is to be doubted in hindsight though if this advantage outweighs the burden introduced by it. Not strictly needed functionality introduced maintenance of unnecessary components, complexity and performance degradation. This burden can be avoided by going for a more lightweight architecture, which can evolve dynamically over time according to newly arising requirements.

ACGT was tied heavily to the use of X.509 certificates. As explained this was not perceived as being user-friendly by end-users. Trusted intermediate certificates might not have been installed correctly or private keys might get lost. End user usability can be greatly improved by using a more well-known (to the end users) authentication mechanism, e.g. password based authentication eventually extended with a hardware token for added security.

The pseudonymisation tool CAT allowed the different sites partnering with ACGT to follow a uniform and auditable data protection strategy. CAT though was not really an integrated service but rather a standalone tool (client and command line tool). ACGT also required manual management of the CAT linking tables and encryption and pseudonymisation secrets, which does not scale in practice.

5.4 Lessons learned from ACGT and challenges for the p-medicine data protection and data security framework

5.4.1 Legal challenges

Summarising the evaluation of the data protection and ethical framework set up for ACGT, the following lessons have been learned and are challenges for the data protection and ethical framework to be designed for P-MEDICINE.

Despite the fact that the legal framework was successfully flanked with a variety of contractual agreements, the negotiations of such agreements appeared to be time consuming and difficult. It therefore will be a challenge for the p-medicine framework to simplify such agreements as much as possible without weakening their impact and to start with contract negotiations at a relatively early stage of the project.

In addition to this effort will be invested to analyse the position of the Trusted Third Party more intensively in order to clarify whether a project participant, not being interested in patient data with respect to his/her business model could serve as Trusted Third Party. This is likely to ease the setting up of the project infrastructure. At the same time the implementation of an identity management system for patient data could progress the whole framework especially if it comes to research with prospective patient data. Still, from a data protection point of view this is highly challenging. The impact that such a system will have on the "context of anonymity" that p-medicine (as ACGT before) aims at, has to be analysed carefully.

Last but not least it will be useful it to provide a solution for non-European partners to take part in this European data protection framework in order to make worldwide medical research and exchange of patient data possible.

5.4.2 Security challenges

The evaluation of the ACGT security framework concluded that the usability could be further improved. Therefore p-medicine will mainly focus on this requirement. One of the challenges is to build a highly secure and user-friendly architecture, which are usually perceived as conflicting requirements. Highly secure systems also tend to be difficult to integrate and to manage. Therefore p-medicine will not only aim to be easy and straightforward from the end-user perspective, but also for administrators and developers (i.e. the p-medicine tool developers who need to integrate the provided security components).

ACGT not always convincingly focussed on security policies and procedures, i.e. password policies, re-identification procedures, etc. Both are an important part though of a security framework, therefore p-medicine will aim for elaborated security policies and procedures.

As the security framework defers from ACGT, i.e. the architecture doesn't rely on GRID middleware, typical security challenges as delegation, federation, policy management, governance, etc. need to be addressed.

The implementation of an identity management also imposes some technical challenges.

- As sending personal identifying information (attributes) in a nominative form to the Patient Identity Management Service (PIMS) might be in violation with the Data Protection Framework, identifying attributes could be encrypted or pseudonymised at the client before sending them to PIMS. However, due to the nature of cryptographic algorithms, similarity between records is not sustained. Therefore very similar attributes (e.g. typos) will be transformed to different encrypted or pseudonymised values. As a consequence fault-tolerant matching based on individual attributes is not possible anymore. This problem can be tackled by

  - using algorithms to match encrypted words.
  - matching records at client side through so called distributed probabilistic matching

- Challenging within the subject of record matching is the goal to reduce the number of false negatives and positives. This can be reduced by introducing manual intervention if two records are not similar enough, i.e. if their matching weight fails between a specific threshold range in which the engine cannot determine whether both records match or not.

6 Comparable data protection frameworks in medical research

6.1 Introduction

ACGT, which started in 2004, was one of the first large scale infrastructure projects for the medical research on the basis of genetic data on a European level. Since then several other projects have been dealing with the use of (sensitive) patient data for scientific medical research. This chapter shall analyse the data protection approaches of these projects and guidelines in order to examine possible improvements or simplifications for the legal framework for p-medicine. As a matter of course it would be way beyond the scope of this Deliverable to analyse all of these projects. Hence we rather try to present a selection of the most important projects in this field, which are of CRIP, BBMRI TRANSFoRm, VPH Share and the generic data protection concept developed by the German TMF. The first two projects are examples for data protection concepts in the field of human biobanking. As biobanks always contain (personal) patient data too, these approaches can also contain useful elements for the p-medicine data protection framework. TRANSFoRm and VPH Share are two of the current infrastructure projects on a European level. The German TMF is not a concrete project but provides a very advanced set of principles for the use of medical data in scientific research that has been elaborated in cooperation with members of ethical committees as well as German data protection authorities. These rules thus reflect a large consensus among the stakeholders in Germany.

In the following the data protection approaches shall briefly be outlined. For this purpose we will start with a short description of these projects and the challenges from a data protection law perspective, before analysing how these challenges are addressed in each project.

6.2 Central Research Infrastructure for molecular Pathology (CRIP)

6.2.1 Description and scope

The CRIP framework constitutes a central infrastructure for biomedical research involving human tissue repositories.[89] CRIP does not aim to build up a new biobank but it tries to foster the access to existing biobanks. Therefore, the main challenge within CRIP is the development of a meta biobank in order to facilitate the search for human tissue samples and related data suitable for the research project envisaged by the users.

The main sources for the CRIP meta biobank are the biobanks located at the so-called **database partners**.[90] These are clinical biobanks willing to provide access to their human biological samples to researchers. The CRIP database partners regularly transfer information as to the human biological samples available in their repositories to the central CRIP database. The data transferred is focused on specific diseases and primarily

[89] For further information see: http://www.crip.fraunhofer.de/en.

[90] Currently CRIP has four database partners: Charité Universitätsmedizin Berlin, Medizinische Universität Graz, Technische Universität München, Universitätsklinikum Erlangen. For a list of the current partners see: http://www.crip.fraunhofer.de/en/partner/datenbank?noCache=201:1324052293. In addition the CRIP framework is open to so called Metabiobank CRIP partners, that do not upload data to CRIP but have web services in place to accept, perform on their database, and answer web requests, which had been entered via CRIP. The search results retrieved and combined from both the meta biobank partner's and CRIP's databases are then displayed over CRIP and the meta biobank partner's website.

contains information on what human samples are available at the database partners. The reported samples have to be classified according to the WHO specifications.[91]

The CRIP database is developed and maintained at Fraunhofer IBMT, that also provides the **CRIP project management** coordinating the cooperation among the database partners and (if requested) managing projects agreed through CRIP. Furthermore, the project management is responsible for informing the media and the public about CRIP.

The project management is assisted by the **Advisory Board for CRIP**, which establishes guidelines, provides advice in ethical, medical and legal data protection, medically-scientific and communication issues. The Advisory Board consists of six members[92] and acts independently for CRIP.[93]

The **users** of the CRIP meta biobank are researchers that access the meta biobank upon registration in order to search for suitable human samples and related data available at the database partners.

The following figure shows the organisational infrastructure of CRIP as well as the interaction of the relevant bodies:

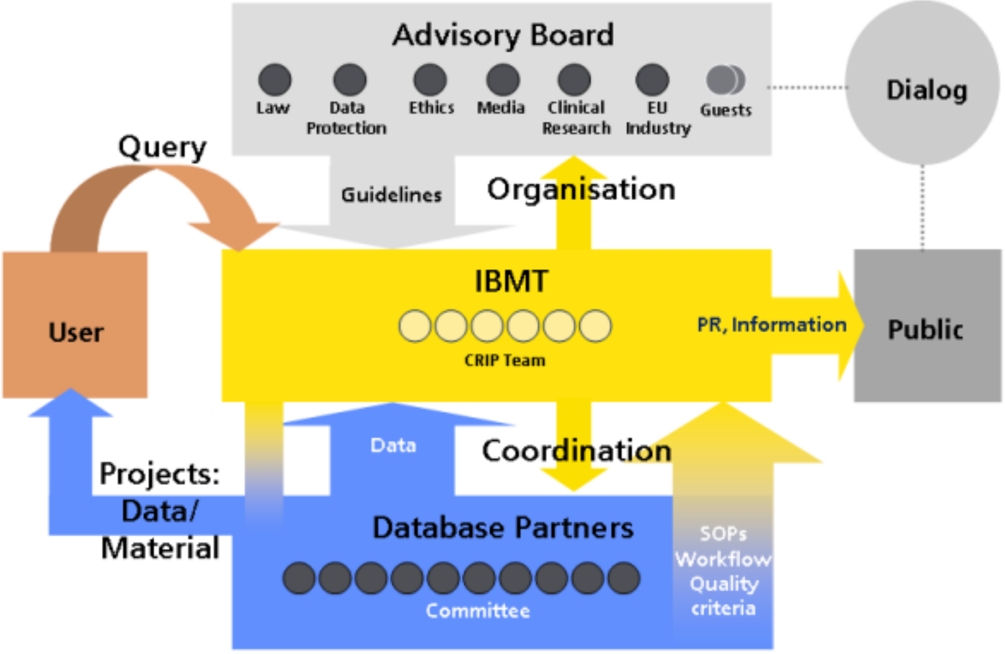

Figure 4: Organisation scheme of CRIP[94]

Within the CRIP infrastructure Fraunhofer IBMT tries to act as an honest broker, focusing on the organisation of the framework in order to facilitate and accelerate access to partner biobanks. Fraunhofer does neither provide biomaterial on their own and thus is no

[91] http://www.who.int/classifications/icd/en.

[92] For further information see: http://www.crip.fraunhofer.de/en/about/advisoryboard.

[93] The role and the rules for the functioning of the Advisory Board are laid down in a special document (Beiratsordnung): http://www.crip.fraunhofer.de/htdocs/pdf/Ordnung_IBMT.pdf (only available in German).

[94] Source: http://www.crip.fraunhofer.de/en/about/orga?noCache=425:1323869596.

competitor of the biobank partners nor does Fraunhofer have a research interest in biomaterial or respective personal data.

6.2.2 Data and data flows

The workflow of CRIP comprises the following six steps that are also shown in figure 5:[95]

In a first step the database partners extract research relevant data from their electronic medical records (patient database) to a so-called Inhouse Research Database (IRDB). The IRDB is a software tool provided by Fraunhofer IBMT that pseudonymises the patient data transferred to the meta biobank. Accordingly the IRDB only contains pseudonymised data.

Furthermore, in a second step, these data are anonymised before being transferred from the IRDB to the CRIP meta biobank. The data export must be authorized on each occasion and triggered by CRIP database partners. Thus, only anonymised data regarding the available biospecimens leave the biobank, whereas all biospecimens and related original data remain stored exclusively at the database partners.

In a third step registered users can send queries to the CRIP meta database and receive statistical information, a so-called "pool of cases", on the overall number of available cases meeting the criteria and requirements specified in their search query (forth step). CRIP exhibits merely aggregated data on how many specimens and data of a specific disease are available for research through CRIP.

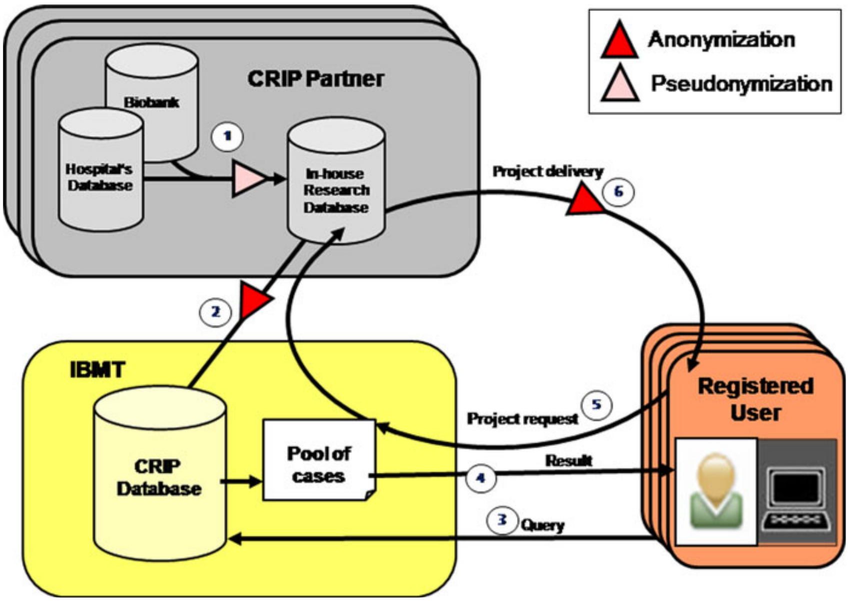

Figure 5: CRIP Structure and workflow[96]

[95] C.f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 235.

[96] Source: Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 235.

If the researcher finds sufficient potential biospecimen and data for the envisaged research project, he/she can send a project request (fifth step) to the respective database partners via the CRIP platform. The database partners have to answer these requests in reasonable time.[97] In case the database partners decide to provide samples and respective data for the envisaged project a bilateral contract will have to be concluded between the researcher and the database partners that will serve as a basis for their cooperation (sixth step).

6.2.3 Data Protection Concept

CRIP provides several measures in order to safeguard the privacy of the patients, the most important being the anonymisation of the data that is sent to the CRIP meta-biobank and in the following made available to researchers. Furthermore the data of the patient shall only be transferred to the CRIP database upon his/her prior informed consent.

The compliance to the rules established for the access and use of data as well as the interaction between the participants in the CRIP platform are ensured by contractual agreements. Database partners require ethical approval before joining the CRIP meta biobank. Access to the CRIP meta biobank is only open to registered users.

6.2.3.1 Pseudonymisation / Anonymisation

CRIP provides for a combination of pseudonymisation and anonymisation in order to protect the identity and the privacy rights of the patients. Furthermore only statistical data are shown to the researchers using the platform. All together the CRIP privacy regime comprises three steps.[98]

1. Pseudonymisation:

Only pseudonymous data shall be extracted to the Inhouse Research Database (IRDB) at the hospital site. Pseudonymisation is effected by a CRIP software tool installed at the database partners. By this it shall be ensured that only the hospital staff (who is bound by both professional confidentiality and specific confidentiality obligations) can re-establish the link from the sample to the patient at a later stage.[99]

2. Anonymisation:

The data exported to CRIP shall additionally be anonymised by said tool before the data are transferred to the CRIP meta biobank database. Data are termed ‘‘anonymised’’ in this context, if the data have been pseudonymised and the user/researcher has no access to the code.[100] Hence the CRIP infrastructure follows a concept of de facto anonymisation, which is also proposed for p-medicine.[101] It is announced that more technical detailed information on CRIP data formats, database structure and software will be published.[102] As for now there is however no detailed description available. In addition, all data from a specific CRIP partner is completely

[97] At the moment it is foreseen that the biobank operators have to reply within 10 working days.

[98] C. f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 237.

[99] C. f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 237.

[100] C. f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 235.

[101] This approach is also pursued in p-medicine. See chapter 7.3.2.

[102] C. f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 235.

replaced by via periodical updates, so that the pseudonym under which the data of the patient is stored within the CRIP database also changes with every update.[103]

3. Organisation in statistical groups:

Researchers sending a request do not receive detailed information as to the samples, but are only shown the number of cases ("pool") fulfilling the researchers search query.

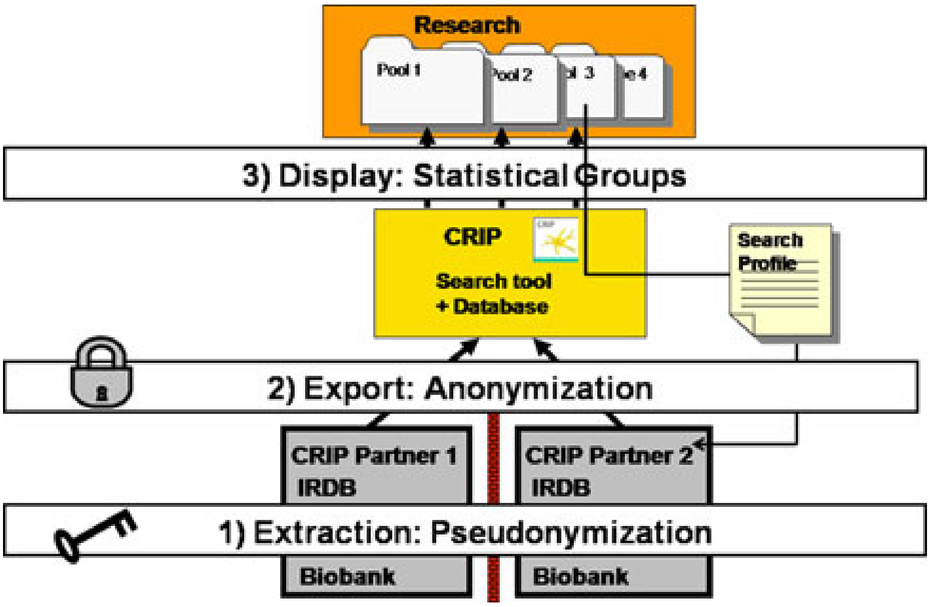

Figure 6: CRIP Privacy Regime[104]

6.2.3.2 Informed consent

Furthermore CRIP requires that the database partners have obtained and obtain patient's appropriate written informed consent. Text and procedure for obtaining this consent have to be approved by the local ethics committees and can be disclosed upon user's project request. However, also samples lacking patient's written informed consent are used in anonymised form. In this regard CRIP refers to an opinion on "Biobanks for research" of the German National Ethics Council according to which "the requirement of consent … the requirement of consent may be waived if the samples and data are completely anonymized."[105]

In case a sample donor withdraws his/her consent, the IRDB allows for deletion of the pertinent dataset before the whole data are anonymised and updated in the CRIP database. In other words, to make withdrawal of consent effective, the CRIP partner deletes the withdrawing donor's dataset in the IRDB and uploads all the respective

[103] C. f. Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 237.

[104] Source: Schröder et al., Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 237.

[105] German National Ethics Council, Biobanks for research – Opinion (2004) p.11.

(updated) IRDB-data to the CRIP database, thereby completely replacing all data previously exported to the CRIP database.[106]

6.2.3.3 Contractual agreements

The interaction between the CRIP partners, users and central CRIP facilities is established by and based on the “Database Contract”, which is concluded between Fraunhofer IBMT and all CRIP partners. The Database Contract applies to all partners in the same wording, and thus serves as the "basic law" for CRIP. It shall regulate and secure the data transfer, the development and regular up-dates of the CRIP database, the development and up-dates of the database partners’ IRDB (provided by Fraunhofer IBMT), the CRIP workflow and the standard operating procedures (SOPs) as well as cost recovery and fees.[107]

In the database agreement, CRIP and database partners commit to obey generally accepted standards and regulations concerning data integration, data transfer to third parties for research use, ethics reviews and patients’ written informed consent and data protection laws.

6.2.3.4 Ethical approval for the database partners

The CRIP partners’ biobanks must be approved by their local Ethical Review Boards (ERBs) and data protection commissioners prior to joining CRIP. Furthermore the databases have to be operated in accordance with all relevant legal regulations. The legal framework applicable to the databases can slightly differ, as the national/local data regimes have to be taken into account. The partners need provide evidence of these approvals before joining CRIP.[108]

6.2.3.5 CRIP Rules for Access

Researchers interested in using CRIP need to sign an application form disclosing affiliation, research group / head of group (if applicable), and contact details. Application details are verified by CRIP staff before access (user name and password) is granted. This is to restrict access to persons with plausible research interests.

6.2.4 Evaluation

The CRIP framework provides for several measures regarding the privacy of the donors, the most important being the pseudonymisation of the human tissue samples that are held in the biobanks of the database partners (Inhouse Research Database) and the anonymisation of the information of the samples available before transferring these data to the meta biobank operator. This shall provide that the meta-biobank only contains anonymous data. Furthermore, the meta biobank operator does not disclose the information of the respective samples to the researchers, but rather displays mere statistical data on how many samples fulfilling the requested criteria are available at the participating biobanks. These measures provide for a high level of data protection. The concept of providing pool data however cannot be applied to the data warehouse within p-medicine since the researchers are supposed to access the medical data of the patient (in anonymised form though).

[106] Schröder et al., Safeguarding donors’ personal rights and biobank autonomy in biobank networks: the CRIP privacy regime, Cell Tissue Bank DOI 10.1007/s10561-010-9190-8, p. 236.

[107] Schröder et al., Safeguarding donor´s personal rights and biobank autonomy in biobank networks: the CRIPS privacy regime, Cell Tissue Bank 2011, 12:233-240, 235.

[108] Schröder et al., Safeguarding donor´s personal rights and biobank autonomy in biobank networks: the CRIPS privacy regime, Cell Tissue Bank 2011, 12:233-240, 236.

Further the CRIP framework establishes relatively high standards for accessing the CRIP meta biobank in order to search for suitable samples and data. This aims to prevent that persons without a serious research interest send requests to the biobank partners. It seems very useful for the CRIP concept.

The CRIP data protection concept is highly suitable for the quick search of human biological samples and respective data for scientific research. The whole framework is governed by only one contract, which is valid for all participants, in order to increase the transparency for all participants. The conditions of the access to the samples and data, however, are left up to the bilateral agreement between the biobank operator and the researcher. Thus the different access regimes of the biobank operators shall be respected. This also allows for flexibility and increases the willingness of hospitals/biobank operators to participate in the framework.

CRIP however provides that biobank operators must at least have the permission of the ethics committee of their organisation to participate in the project. This shall ensure that only the existence of human samples is shown in the meta biobank that are really available for research. It also seems recommendable for p-medicine to ensure by contractual agreement that the transfer of data to the p-medicine data warehouse shall only be effected upon permission of the ethic committees of the hospitals.

Another measure that potentially can be useful for the creation of the p-medicine legal framework is the creation of an independent Advisory Board, that is not only responsible for advising the CRIP management and external communication, but also for the development of ethical and legal standards for the framework. The ethical and legal framework for p-medicine is designed in this deliverable. This framework shall serve as a basis for the p-medicine and shall in principle not be changed. Nevertheless possible changes of the ethical and in particular legal rules in the fled of data protection law have to be thoroughly supervised in order to react to possible changes within the duration of the project. Moreover it is necessary to supervise the compliance to these rules and standards by the partners. For p-medicine this task will be provided by the CDP.[109]

6.3 Biobanking and Biomolecular Resources Research Infrastructure of the European Union (BBMRI)

6.3.1 Description and scope

The Biobanking and Biomolecular Resources Research Infrastructure of the European Union (BBMRI) aims to create an interface between human specimens and related data and biological and medical research[110] and thus resembles the CRIP framework. Like CRIP, the BBMRI project shall provide an IT-infrastructure through which authorised researchers shall have the possibility to search and obtain required material and data from all participating biobanks for their research. BBMRI however shall not only include samples from patients but also samples from healthy persons, representing different European populations (with links to epidemiological and health care information), molecular genomics resources and bio computational tools to optimally exploit these resources for global biomedical research. It thus differs from CRIP and Biobank Suisse that only contain data and samples for persons carrying a (specific) disease.[111]

[109] Further information on the CDP is provided in chapter 7.3.5.

[110] For further information see http://www.bbmri.eu.

[111] Also see the other examples of already existing biobanks in Scandinavian countries (BBMRI NORDIC), http://www.bbmri.se/en/About-BBMRIse/BBMRI-Nordic.

6.3.1.1 Legal structure

The goal of setting up comprehensive collections of human biological samples of different (sub-)populations of Europe linked with continuously updated data on the health status, lifestyle and environmental exposure of the sample donors shall be achieved in a federated network of centres, so-called hubs, established in all (or at least most) European Member States. These hubs shall coordinate activities, including collection, exchange and analysis of samples and data for the major domains. The biobanks, biomolecular resources and technology centres, which are members of BBMRI, are associated with their specific domain hub.[112] Furthermore, a variety of public or private institutions (e.g., universities, hospitals, companies), which provide biological samples, data, technologies or services, may be associated with certain BBMRI members (associated partners).[113]

This structure shall provide great flexibility giving new participants the possibility to connect at any time and that it can be easily adapted to emerging needs in biomedical research.[114] The national hubs will link the national scientific community to the network as a whole. A headquarter shall be established in one Member State in order to coordinate the interaction of national hubs established in the Member States. This headquarter will provide a common access portal to resources available in Member states as well as appropriate facilities and expertise.

[112] BBMRI members will be the key providers of resources and technologies. Membership is non-exclusive so that members link BBMRI to other national, European and global initiatives (e.g., the emerging OECD global network of Biological Resource Centres or WHO programmes).

[113] Associated partners and subcontractors provide certain resources (services, data, samples, materials) to BBMRI. An associated partner, for instance, a hospital or research institute which provides biological samples and data, may be either reimbursed or compensated for its contribution by being granted free access to resources and technologies of the BBMRI. Associated partners may also be ministries, governments, research councils, and funding agencies from interested countries whether or not they currently support biobank or biomolecular resource infrastructure projects.

[114] BBMRI, D6.10: Overall report and Recommendations for BBMRI ELSI, including educational proposals, p. 3 s..

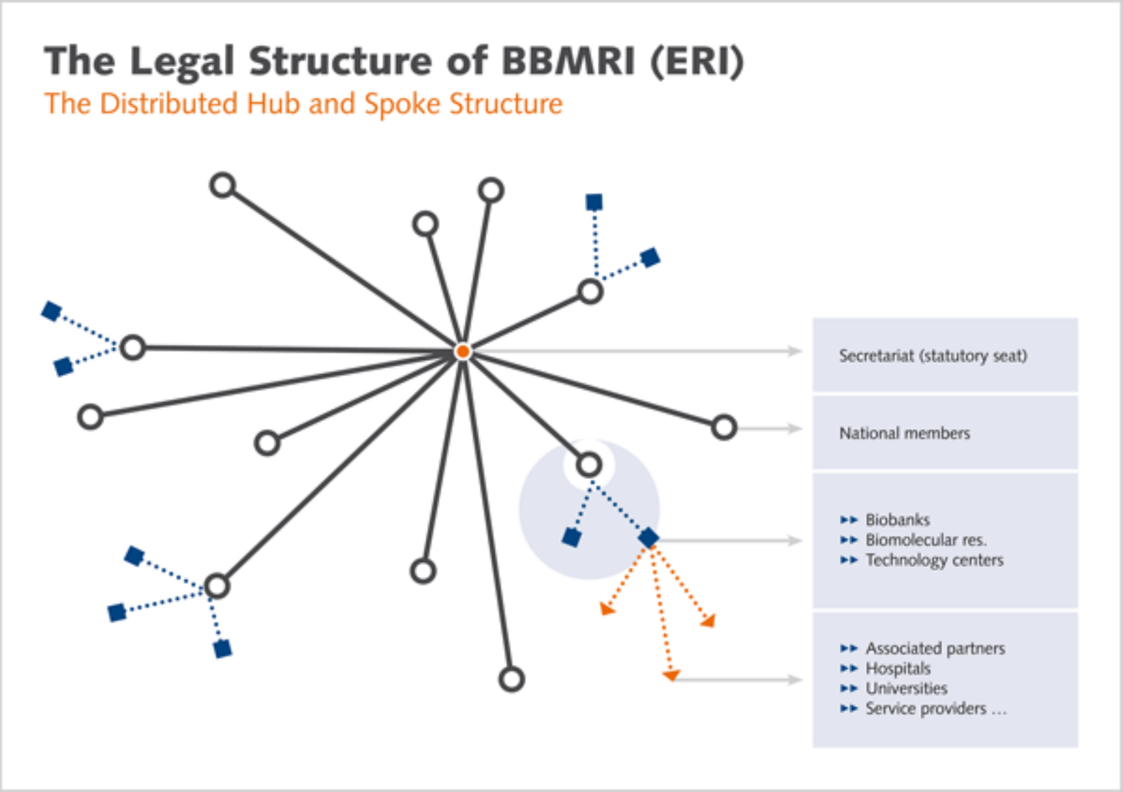

Figure 7: The Legal Structure of BBMRI (ERI)[115]

Legally the BBMRI headquarter as well as the national hubs shall be established under the European Research Infrastructure Consortium (ERIC) legal entity. This specific legal form (legal person sui generis)[116] is designed to facilitate the joint establishment and operation of research infrastructures of European interest.[117] It has been established by Council Regulation (EC) No 723/2009 of 25 June 2009,[118] laying down the procedure for the establishment and operation of pan-European research facilities. This regulation aims to provide an alternative to the existing legal forms for the establishment and operation of common (European) research facilities. Whereas these facilities mainly have been established either as an international treaty organisation or as an organisation under national law, the Regulation 723/2009/EC recognises that these two procedures are not equally suitable for all research facilities.

6.3.1.2 BBMRI-ERIC Partner Charta

The BBMRI-ERIC Partner Charta shall define the most important cornerstones for the participation of biobanks or biological resource centres (Partner) that are associated with BBMRI-ERIC setting out principles regarding the access policy, the data protection and management policy, informed consent, infrastructure management, quality management, reporting and charges.

[115] Source: http://www.bbmri.eu/index.php/about-bbmri/background.

[116] *An ERIC derives its legal personality directly from the Regulation. An ERIC is therefore neither an organisation pursuant to international law, nor a public or private legal person pursuant to national law.*

[117] http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric.

[118] Published in Official Journal L 206, 08/08/2009, p. 1.

As regards the access policy it is foreseen that samples and data need to be accessible through a clear access procedure. Access in the context of research projects performed within BBMRI-ERIC shall only be provided for specified research projects, in accordance with the terms of the consent given by the participant and after approval of the research project by a Research Ethics Committee (REC), leaving the decision, whether access will be granted for a specific project, to the partners. This decision, however, has to follow transparent decision making procedures. Furthermore it is held that information on individuals shall only be made accessible to authenticated scientific users in a coded or anonymised form in compliance with national and EU legislation, and subject to the BBMRI data access conditions. Partners will support integration of their data management system with that of BBMRI-ERIC by complying with the BBMRI-ERIC information requirements. The initial information requirements are realised as the expected minimal common data content and data structure in relevant databases. No access will be provided for non-research purposes (such as forensic, insurance or employment purposes), except pursuant to a court order.

Regarding the patients' rights, the Partner Charta contains a clear statement to honour commitments owed to donors, taking into account the principle of informed consent. It, however, does not develop its own standards but refers to the OECD Guidelines for Human Biobanks and Genetic Research Databases for issues related to informed consent, as appropriate and subject to the primacy of national and EU legislation.

The Partner Charta lays down a policy for quality management, which shall be compliant with OECD best practice guidelines for Global Biological Resource Centres Networks. The standard operating procedures (SOPs) shall be established and made publicly available for all processes related to sample collection, processing, storage, retrieval and despatch. It is recommended that SOPs should follow the procedures as specified in the WHO/IARC guidelines for biological resource centres for cancer research whenever feasible. Finally it is envisaged to establish a unique BBMRI biobank (collection) identifier.[119]

Within the Partner Charta it is stated that BBMRI-ERIC shall pursue its principal task on a non-economic basis. Limited economic activities, however, may be carried out provided that they are closely related to BBMRI´s principal task.

6.3.2 Data protection concept

Achieving a solution for the data protection issues related to the implementation of a cross-border exchange of biobanking material is seen as a prerequisite for the realisation of BBMRI.

The „most important cornerstones“[120] for the participation of biobanks or biological resource centres (Partner) that are associated with BBMRI-ERIC shall be laid down in a BBMRI-ERIC Partner Charter.[121] The Partner Charter is binding for any partner of the BBMRI-ERIC and shall be agreed between national BBMRI-ERIC nodes and the partners. In addition it is proposed to regulate the data protection and data security issues at stake by further contractual agreements in a more detailed way.

[119] See Kauffman, F & Cambon-Thomsen, A. Tracing biological collections: Between books and clinical trials. JAMA 2008, 299: 2316-2318.

[120] The European Research Infrastructure for Bio-Banking and Biomolecular Resources Partner Charter (Draft version 4; 6.12.10), p. 1.

[121] The European Research Infrastructure for Bio-Banking and Biomolecular Resources Partner Charter (Draft version 4; 6.12.10), available at: http://www.bbmri.eu/bbmri/index.php?option=com_docman&task=doc_download&gid=329&Itemid=97.

The data protection framework was subject to a joint deliverable of WP5 and WP6 of BBMRI.[122] On this basis the BBMRI framework proposes the following standard data protection documents and measures:[123]

I. Use de-identified data; and

II. Use k-anonymity to ensure anonymity of those individuals whose data are part of a data set made accessible to another (cross-border) legal entity; and

III. Adopt and apply:

a. the Standard Form Contract to assist compliance with Article 17 of the Directive; and

b. the Basic Information Security Measures set forth in the Standard Form Contract; and

c. the BBMRI-EU Model Data Access Policy ;

d. the National Data Processing Notification Requirements; and

IV. Condition access to data on the Standard Contractual Clauses adopted by the European Commission.

6.3.2.1 Pseudonymisation/Anonymisation

This shows that the BBMRI data protection framework is mainly based on the use of anonymous data ("de-identified data") and the use of k-anonymity[124] on the one side and the contractual agreements and policies that have to be respected on the other hand in order to ensure compliance to the rules on a contractual basis. As the project is only in its preparatory phase there are no further details on how pseudonymisation/anonymisation of human tissue samples and data will be provided technically.

6.3.2.2 Contractual agreements

In addition to the rather programmatic rules of the Partner Charta, the conclusion of several contractual agreements is proposed, that shall regulate the compliance to the data protections laws in a more detailed way. These agreements shall briefly be outlined in the following.

6.3.2.2.1 Standard Form Contract to assist compliance with Article 17 of Directive 95/46/EC

In order to comply with the requirements regarding confidentiality and security of processing established in Art. 17 para. 3 Directive 95/46/EC, it is proposed that the contracts concluded by the BBMRI members shall be based on the "Standard Form Contract to assist compliance with obligations imposed by article 17 of the Data Protection Directive 95/46"[125] prepared by the European Committee for Standardisation (CEN).[126] The contract form has been defined according to the requirements of data controllers who employ or use third party processors to use such contracts. It may be

[122] Joint Deliverable of WP 5 and WP 6: To explore pan-European solutions for the cross-border data protection issues associated with BBMRI, available under: http://www.bbmri.eu/bbmri/index.php?option=com_docman&task=doc_download&gid=316&Itemid=97.

[123] BBMRI-EU Data Protection Group-Karolinska Institutet & Legal Pathways-Report DATED 08-03-2011, p. 6 s.

[124] The concept of k-anonymity demands that every tuple in the microdata table released be indistinguishably related to no fewer than k respondents.

[125] ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/DPP/CWA15292-00-2005-May.pdf.

[126] http://www.cen.eu/cen/pages/default.aspx.

used as a complete agreement to accompany a separate data processing services agreement or the operative clauses can be extracted and incorporated into the processing services agreement.

6.3.2.2.2 Basic Information Security Measures set forth in the Standard Form Contract

In addition the Basic Information Security Measures for the processing of personal data set out in the Appendix to the CEN Standard Form Contract[127] shall be taken into account. These measures can also be taken into consideration by BBMRI-EU members processing personal data strictly on their own behalf, without using (external) processors, in order to comply with the requirements established in Art. 17 para. 1 Directive 95/46/EC.

6.3.2.2.3 BBMRI-EU Model Data Access Policy

According to BBMRI-ERIC Partner Charta samples and data need to be accessible through a clear access procedure compliant with the general access procedures and conditions of BBMRI-ERIC. Within the WP5/WP6 working document a Model Data Access Policy has been set up, comprising a set of formal rules, criteria and priorities, which shall facilitate a competent and fair review of all incoming proposals for the use of the data available at or through BBMRI-EU members. The document underlines the importance of clear and objective access criteria that are clearly communicated to the potential users.

Accordingly the policy should make clear what data exactly could be accessed or transferred and also what data will NOT be accessible or transferable. The policy moreover has to provide for a clear and transparent procedure for the handling of requests for access to or transfer of data to third party researchers.

The procedure for assessing requests (data access procedure) could provide for the assessment of any requests by the pertinent scientific and/or governance committee or a separate access committee.

All requests for utilisation of the available data, including requested analyses, access or transfer should be submitted in a standardized application form. By the means of this form the requester should disclose, at least, the following information:

I. *name and position of Applicant, including employment or affiliation with any organisation, either for-profit or non-profit;*

II. *credentials of the Applicant;*

III. *the nature and amount of data requested;*

IV. *proposed or requested uses of the data, including ultimate objectives of the use.*

Whenever the proposed use concerns scientific investigation the applicant should be required to describe in addition:

I. *the scientific aims of the investigation,*

II. *the outline of the study design,*

III. *give an indication of the methodologies and*

IV. *specify preceding peer-reviews (if any present);*

V. *presence of all required approvals by local ethics committee or institute research board.*

[127] ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/DPP/CWA15292-00-2005-May.pdf.

The information disclosed in the process of the request should be treated as confidential and should only be disclosed to the persons evaluating the request. Unless explicitly consented for by the requester, this information should not be used for purposes other than evaluation.

6.3.2.2.4 National Data Processing Notification Requirements

Finally it has to be analysed whether the applicable national data protection law provides for a duty to notify the supervisory authority (Data Protection Agency) before carrying out wholly or partly automatic processing operations in the sense of Art.18 para. 1 Directive 95/46/EC. Furthermore, Directive 95/46/EC states in its Art. 20, that Member States shall determine the processing operations likely to present specific risks being subject to “prior check” by supervisory authority, which also have to be taken into consideration.

6.3.3 Summary

In BBMRI, like in CRIP, the data protection concept shall be based on pseudonymisation and anonymisation of the data. Regarding anonymity BBMRI tries to focus on k-anonymity, which shall guarantee that every set in the database released has to be indistinguishably related to no fewer than k respondents (k-anonymity).

Apart from these measures great emphasis is laid on a clear and transparent policy for the access to BBMRI. As in CRIP the biobank operators shall not be forced to provide access to their human biological samples and data to every participant, so that it is up to the biobank partners to decide who will be granted access in each case. This decision however has to be transparent. These general principles are already stated in the BBMRI Partner Charta. More detailed rules for the access are provided for by a model data access policy, that shall clearly specify what data are accessible and which data are not accessible. Further a clear description of the access procedure shall be provided. For the access to data and samples for the purposes of scientific research the researcher shall provide detailed information about the envisaged project. As clear access rules provide transparency and thus are suitable to generate trust, the provision of a clear access policy to the data warehouse shall also be considered for p-medicine although this might increase the effort for the registration.

Another important point within BBMRI is the compliance to the data security requirements imposed by Art. 17 Directive 95/46/EC. In this context the framework of BBMRI refers to the standard contractual clauses designed by the European Committee for Standardisation. Within p-medicine we propose to implement specific rules regulating the data security issues in the contracts with the data exporters[128] as well as the end users[129] in order to avoid the conclusion of an additional contract. This allows us to focus on the specific data security issues for the use of patient data within the project. Only these issues will be regulated by contractual agreement, what will shorten the contracts and increase their readability.

[128] See Annex A of this deliverable.

[129] See Annex B of this deliverable.

6.4 Translational Research and Patient Safety in Europe (TRANSFoRm)

6.4.1 Scope

The overall aim of the TRANSFoRm (Translational Research and Patient Safety in Europe) project is to develop a "rapid learning healthcare system" driven by advanced computational infrastructure that can improve patient safety as well as the conduct and volume of clinical research in Europe.[130] The project shall provide interoperability between different clinical systems, across national boundaries, and integration of clinical systems and research systems. TRANSFoRm will support clinical studies with potential patient safety value and directly support the use of evidence for diagnosis, reducing diagnostic error.

There are very few documents and deliverables about the TRANSFoRm project available to the public as for now. A first analysis of the data protection issues within the project is provided in deliverable 3.2 "Report on regulatory requirements, confidentiality and data privacy issues".[131] The objective of this deliverable is to develop an extensible privacy and confidentiality framework that supports the different stages of the clinical trials process for finding and recruiting eligible patients while maintaining their privacy. Therefore, the framework must distinguish between different types of data including anonymised and identifiable clinical data providing approved mechanisms of data access at different levels and at different stages. The overall framework needs to support and preserve different local data sharing policies in different health organisation while enabling them to benefit from TRANSFoRm services.

6.4.2 Data flows and data protection approach

TRANSFoRm deals with heterogeneous data from different data sources (e.g. primary, secondary), related to different context (medical care, research) and associated with different degrees of risk of identification (personal identifiable, pseudonymous, anonymous). The project chose a formal approach. The data flow within the project is described on the basis of data flow schemes within the TRANSFoRm use cases. These schemes basically distinguish two zones in which data can be processed.

Zones are defined as areas for data sources that are comparable and similar with respect to purpose, rules and regulations for use. The idea is to have personal identifiable data, pseudonymised data and anonymous data in different zones to structure the data sources zones according to risk for confidentiality and data privacy issues, representing areas with **low, medium and high risk of identification**.[132] In the framework two major zones are defined:

- data source zone
- research zone

The **data source zone** contains data sources available and needed for research. This zone can again be split into a care zone and a non-care zone.[133] Whereas the **care zone**

[130] http://www.transformproject.eu.

[131] https://transform.kcl.ac.uk/sandbox/groups/publicdeliverables/wiki/welcome/attachments/7183f/Transform_Confidentiality_Framework_Version1_fin_31032011.pdf?sessionID=4018b7d477f788ae1e69b3c94f7aba0bb27c729f.

[132] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), p. 20 s..

[133] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), pp. 20 ss..

is dedicated to data for patient diagnosis and treatment (having been collected during the context of medical care), the **non-care zone** contains research databases, registers, etc. The data stored in the care zone are generally personal medical data that are used within the care context by the treating physician. It needs explicit consent by the patient and/or authorisation for use by researchers outside the care context.

Data sources in this zone are normally pseudonymous. They are either based on explicit consent or the research use will be based on country-specific regulations (e.g. exemptions to consent for research) usually allowing for an opt-out mechanism. The use of these databases shall be authorized and controlled by the data controllers of these sources based on a defined policy. Data sources in the non-care zone may contain primary data (e.g. clinical trial database, cohort study database) or secondary data (e.g. GPRD, cancer register).

As databases in different countries operate under different rules and regulations concerning confidentiality and data privacy these zones can be divided in **subzones** that shall reflect the different level of data protection within the different countries.[134] These subzones shall contain data that are comparable and can be used for the same or a very similar purpose and with similar applicable rules and regulations for their use.

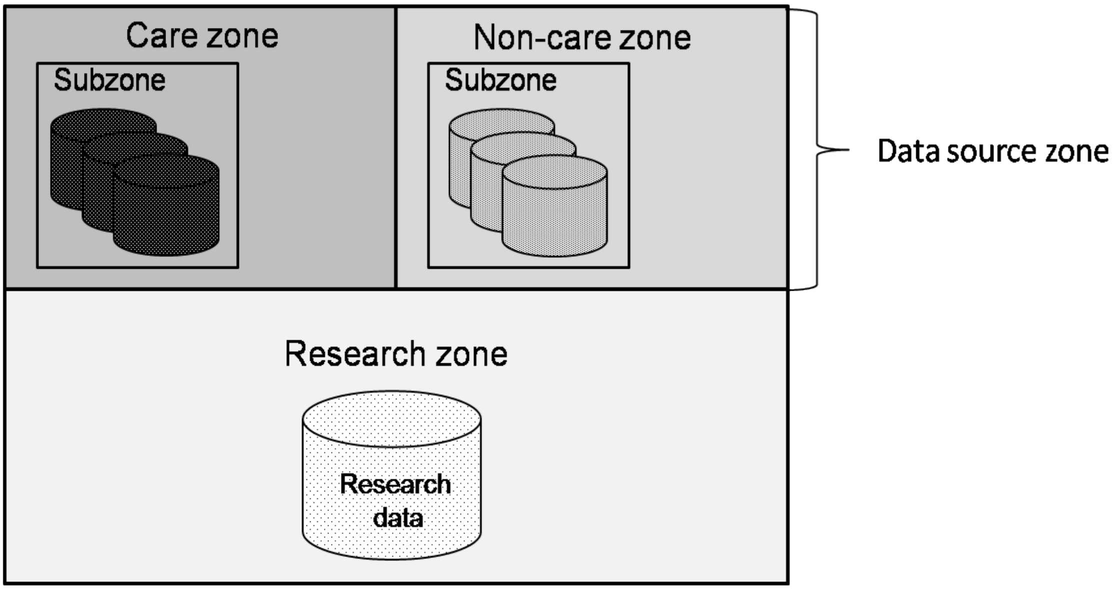

Figure 8: Separation in different zones

6.4.3 Data protection concept

The TRANSFoRm data protection concept intends to find a balance between individual privacy interests on the one hand and research with health care data for the public good on the other hand. The framework tries to distinguish between different types of data including anonymised and identifiable clinical data providing approved mechanisms of data access at different levels and at different stages. The overall framework needs to support and preserve different local data sharing policies in different health organisation while enabling them to benefit from TRANSFoRm services.

[134] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), pp. 22 ss..

A description of the data protection framework is provided by Chapter 4 of Deliverable 3.2. Accordingly the primary responsibility shall lie at the treating physician, who is the first in collecting the data. As the treating physician generally shall try to raise these care standards as well, he/she should try to use patient data acquired in the medical care context for medical research, or at least allow such use, according to the applicable national regulatory conditions. In most systems these conditions allow for such research *by* the treating physician.[135] The treating physician must only make data available with the consent of the patient concerned or if a research exemption applies under national laws. As data controller the treating physician retains the ultimate responsibility for allowing use of these data by third parties in whatever form.[136]

This transfer of data to a third person leads to a so-called "data chain" or "chain of data". The framework defines several conditions for the transfer of data, the most important being the use of pseudonymisation and anonymisation technologies. The data chain should be generally transparent to the patient when consent is required, and just as long as necessary in order to achieve the envisaged research purposes. Within the chain it should always be clear who is the controller of personal data (when those are used) and who is the processor of the data. Each data controller of a database shall ensure compliance of the data processes with the applicable national/regional data protection legislation. In the data chain a third party receiving data should not be responsible for the regulatory compliance of the sending organisation. Access to existing databases and use of such data is the responsibility of the data controller of the existing database until the third party has taken over responsibilities as data controller by forming a new database.

The transfer within the data chain shall be regulated by contractual agreements, so-called Data Transfer Agreements (DTA). Within the TRANSFoRm project a model DTA will be developed for the TRANSFoRm partners, which can be modified according to the needs of the partners. Furthermore database policies shall be established that stipulate certain cornerstones, such as the acceptable sources of data, the pseudonymisation and anonymisation measures used for collecting those data in the database, the aim of the database, the conditions under which the database can be used for research.

The establishment of clear database policies shall facilitate the process of creating individual DTAs between the partners. If there is a database policy (or statutes) of the recipient and controller or holder of the database, which comply with the standards mentioned in this document, it should be sufficient that the provider of the data declares that the data are transferred in accordance with that policy.

The role of explicit consent within the data protection framework is rather marginal. In principle, explicit consent shall be obtained whenever personal data are used by a third party. However, the prevailing general rule is that, third parties shall only use "coded-anonymous" data for research. Thus explicit consent (together with the processing of data on the basis of national exemptions) only constitutes an exception to this principle.[137]

6.4.4 Evaluation

The TRANSFoRm project aims to provide data protection solutions for a variety of projects dealing with patient data for the purposes of medical treatment and for medical

[135] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), p. 27.

[136] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), p. 27.

[137] TRANSFORM, Report on regulatory requirements, confidentiality and data privacy issues Part B: Confidentiality and data privacy framework (Deliverable 3.2), pp. 31 s..

research. This project demonstrates the importance of a clear separation between these two areas of data processing (source zone and research zone). Furthermore it shows that data shall be kept in a structured way according to their sensitivity in order to provide adequate data protection measures for each group of data.

Within p-medicine the source and the research zone shall be divided from each other in a similar way (hospital domain and a research domain). This division shall clarify the responsibilities for the data processing and the applicable data protection policies. As the medical data of a patient transferred to p-medicine shall be stored together in a data warehouse we will refrain from a further creation of different groups of data. Hence all data transferred to the framework will underlie the same data protection policy.

6.5 VPH Share

6.5.1 Data flows, concept, data protection approach

The main idea of VPH Share is to develop, integrate and maintain an environment which will enable the VPH-Share workflows, as well as any application making use of VPH-Share resources, to operate on top of the cloud and high- performance computing infrastructure provided by the project. For this aim a consistent service-based system shall be developed. This system shall enable end-users to deploy the basic components of VPH-Share application workflows (known as Atomic Services) on the available computing resources, and then enact workflows using these services.[138]

Data sources are usually clinical data, such as medical images and/or biomedical signals from individual patients. The operations range from secure access and storage through annotation, data inference and assimilation, to complex image processing and physics-based mathematical modelling, to data reduction and representation. Thus, VPH Share shall provide the essential services, as well as the computational infrastructure, for the sharing of clinical and research data and tools, facilitating the construction and operation of new VPH workflows, and collaborations between the members of the VPH community. In order to achieve this, a number of specific objectives are formulated:

I. **Accessibility:**

   Technologies need to be accessible to the general community, providing a low bar to entry. Further simple user interfaces and tool sets shall allow the use of the tools without understanding the underlying technology,

II. **Utilisation:**

   The project shall facilitate the utilisation of the infrastructure by VPH-Share as well as the facilities produced by the VPH community,

III. **User Guidance:**

   Furthermore VPH Share shall provide guidance to the data providers on how to produce data sharing agreements. Furthermore a governance framework shall be designed to facilitate the bilateral contracts between providers and users within the network.

IV. **Large Users:**

[138] See Work Package: WP2 Data and Compute Cloud Platform Deliverable: D2.1 Analysis of the State of the Art, Version: 1.5, p.7.

For larger sites wishing to deploy their own cloud infrastructure for security reasons support will be provided. A seamlessly integration of these sites with the network shall be ensured.

6.5.2 Evaluation

As for now there are very few VPH Share deliverables publicly available. Hence this analysis has to be done on very basic information. The analysis shows that the work VPH Share is focused on the development of tools and technologies that provide for the cooperation for the use of VPH technologies in a very flexible way. Unlike p-medicine VPH Share therefore does not aim to establish a common data warehouse for all data that shall be available for the project partners, but rather tries to provide the technologies and know how necessary in order to establish a cooperation including the transfer/exchange and processing of data, including patient data, on a bilateral level. The legal aspects related to the realisation of this cooperation shall be taken into account by providing guidance to the data providers on how to produce data sharing agreements as well as by designing a governance framework to facilitate these bilateral contracts. The content of these documents is however not publicly available. Considering that VPH Share shall also be based on the deployment of cloud computing technologies, special attention has to be directed on legal data protection and data security aspects of the storage and processing of mostly sensitive patient data in cloud computing infrastructures.[139] This will require a sound assessment of the question whether to transfer of patient data to public clouds, such as Amazon Cloud services, can at all be carried out in conformity with the current data protection legislation. For p-medicine the use of (public) cloud infrastructures is not envisaged. There is also no need to provide model contracts as all the partners and the end users will have to conclude the same contracts for the transfer of the data to the data warehouse and to access these data over the p-medicine infrastructure.

6.6 Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (TMF)

The Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (Technology and Media Platform for integrated medical research, TMF) is an umbrella organisation for medical research networks. The TMF tries to improve the organisation and infrastructure of medical research in networked structures and supports researchers in jointly identifying and solving problems of an organisational, legal and technical nature regarding the interconnection of research and healthcare, standards and terminology, legal and ethical frameworks, quality management, technology assessment, public relations.[140] Members include medical competence networks, networks dealing with rare diseases, psychotherapy networks, zoonosis networks, coordination centres for clinical trials, the German National Genome Research Network (NGFN), Fraunhofer institutes as well as a patient organisation in the form of the German Cystic Fibrosis Institute.[141]

[139] A detailed analysis of legal data protection and data security aspects is provided by the respective deliverables in OPTIMIS. C.f. Forgó et al., OPTIMIS, Cloud Legal Guidelines (Part I), http://www.optimis-project.eu/sites/default/files/content-files/document/optimis-cloud-legal-guidelines-part-i.pdf, and Forgó et al., Cloud Legal Guidelines: Data Security, Ownership Rights and Domestic Green Legislation (Part II), http://www.optimis-project.eu/sites/default/files/content-files/document/optimis-cloud-legal-guidelines-part-ii.pdf.

[140] http://www.tmf-ev.de/EnglishSite/AboutUs.aspx.

[141] A list of current TMF member networks is available under: http://www.tmf-ev.de/EnglishSite/MemberNetworks.aspx.

In order to facilitate medical research and the exchange of medical data the TMF developed generic data protection concepts for networked medical research projects and took over their coordination with of the competent employees of the Data Protection Officers at both federal and state level. These generic data protection concepts have served as a basis for several competence networks and numerous other cooperative research projects to develop and accelerate the progress of their specific data protection concepts. Furthermore, these data protection concepts and any necessary changes and additions were submitted to the respective state Data Protection Officers with reference to the TMF template. Thus these model data protection concepts consequently underwent a simplified assessment process.

The Working Group on Data Protection of the TMF distinguishes two ways of medical research networks: clinically focused research networks on the one hand and scientifically focused research networks on the other.

Clinically focused research networks are focused on the direct derivation of scientific data from the treatment process.[142] Scientifically focused research networks are mainly based on the available research-relevant data within the network. These networks shall provide temporally and spatially largely unrestricted access to research data including the ability to online research and analysis. The collected data are not only obtained in the process of treatment but may also be generated in specific surveys. As these data do not underlie the clinically motivated quality control, they have to pass a quality control system be prior transfer to the research database, which minimizes defects in feedback to the clinic in the reasonableness and completeness of the data.[143]

6.6.1 Commonalities of the generic concepts

Although these models differ, they show a number of similarities from a data protection and technical perspective, which shall be summarized in the following.

6.6.1.1 Contracts and legal personality

In both scenarios it is proposed to conclude contractual agreements concretising the data protection relevant legislation and professional ethics codes applicable to the research project, so that all parties involved can easily look up to what rules the medically treating and researching staff is legally bound to.[144]

Furthermore it is held essential for a legally secure implementation of the rules for data protection and data security that the research network has the status of a legal person, such as an association. This legal entity may conclude contracts for the providers of central services and thus bind them to the compliance to the organisational and data protection relevant rules.[145]

The legal entity shall conclude contracts with the physicians and their staff laying down the requirements for research data and their transfer to the legal entity. Furthermore the legal entity shall conclude contracts with the scientists regulating the procedures to access the data as well as the conditions of proper use of data and biological samples. Finally contracts shall be concluded with the providers of central services regulating the duties and obligations associated with data processing. The contracts shall also ensure the independence of database administrators from research-based personnel.[146]

[142] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, pp. 4 s..

[143] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 5.

[144] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 7.

[145] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 7.

[146] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 8.

6.6.1.2 Privacy Committee

In every research network there shall be a committee being responsible for managing the exchange of and the access to patient data. This committee shall inter alia be responsible for the review and approval of requests from researchers to access the data (also defining the extent and the conditions of access), the review and approval of requests to inform the patient of research results by their physicians, and the assignment of the provides of central services and the adoption of respective regulations for use for these services, including data protection and data security-related rules.[147]

6.6.1.3 Rules for the use of data

It shall be guaranteed by contractual agreement that the researchers may use the data made available exclusively within the research framework and according the conditions established by the approval. The transfer of received data to third parties shall be prohibited. The use of data outside the research framework shall be subject to specific agreements with the data protection committee of the network.[148]

6.6.1.4 Safety policy - terms of use

Finally the terms and conditions for the use of the central services shall be laid down in order to bind the users and operators to the necessary measures and procedures.[149]

6.6.1.5 Right of the patient, withdrawal of patient cooperation, time limits for storage of data

The patients shall have the right of information regarding the data that are stored on them by addressing the treating physician who then will initiate the process that leads to the requested information.

Whenever a patient who´s data are stored within the research network withdraws his/her consent or dies the data of the patient has to be erased or made anonymous depending on the consent.

Furthermore a retention period shall be determined taking depending on the research objectives. As a general rule, it is proposed to store the data at least for a term of six years from the last treatment, considering that research findings might be of direct use for the patient.[150] After this determined period of time, the patient data shall only be used in anonymised form. A longer retention period of non-anonymous data shall require an explicit justification and the specific consent of the patient.[151]

6.6.1.6 Use of biological samples

Within these generic models it is also possible the establish collections of biological samples from the treatment process for scientific purposes, provided that appropriate regulations are implemented. For the extraction and processing of biological samples and their analysis for scientific purposes the specific information and consent of the patient shall be required, in particular when samples shall be analysed genetically.[152] The delivery of biological samples to a researcher shall in addition require the approval of the research network´s Privacy Committee. In it´s decision the Privacy Committee

[147] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 8.

[148] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 8.

[149] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 9.

[150] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 10.

[151] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 10.

[152] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 10.

shall also consider the re-identification risk, which is generally associated with the use of biological samples.[153]

6.6.2 Specific conditions for research focused networks

Apart from these general rules the TMF generic data protection concept provides specific rules for the two types of research networks defined. As p-medicine is focused on the transfer of patient data to a central data warehouse it falls under the category of research focused projects. These specific conditions of the framework shall not regulate the gathering of data through the treating physician, but their transfer and further use for the purposes of scientific research.

6.6.2.1 Fundamental requirements

The TMF framework establishes three **fundamental requirements** in order to ensure the quality and confidentiality of the data:

1. Implementation of a Patient Identification system.

A system for the secure identification of the patient is required, that allows for the identification of the patient without using the patient´s name in the context of scientific research. The system shall generate unique strings for each patient as a patient identification. The creation of synonyms or homonyms shall be avoided.[154]

2. Implementation of a high security pseudonymisation tool.

An adequate high security pseudonymisation tool shall be used. The links to the patients shall be stored by a trusted third party that holds the keys all data within the research framework.[155]

3. Implementation of a data quality management.

A data quality management shall be foreseen in order to ensure the plausibility and completeness of the data.[156]

6.6.2.2 Tools

From a technical-organisational point of view three tools are required.

6.6.2.2.1 Patient list

The identification of each patient by attributing a single and unique ID is regarded as a measure to ensure data quality, as it avoids that patients that have been treated in different hospitals are regarded to be different persons.[157] Accordingly each patient receives a so-called patient identifier (PID). The PID is generated by a special tool, the PID generator, on the basis of the patients name and other data (IDAT) that is then stored together with the PID in a patients list.[158] On the basis of this list, the PID

[153] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 11.

[154] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 44.

[155] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, pp. 44 s..

[156] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 44.

[157] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, pp. 46 s..

[158] The elevation of the IDAT has to be uniform. It is recommended to refer to the data set on the insurance card, since hereby the greatest degree of standardization is achieved. In addition to the name also the birth name and former names shall be stored in the patient list. In addition, other data (notifying hospital etc) can be included in the list. Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 47.

generator verifies whether a specific patient has already been assigned a PID. If this is not the case, a new PID will be generated and included in the patient list.[159] As soon as the patient declares his/her consent to participate in a trial/study, the hospital can register the patient in the patient list by transferring his/her identification data.[160]

The function of the patient list is generally automated. Depending on the importance of the assignment of the data to the "right" patient, a manual intervention can be established, e.g. when registering a patient was possible, but not secured.[161]

6.6.2.2.2 Pseudonymisation

The pseudonymisation of medical data consists in the cryptographic transformation of the PID. Only the pseudonym shall be forwarded together with the medical data to the research database, whereas the PID as well as the pseudonym remain stored at a trusted third party.[162] In the generic data protection concept of TMF there shall be only one

The TMF describes five scenarios different models of pseudonymisation of increasing complexity.[163] The first three relate to the unique use of data for a specific predefined purpose of research. Starting with the use of single data source for one-time usage (e.g. the a simple statistical analysis of patient data)[164] and multiple data sources for one-time use (e.g. follow up data), where no identification is needed and databases for one-time use with the possibility for re-identification, up to the establishing of pseudonymised databases for research or a central clinical database for long term use.[165] The last two scenarios are in the TMF generic data protection concept will help to build research networks for long-term data. For the construction of such long-term research infrastructures the mere distinction between the treatment and research context does not suffice. In these cases more complex solutions are needed, in which up to four different areas have to be distinguished: the treatment context, the local collection of research data, the central data pool for a research and the use of data from these pools as a data base for evaluation specific obligations or recruitment of cases for new trials.

6.6.3 Evaluation

The TMF network designed an advanced set of principles for the use of medical data in scientific research. As a general principle for every type of medical research it is stated that protection relevant legislation and professional ethics codes applicable to the research project are concretised by contractual agreements. These agreements shall specify the rules for the access to the data in the project as well as the data security policy applying to the project. Hence the conclusion of detailed contracts is of crucial importance for the setting up of a medical research framework according to this approach. In addition a privacy committee shall be established being responsible for managing the exchange of

[159] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 46.

[160] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 46.

[161] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, p. 49.

[162] Reng et al., Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, pp. 49 ss..

[163] Pommerening, Reng, Debold, Semler, Pseudonymisierung in der medizinischen Forschung - das generische TMF-Datenschutzkonzept. GMS Med Inform Biom Epidemiol. 2005;1(3):Doc17.

[164] This is the typical use case for anonymity. As an example, one can imagine a simple statistical analysis of patient data.

[165] Pommerening, Reng, Debold, Semler, Pseudonymisierung in der medizinischen Forschung - das generische TMF-Datenschutzkonzept. GMS Med Inform Biom Epidemiol. 2005;1(3):Doc17.

and the access to patient data. In order to respect the patients´ wishes the patients should be granted the right to withdraw his/her consent to participate in the research study or to limit the storage period of the respective data. Respective contractual rules will be established for p-medicine by the contracts provided in this deliverable. The legal body concluding these contractual agreements will also be responsible for managing the exchange of and the access to patient data.

These TMF guidelines do not only provide general rules for the use of patient data for medical research but also for specific rules according to whether the networks are focused on clinical or scientific research. For p-medicine principles developed for research focused frameworks are relevant. Within this kind of projects high emphasis shall be put on the quality of the data. Accordingly a patient management shall be implemented in order to ensure that the data of the same patient can be stored together under a unique patient ID in the common data warehouse. This shall prevent the creation of different datasets for the same patient and thus increase the data quality on the long term. Moreover the project has to provide for a state of the art pseudonymisation in order to protect the identity of the patient.

6.7 Outcomes for p-medicine

Pseudonymisation and distinction between treatment domain and research domain

All the analysed projects showed that a clear distinction has to be made between the treatment zone and the research zone. Within the treatment zone the treating physician needs to identify the patient, so that the use of personal data can generally be based on the contract for treatment concluded with the hospital. In order to ensure the privacy of the patient already the patient data shall be stored in pseudonymised way already in the treatment context. According to all analysed projects the data may only be transferred from the treatment zone to the research zone through another pseudonymisation procedure or upon complete anonymisation, if there is no need to re-identify the patient in the future (e.g. about potential findings that could be of relevance for his/her case). According to the sensibility of the data and the steps of processing several pseudonymisation procedures may be necessary.

Patient identity management

The TMF (in cooperation with the German Data Protection Authorities) recognises the need for a patient management system for research-focused networks, as these networks shall build up a long-term database for research purposes. It, therefore, may be essential to attribute the same pseudonym to the same patient in order to combine data of treatments carried out in different hospitals or to combine the data of previous treatments with follow-up data. Thus the quality of the database shall be ensured. The patient identity management proposed by the TMF is based on patient list fed with a (relatively) uniform dataset in order to verify whether the patient already has been registered and attributed a PID. If this is not the case the PID-Generator calculates a PID for the patient that is then used as a pseudonym for the patient within the study. Another pseudonymisation procedure will follow when the data are transferred to the common research database. As within p-medicine also a patient identification management shall be established this approach will have to be taken into account. The inclusion of a patient in this patient list may only be carried out based on the patient´s informed consent.

Contractual agreements

In every project the need for clear contractual agreements for processing and the transfer of data is underlined. Therefore the data protection policies and contracts shall define the duties for the network staff and external partners in detail. Special emphasis is put on the access policies, meaning both the access to the network and the access to the patient data and human material. In some cases a request for access shall be approved an Ethical committee or advisory board of the project. As also within p-medicine an ethical committee is foreseen, its role within the access procedure established shall be considered.

Within BBMRI the basic guidelines and policies for the framework are laid down in the Partner Charta, summarizing the important key policies in a comprehensible and easily accessible form. This document has a very prominent place in BBMRI and it seems recommendable to provide such a text also for the use of patient data within p-medicine. However, this text does not necessarily have to be in a document on its own. Already within ACGT a general text had been provided as a preamble to the informed consent forms as well as the contracts concluded with the data exporting institutions and the researchers. In order to limit the number of documents that have to be agreed on and signed by the partners it seems recommendable to follow this approach also within p-medicine.

7 Data Protection and Data Security Framework for p-medicine

7.1 Introduction

In this chapter the data protection and data security framework for p-medicine shall be set up. Using the data protection and data security framework established within ACGT as a starting point, this framework shall be extended according to the needs and requirements of p-medicine. The ACGT framework has already thoroughly been evaluated in chapter 5 of this Deliverable, where we identified the strengths of the ACGT framework as well as the potential for improvement. In addition new legal challenges for p-medicine have been defined.[166]

On the basis of this evaluation we try to strengthen the elements that have been positively evaluated and to provide solutions for improvement for those ones that have been identified as weaknesses. Accordingly the key issues will be to maintain a high level of data protection, while at the same time simplifying the contracts between the partners and implementing a patient identity management system. Furthermore the requirements for a Trusted Third Party (TTP) will be evaluated and analysis will be provided as to the participation of third countries in the framework of p-medicine.

In compliance with the legal rules applicable to the p-medicine framework as well as in line with the standards set up in comparable research projects in the field of medical research involving patient data the framework of p-medicine will be based on the de facto anonymisation of all data processed for research purposes. Since compliance to data protection is of crucial importance for the scientific research involving medical data, the data protection and data security framework shall, however, not only be based on one single pillar. We rather aim to build up a “safety network” based on de facto anonymisation being only one of three pillars, the other two pillars being the informed consent of the patient and the existence of research exemptions provided by the national laws applicable to the respective data controller.

Finally it has to be pointed out that the range of p-medicine is wider than that of ACGT as regards the access to biobanks, patient empowerment and the legal analysis of the international clinical trials. These questions are subject to upcoming Deliverables of WP5.[167]

7.2 Data flows in p-medicine

In order to develop a data protection and data security framework for p-medicine, we need to identify the data flows in order to highlight where personal data might be exchanged and what safeguards need to be taken.

Health data of the patient is collected by the treating physician in the hospital and analysed and stored within the hospital. As the following figure of the p-medicine architecture shows these data shall be transferred to the infrastructure in two ways. Data shall be pushed to the so-called data warehouse, where data is stored. Furthermore data are transferred to the Clinical Trial Management System, where the data is stored and further transferred to the

[166] See chapter 5.4 above.

[167] The data protection law issues as well as the other legal questions related to the setting up of the biobank access tool for p-medicine will be analysed in D5.3. For a rough overview of the legal and ethical requirements for biobanks see chapter 6 of D10.1. The legal and ethical issues regarding patient empowerment are subject to D5.4. The legal and ethical issues regarding international clinical trials will be analysed in D5.5.

data warehouse. These databases serve as a basis for the tools provided within the p-medicine workbench. The users/researchers of the p-medicine infrastructure will have the possibility to work with these tools and data over the p-medicine portal.

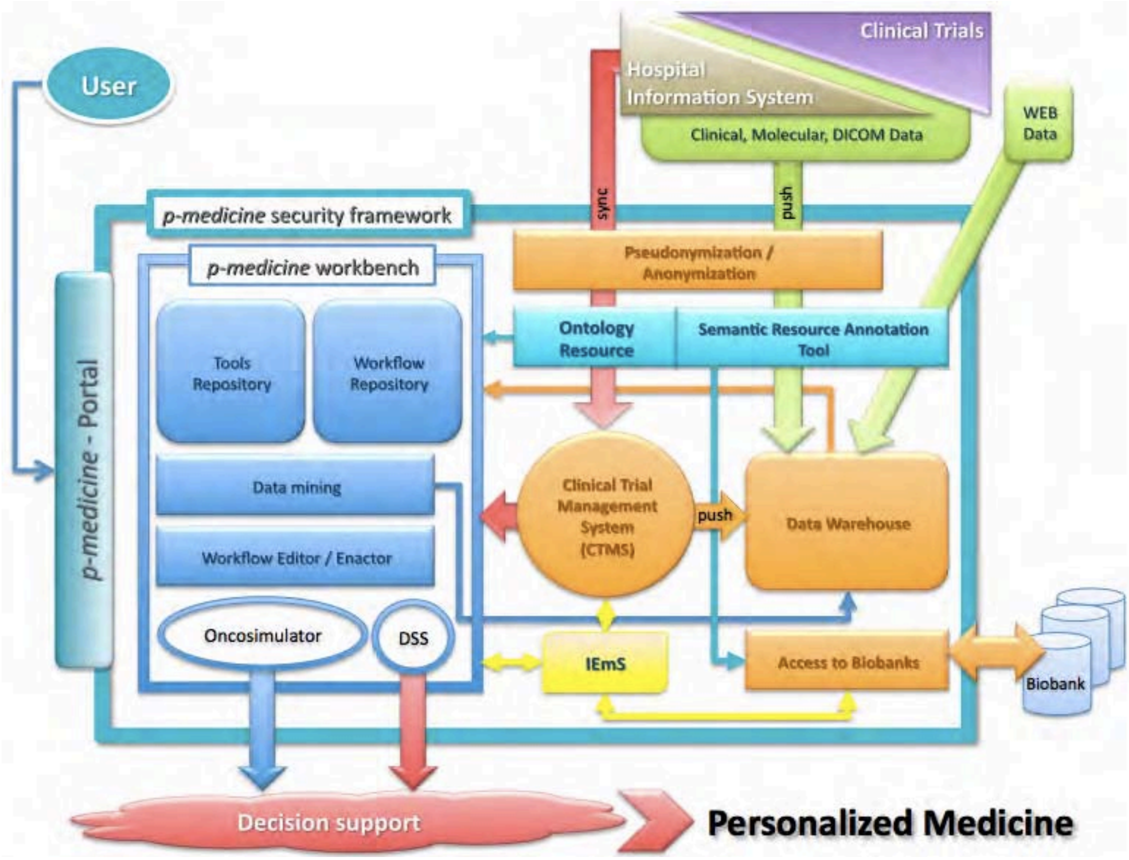

Figure 9: The architecture of p-medicine from a clinical perspective[168]

7.3 Data protection framework of p-medicine

7.3.1 General distinction between treatment domain and research domain

When establishing a data protection framework for clinical research a clear distinction has to be made between the treatment domain (in the hospital) and the research domain (p-medicine infrastructure).[169] This distinction is fundamental considering that different rules and data protection requirements apply to the storage and processing for the purpose of medical treatment and the storage and processing for the purpose of medical research.

7.3.1.1 Treatment domain

Within the treatment domain (hospital/healthcare organisation) the treating physician generally collects information about the patients in the course of medical treatment. As these data contain medical data (data on the treatment and the disease) and thus contain information concerning the health of a person, these datasets contain sensitive information (in the sense of Art. 8 para. 1 Directive 95/46/EC) about the person

[168] Source: p-medicine, Description of Work, p. 12B of 246.

[169] See also the outcomes of chapter 6.

concerned. Accordingly the special requirements for the processing of sensitive data stated in Art. 8 Directive 95/46/EC apply.[170]

The hospitals are obligated to store and process patient data in a pseudonymous way, whenever the physical examinations do not require the direct identification of the patient. However, the treating physician will generally have access to the link between pseudonym and real name. The collection, storage and processing of personal and pseudonymised patient data within the hospital is covered by the treatment contract with the patient as long as these data are necessary for the purpose of treatment. In addition the use of the patient data for the purposes of treatment will generally be covered by the patients´ informed consent.

Pseudonymisation within the treatment domain is generally carried out by the local pseudonymisation tools of the hospital. As a consequence it is very likely to happen that a patient that has been treated in different hospitals can have two or more different pseudonyms. The storage of data of the same patient under different pseudonyms in a large database without any possibility to verify that these datasets belong to the same patient can affect the data quality in a negative way influencing not only the statistical output of the database but also the long term prediction of patient treatment. In order to prevent these negative effects the hospitals will have the option to refer to the patient identity management system (PIMS) of p-medicine, which shall ensure that the same patient gets the same pseudonym when his/her data are transferred to the framework by verifying whether the specific patient had already been registered in the common PIMS-database and creating a new pseudonym only when there is no such patient registered so far.[171]

7.3.1.2 Research domain

For the purposes of scientific research data shall generally be used in anonymised form, whenever this is possible. In general anonymisation is the best way to protect the patients´ privacy.

Anonymisation, however, is not an option for the data in p-medicine for various reasons, the most important being that (complete) anonymisation can only be used when there is no need to re-identify the patient. It is, however, a clear task of p-medicine to provide for the re-identifiability of the patient in order to give the best treatment to the patient in case the research within p-medicine reveals that a certain therapy is highly effective for a certain disease. Therefore the data within p-medicine cannot be completely anonymised.

Hence the framework shall be based on de facto anonymous data. Pseudonymous data can be regarded as de facto anonymous data whenever the researchers working with the data do not dispose of the link back to the patient and there are certain data protection and data security measures in place that prevent the researchers processing from trying to re-identify the patients. This concept of de facto anonymous data held within a network of trust has been introduced in ACGT. This concept has already been described and positively evaluated in chapter 5 of this Deliverable (see 5.2.2.1). It shall, therefore, be followed as a guideline for p-medicine.

Accordingly a framework shall be set up where only de facto anonymous data are stored and processed. The de facto anonymisation will be achieved by a state of the art pseudonymisation of all data entering the “network of trust”. This network of trust will be built up by contractual agreements in which the participants declare to comply with data protection and data security standards of p-medicine. A central point within these

[170] See subchapter 4.2.1.3.2.

[171] For more information on the PIMS see 7.3.4.

contracts is the complete prohibition of any attempts of the researchers working with de facto anonymous data to re-identify the patients by matching the de facto anonymous data with other data sets or to transfer the data to any third person outside the network of trust. The necessary contracts will be concluded by the Center for Data Protection (CDP), a legal body that will serve as central data controller for all data entering the framework. Furthermore the CDP will serve as a central contact point for patients.

The patient data coming from the hospitals (already in pseudonymous form) will be de facto anonymised as soon as they leave the treatment domain by running through a second pseudonymisation procedure before being transferred to the research domain. The link back to the patient will be stored safely at a trusted third party, which serves as a “vault” for the links to the patient and does not dispose of any health data.

As the data within the environment controlled by the CDP can be regarded as anonymous data in the sense of the Directive, the processing within the framework does not fall into the scope of Directive 95/46/EC and, thus, may be processed without the processing restrictions of the Directive. The fact that de facto anonymised data does not underlie the rules for processing of personal data stated by the Directive, however, does not mean that there are no data protection and data security rules in place governing the data processing within p-medicine. These rules just do not directly derive from the data protection laws, but from the contractual agreements that are necessary to set up the network of trust within the project binding all partners to comply with the data protection and confidentiality rules of p-medicine. The main advantage from a research perspective is that the data can be used freely within the scope of the research purposes of p-medicine by all registered users of the p-medicine infrastructure. Moreover this provides for a high flexibility for the p-medicine infrastructure as a whole regarding the physical placement of the p-medicine databases. Accordingly the data can be shifted from one database to another database under the control of the CDP located in another Member State of the EU without any restrictions. However, the data can only be considered as de facto anonymous data as long as they are stored and processed within the de facto anonymous environment. A transfer of de facto anonymous data to third parties that are not bound to the data protection policies of p-medicine, however, would fall under the scope of the Directive.

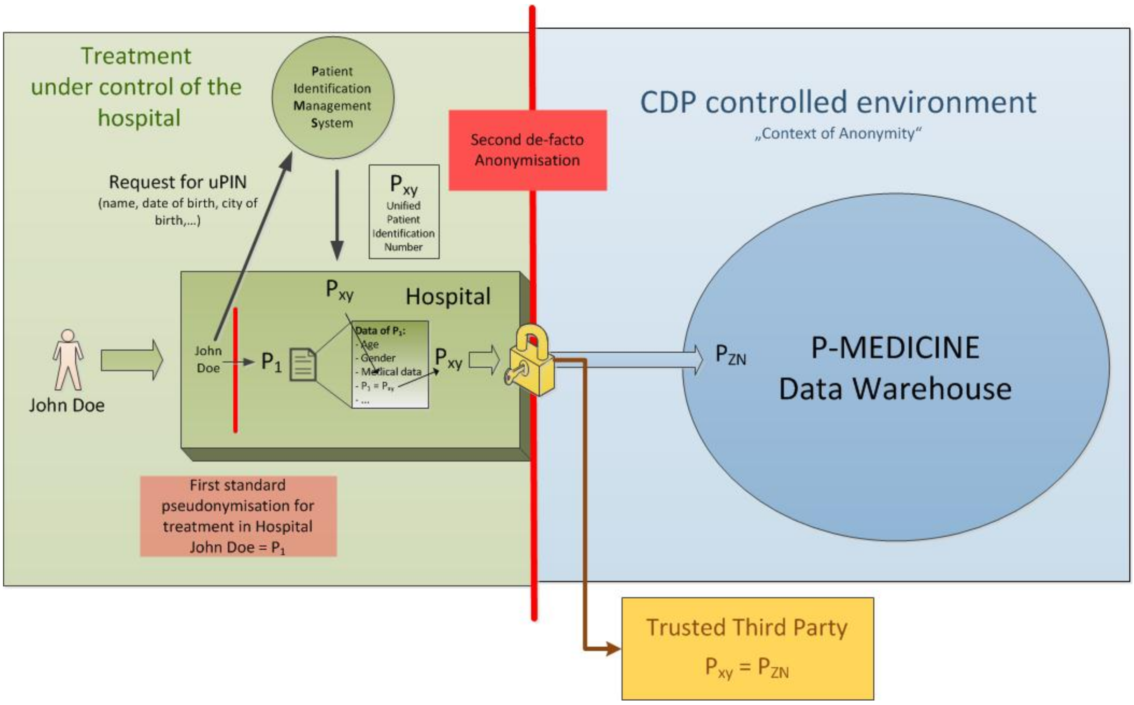
Treatment under control of the hospital
Patient Identification Management System
Request for uPIN
(name, date of birth, city of birth,...)
P_{xy}
Unified Patient Identification Number
Second de-facto Anonymisation
CDP controlled environment
„Context of Anonymity“
P_{xy}
Hospital
John Doe
P_1
Data of P_1:
- Age
- Gender
- Medical data
- $P_1 = P_{xy}$
- ...
P_{xy}
John Doe
First standard pseudonymisation for treatment in Hospital John Doe = P_1
P_{ZN}
P-MEDICINE Data Warehouse
Trusted Third Party
$P_{xy} = P_{ZN}$

Figure 10: Pseudonymisation, TTP, and PIMS

7.3.2 Pseudonymisation and de facto anonymous data

Within the hospitals patient data are stored in pseudonymised form, whenever the physical examinations do not require the direct identification of the patient. If the patient agrees to participate in a p-medicine trial, the physician transmits the respective medical data to the p-medicine data warehouse. During this transmission a state of the art pseudonymisation tool that guarantees an equivalent high pseudonymisation standard for all the data transmitted from the clinical partners to the p-medicine infrastructure will de facto anonymise the data. The standard anonymisation tool will be developed by Custodix and shall serve as a benchmark. Hospitals, however, are not bound to use the Custodix Anonymisation Tool (CATS), as long as they provide a state of the art anonymisation tool themselves that is approved to be sufficient by the CDP.

In both cases the pseudonym created by the state of the art pseudonymisation tool together with the pseudonym of the hospital will be sent to a Trusted Third Party (TTP). The TTP stores that link and assures that as few persons and entities as possible and only if necessary get access to data revealing the local (hospitals/PIMS) pseudonym of a patient. This link therefore will only be provided under certain circumstances. The data set containing medical data are transferred to the p-medicine infrastructure only under that pseudonym. Thus the data and the links that allow for the re-identification are stored in different databases. Whereas the data set containing medical data is stored in the p-medicine databases under the control of the CDP, the re-identification link to the first pseudonym of the patient is held by the TTP. The p-medicine end users will only have access to the data in the p-medicine database and will not be able to access the link stored at the TTP. Thus the data sets entering p-medicine can be considered as de facto anonymous from the moment of pseudonymisation.

7.3.3 Trusted Third Party

Within the research domain a Trusted Third Party (TTP), located between the persons concerned, respectively the institution holding the data and the researcher, serves as a vault for the pseudonymisation link. The researcher only receives data in pseudonymous form, so that he/she cannot directly link the data received to a specific person. For most researchers participating in p-medicine this information may not be needed as their research might also be carried out with anonymous data. However in certain cases, e.g. if a new treatment is developed that might help a specific patient whose data are stored in the database, a linkage back to the patient may be required. Therefore, it is the duty of the TTP to hold the link between the first (hospital/PIMS) pseudonym and the p-medicine pseudonym in order to safeguard such link and to exceptionally enable the de-anonymisation of the pseudonymised data. Further the TTP has to assure that the identity of a patient is only revealed under specific predefined circumstances.

In order to fulfil this function the Trusted Third Party has to meet certain requirements.[172]

The data custodian (or more precisely: the custodian of the link back to the patient) must be trustworthy and independent. Therefore it has to be ensured that only the Trusted Third Party can establish the link to re-identify the patient. The TTP may only do this under certain conditions, so that the patient can be assured that linking of information will not be undertaken under other conditions. Furthermore it has to be guaranteed that the information stored at the TTP is not used for any other purposes.

As the TTP shall act as a trustee, it is of high importance that it acts independent from every other participant of the project. It has to be a data controller next to the central data controller of p-medicine, not having to justify its decisions to the central data controller or anybody else. Only if the TTP is data controller itself it can be seen as a THIRD party guaranteeing the safety of the links to the data subjects as a security authority. The trustworthiness and independence of the TTP can be established by a contractual agreement or by legally binding rules provided by statutory law.

In order to ensure the protection of the linking information from third parties access, including the access by courts or the prosecution, it seems recommendable to choose a trustee that has the right to refuse to give evidence in court or is bound by special obligations of secrecy. Therefore, in many cases lawyers or notaries are selected as a trustee. In medical research it also can be thought of entrusting a medical doctor, who in general is bound to his/her medical secrecy as a physician. In this case however it has to be considered that a physician might eventually have own interests in the data what could compromise the trust of both the patients as well as the researchers in his/her neutrality. Furthermore it seems unlikely that the medical secrecy also comprises information about patients that do not result from his/her treatment but his/her role as a trustee in a research trial. Considering that it seems recommendable to choose a trustee that does not form part of the medical community.

Furthermore a trustee must not be bound by directives or has to be clearly separated from the researchers. Finally it shall be ensured that the TTP does not act as a trustee in a large number of medical research projects in order to prevent the creation of huge trust centers responsible for holding pseudonyms of many different research projects.

Custodix that shall serve as TTP within the p-medicine framework meets all these requirements. Custodix is an expert in data security who has already successfully served as TTP within ACGT. By its reliable and independent work as data custodian for the project Custodix has gained the trust of both the researchers and the clinical partners delivering the data to the common database. The independence of Custodix from the other partners of the project will be further guaranteed by a contractual agreement stating

[172] For further information and references see Pöttgen, Medizinische Forschung und Datenschutz, pp. 89 ss.

also the conditions and procedures for the de-anonymisation of patient data within p-medicine.[173]

7.3.4 PIMS

As pseudonymisation tools used by various hospitals differ, the same patient will be attributed different pseudonyms from different hospitals. In general this does not cause any problems, as it is not necessary to combine pseudonymised information about the patients from different hospitals in a common database The attribution of different pseudonyms for the same patient as well as the attribution of the same pseudonym to different patients may however affect the quality of the data in large scale collaborations like p-medicine, where the database does not only contain data from one single hospital but where geographically dispersed data are to be shared. In order to provide a valid dataset, in particular for the long-term evaluation of data on the effects of a specific treatment, a secure and lawful identity/pseudonym management system will be needed.

In order to fulfil this task a patient identity management system (PIMS) will be provided by Custodix. It is the main functionality to guarantee that the same patient gets the same pseudonym when his/her data are transferred to the framework. This is done by checking whether the specific patient had already been registered in the common PIMS-database. A new pseudonym will only be created when there is no such patient registered so far. If the patient had already been registered the existing pseudonym will be attributed to him/her. Thus PIMS avoids the creation of different pseudonyms for the same patient (synonyms) as well as the creation of the same pseudonyms for different patients (homonyms).

The technically easiest way to provide such a tool is to create a common database, where the names and former names (e.g. maiden name) of the patient, date and place of birth, and other identifying information of every patient is stored the first time he/she registers. A unique pseudonym for each newly registered patient would then be created. If a patient decides to participate in a clinical trial, these data would be sent to that common database, where the system will check whether a corresponding dataset already exists. If this is not the case a new pseudonym will be created and sent to the hospital. Otherwise the hospital will receive the already existing pseudonym.

From a data protection perspective it has to be pointed out that this approach requires the transfer of personal data of the patient to a third party holding the common PIMS database. Such a transfer of personal data requires a legal basis, which will have to be the patients´ prior informed consent, since the data protection law does not allow for the creation of a comprehensive patient identity management database for various hospitals. Thus, this tool can only be optional depending on the patient’s consent.

Another possibility might be to generate a unique value on the basis of the said data set that is stored in the common PIMS database, given that only on the basis of this value it is possible to verify whether the patient has the same personal data. These values must be unique for a patient with exactly the same personal information and provide for possibilities to match patients-values that slightly differ, but might belong to the same person (e.g. if there is a spelling or typing error or the person has changed the name in case of a marriage etc.). From a legal perspective this second alternative on the one hand has the advantage that no personal data has to be transferred outside the hospital. On the other hand it also has to be taken into account, that the data quality within p-medicine database may be influenced in the negative if the matching of these values in PIMS may not lead to reliable results. Whether the implementation of a PIMS based on this

[173] See also 7.3.7.3.

technology would deliver reliable results is subject to on-going research by WP5, so that at the current stage only the first option is available.

This pseudonym created for the patient by the PIMS will then be transferred to the hospitals where it is stored together with the medical patient data. The PIMS pseudonym shall however not be used for the internal use within the hospital, but shall rather be reserved for the transfer of data to the research domain solely. In order to ensure the de facto anonymity of the patient data delivered to the p-medicine data warehouse the already pseudonymised data will now be pseudonymised a second time by a state of the art pseudonymisation tool. This tool guarantees that each pseudonym created by the PIMS will be converted to a corresponding pseudonym for the p-medicine data warehouse, so that data belonging to the same patient can be stored under the same pseudonym although they are delivered from different hospitals.

7.3.5 Center for Data Protection (CDP)

In order to ensure the de facto anonymity of the data transferred to the p-medicine framework a so-called “network of trust” shall be established. Every participant in this network shall be obligated to respect the patients´ privacy rights and to comply with the necessary data protection and data security rules. In particular it shall be assured that the partners and users within the network do not undertake any measures to reveal the patients identity by matching patient data transferred to the p-medicine data warehouse or received from it with other data (personal) data sets.

Such a “network of trust” could be achieved by contractual agreements between all participating partners and end users. This would, however, multiply the contracts that need to be concluded. In order to prevent this, a central body will act on behalf of the consortium. This central body will conclude all contracts so that all participants only have one contractual partner. Furthermore that central body shall be responsible for the data processing within the data warehouse and controls compliance of the participants with the data protection and data security policies set up.

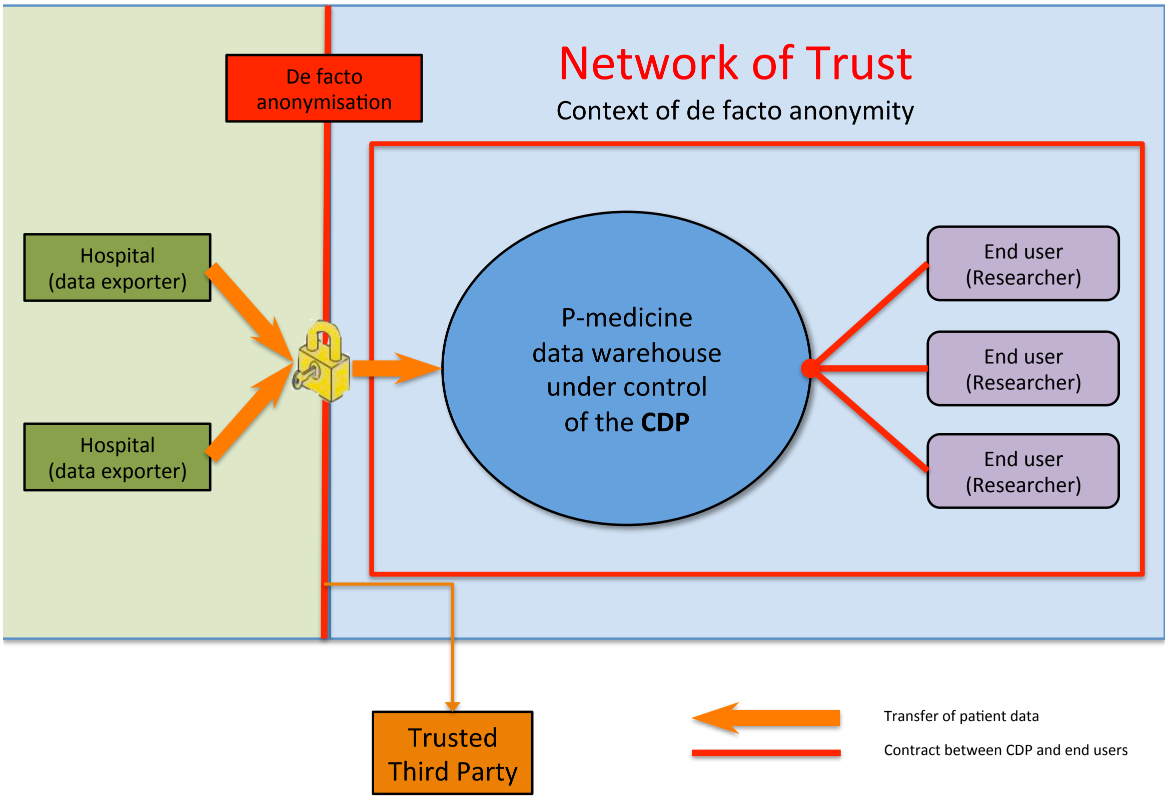

Figure 11: Network of Trust

For building up such a data protection network a legal entity is required that is both legally able to conclude binding contracts with the participants in the project and empowered to inflict a penalty for infringement. In order to conclude the necessary legal agreements and to inflict claims for violation of the policies agreed, this entity has to be carrying legal personality. In addition this legal body needs to be empowered by the consortium. As this entity is going to act on behalf of the consortium and as central data controller for the framework, it needs to be independent in its decisions.

Already within ACGT such a body had been founded in August 2007. It is a non-profit organisation under Belgian law named "Center for Data Protection".[174] The work of the CDP within ACGT has been evaluated positively under chapter 5. We therefore decided to include such a legal body also for p-medicine. As p-medicine is a follow-up project of ACGT, a new creation of a very similar legal body shall be avoided in order to save time and resources. We therefore decided also for p-medicine to assign this task to the CDP. Accordingly the CDP will serve as the data controller for the data transferred to the p-medicine infrastructure and therefore has to ensure the compliance to the current data protection legislation.[175]

[174] For more information please see http://www.privacypeople.org.

[175] For a further description of the role of the CDP as a central data protection authority within p-medicine see 7.3.7.2.1.

7.3.6 Re-identification procedure

As explained in chapter 7.3.2 above, p-medicine shall offer a possibility to re-establish a link from the de facto anonymised data back to the patient in the event that the research reveals findings that might be beneficial for the treatment of the patient concerned. In order to protect the patient's privacy rights, the end user, i.e. researcher using patient data received over the p-medicine data warehouse, cannot directly address the patient. Re-identification within p-medicine rather requires four steps, so to ensure that the patient's identity is not disclosed to the p-medicine end users.

In a first step a researcher has to address the CDP and to explain the reasons for the request to re-identify the patient concerned. Re-identification shall only be granted where it is clinically beneficial and ethically appropriate for the patient. Accordingly the re-identification procedure shall only be initiated when the researcher shows credibly that feedback would be in the best interest of the patient. When this requirement is fulfilled the CDP will send a request to the TTP regarding the key to the patient concerned. In a third step the TTP will transmit the original pseudonym received from the hospital when the data was transferred to the p-medicine data warehouse to the hospital where the patient was treated. Since the TTP only disposes of a pseudonym of the patient still at this stage within the re-identification procedure no information that could be used to directly identify the patient concerned, such as the name and address of the patient, are processed. Finally the hospital informs the patient about new developments that could be beneficial for his/her treatment. The patient will only be addressed if he/she had agreed to receive feedback within the consent given when agreeing to participate in the trial. The treating physician will then give the feedback.

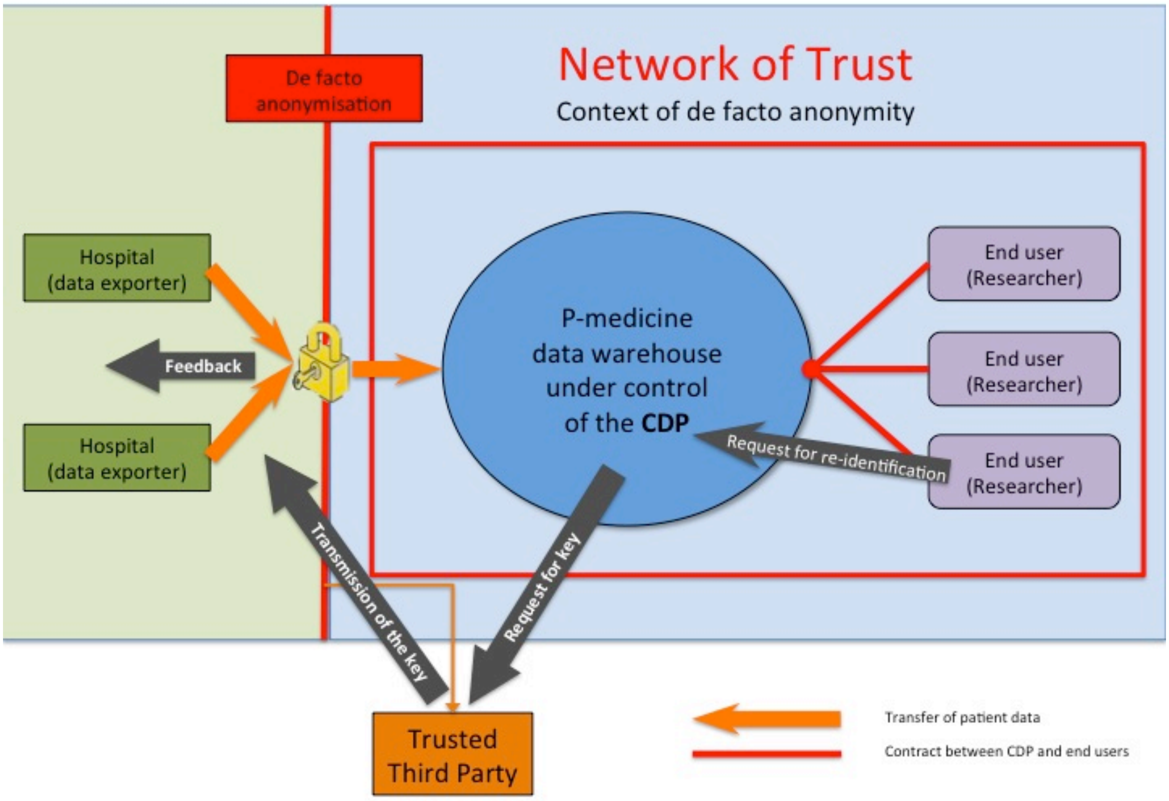

Figure 12: Re-identification procedure

7.3.7 Initial contracts

This subchapter shall analyse which contracts have to be concluded in order to establish the legal framework for p-medicine and draft (with respect to the infrastructure still in progress) initial versions. Furthermore we will introduce the essential provisions of the contracts drafted, such as the question of data control within p-medicine, obligations concerning the de facto anonymity, third beneficiary rights and the applicable law and we will explain for what reasons it has been decided to design the different provisions the way they are.

As stated above the p-medicine data protection and data security framework will be based on a de facto anonymisation of all data entering the network, the implementation of a Trusted Third Party and the establishment of a data protection authority within p-medicine (Center for Data Protection, CDP) responsible for signing the contracts on behalf of the consortium for the processing of data within p-medicine.

In order to set up this framework three contracts are needed. Two of these contracts will regulate the transfer of data to the p-medicine framework (Data Transfer Agreement) and the use/processing of these data by the researchers for purposes of scientific research (End User Agreement). The Data Transfer Agreement will be concluded between the CDP and the healthcare organisations/hospitals transferring data to the p-medicine databases. The End User Agreement will be concluded between the CDP and all p-medicine end users doing research on these data. It might very well be the case that clinical partners in p-medicine not only deliver data to the infrastructure but also use the p-medicine network to receive more data from other clinical partners for research, so that their participation in p-medicine could also be regulated in one single contract. However, it does not seem recommendable to provide only one contract regulating the transfer to the p-medicine network and the use of these data for research. This decision takes into account that the obligations that have to be established for the transfer on the one hand and the processing on the other hand differ fundamentally so that it would be too complex, even confusing, to rule everything in only one contract. Furthermore, we would also like to provide the possibility to either only deliver data to the p-medicine framework or to merely receive patient data for research, so that in any case two agreements will have to be offered to potential participants of the framework.

A third contract is required in order to regulate the duties and the role of the Trusted Third Party (TTP) within p-medicine. This contract shall mainly guarantee the independence of the TTP from the other partners of the project. In addition it shall regulate the conditions and the procedure for the de-anonymisation process as well as rules regarding the storage of the links, the access control to the database and other data security issues.

In the following the main rules and the concept of these contracts will be outlined. The text of the initial contracts will be added as an Annex to this document.

7.3.7.1 Data Transfer Agreement (Annex A)

7.3.7.1.1 Scope

The Data Transfer Agreement (DTA) to be concluded between the CDP and the hospitals/physicians does not cover the framework as a whole, but rather sets the rules and provides for guarantees and measures in order to protect the patient´s privacy rights. According to the data protection laws the person or entity holding data under its control is responsible for the compliance to the data protection rules. The transfer of data from the hospitals to the p-medicine data warehouse leads to a change in responsibility. Therefore it is of utmost importance to know when the responsibility for the data shifts from the data exporter to the CDP.

Clause 2 shall clarify that the processing of data within the hospital domain will be under the sole responsibility of the hospital. This includes the data transfer to the p-medicine data warehouse. As soon as the data are successfully transferred to the data warehouse their processing is under the sole responsibility and control of the CDP.

7.3.7.1.2 Data exporter – legitimate data controller

The (de facto) anonymisation in order to transfer data from the hospitals to the p-medicine database constitutes a processing of (personal) data in the sense of Art. 2 Directive 95/46/EC and therefore needs a legal permission. One possible way to receive that permission is to obtain the patient´s consent, which is needed for ethical reasons anyway. Therefore, a patient willing to participate in a p-medicine trial needs to sign a consent form after having received all information wanted from the treating healthcare organisation regarding the processing of his/her data within p-medicine. The patient´s informed consent is the basis for the participation in p-medicine. Hence, it has to be provided that the clinical partners transferring data to p-medicine are obligated to transfer data only upon the informed consent of the patient. In order to prove this, the clinical partners shall additionally be obligated to send copies of the consent forms signed by the patients to the CDP. Furthermore a transfer of data is only licit if it the data transferred has been collected and processed in accordance with the applicable data protection law by the transferring party.

By Clauses 3.1. and 3.2. the data exporters therefore have to warrant and guarantee that the data transferred to the p-medicine data warehouse have been collected, processed and transferred in accordance with the laws applicable to the data exporter and that the patient has expressed his/her informed consent in writing to this transfer (and the pseudonymisation of their personal data). Copies of the signed consent forms have to be sent to the CDP (3.7.).

7.3.7.1.3 Pseudonymisation

The first pillar of the setting up of the de facto anonymous data environment is the pseudonymisation of all data entering the research domain. For these data a state-of-the-art pseudonymisation will be required in order to guarantee an equal level of pseudonymisation for all data within the p-medicine infrastructure.

With the Custodix Anonymisation Tool (CATS), such a pseudonymisation tool will be made available to the data exporters. However the hospitals shall be free to choose another state-of-the-art pseudonymisation tool. In this case the use of this tool needs to be indicated to and accepted by the CDP. The CDP will accept any pseudonymisation tool chosen by the data exporter, provided that it ensures a state of the art pseudonymisation, regarding CATS as a benchmark.

7.3.7.1.4 Patient Identity Management System (PIMS)

The patient identity management system for p-medicine will be developed in course of WP5 and can be implemented in the treatment domain under the responsibility of the participating hospitals. The use of PIMS is not obligatory for the hospitals.

The management of patient identities in general requires a central body (e.g. a participating hospital) holding a database where each patient is registered by name and other identifying information. The hospitals will have to conclude a contract with this central body regulating the terms and conditions of the participation at the PIMS, when they decide to use PIMS. This contract is outside of the scope of p-medicine.

The transfer of personal patient data to the PIMS constitutes a transfer of data in the sense Art. 2 lit. a Directive 95/46/EG. Accordingly the transfer may only be effected if there is a legal basis for it. In this case the legal basis will have to be included in the generally given consent of the patient. Therefore the hospital is only allowed to transfer personal patient data to the PIMS on the basis of the informed consent of the patient.

Furthermore, if the hospital decides to use PIMS, it shall only use the pseudonym generated by the PIMS for transferring patient data to the data warehouse. Whereas for storing patient data for treatment purposes the hospitals shall on the other hand only use the internal pseudonym of the hospital.

7.3.7.1.5 Third beneficiary rights

In general only the parties of a contract can enforce clauses of a contractual agreement, whereas the third party normally could only refer to rights provided by general legal rules or own contractual agreements. The contracting parties may, however, provided that a third party not being part of the contract may also enforce certain clauses of a contract. Such a contractual provision is called third party beneficiary clause.

The contract between the data exporter and the CDP contains such a third party beneficiary clause, granting the patients the right to enforce certain clauses of the Data Transfer Agreement against either the data exporter (e.g. hospital) or the CDP, if the patient´s rights guaranteed by these clauses are violated by one of these parties. Thus the patients can directly deduct rights from the contractual agreement between the data exporter and the CDP, although they are not a contracting party. Accordingly the inclusion of a third party beneficiary clause constitutes an enhancement of the patients´ rights, which is supposed to generate trust in the p-medicine project.

7.3.7.1.6 Applicable law

When concluding contracts in an international environment the question of the applicable law is of high importance. We distinguish between rules stating which national law will be applicable to data protection issues and rules stating which will be the national law applicable to law of obligations issues.

According to Art. 4 para. 1 Directive 95/46/EC the data protection law of that state is applicable on which territory the data controller is established. This provision is binding and cannot be changed by any contractual agreement. The data controller for p-medicine is established in Belgium, Belgian data protection law is applicable for all processing operations the CDP is responsible for. For the processing of patient data under the responsibility of the hospitals (including the transfer of the data to that p-medicine data warehouse) the national data protection law of that state is applicable in which the hospital is established.

Regarding the applicable law of obligations a choice of law can be made between the parties. This choice of law is generally valid when it is taken between legal entities. To determine the applicable law by contractual agreement also seems recommendable as otherwise the applicable law can vary according to the claims and the seats of the parties involved, what would cause legal uncertainty for all participants. Hence it was decided to include a choice of law in this contract. As the legal partner of the consortium is expert in German national law, it was decided that German law should govern these clauses.

As regards the competent jurisdiction it was decided that the courts of Hanover/Germany should have exclusive jurisdiction with regard to this contract. The courts of Hanover then might have to use Belgian data protection law, but the main legal issues the court would have to rule about in the context of this contract, would very reasonably be in the field of law of obligations, which will be governed by German national law.

7.3.7.2 End User Agreement

The second important contractual agreement to be concluded is the contract between the CDP and the end users/researchers using the data via the p-medicine infrastructure. This contract is of high importance for the p-medicine framework, as these provisions shall ensure that only de facto anonymous data are processed within the framework. In addition this contract shall guarantee the patient's rights of access to data as well as transparency of the processing and use of data at the same time.

7.3.7.2.1 CDP as central data protection authority

Already in ACGT two possible approaches for the role of the CDP within the ACGT framework have been thoroughly analysed.[176] According to the first approach the CDP would have been the central data controller for all data within the framework whereas the end users working with the data would only have been treated as data processors working with the data on behalf of the central data controller.[177] The advantage of the approach was a clear and transparent responsibility for the processing of the patient data within p-medicine, as only one entity would have been responsible for all the processing of data within ACGT. In addition the patients would have had, already by the law, only one data controller that is responsible for the processing of the data and thus could be held liable for possible data protection law infringements.

However, this approach was not pursued, as a data processor is defined in Art. 2 lit. e Directive 95/46/EC as a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller, meaning that data processor shall only act on instructions from the data controller (see Art. 17 para. 3 Directive 95/46/EC). In ACGT the end users were researchers that, of course, had to decide on their own how to do their research.[178]

Therefore, a second approach had to be developed in which the data researchers themselves are also qualified as data controllers. The CDP however acts as a central data protection authority. This approach confers more responsibility on the researchers as end users but also more freedom in how to deal with obligations provided for by the data protection legislation. The advantage of this approach is that the end users are responsible for the data processing and as such are also free to choose how and with which means they want to process data received, within the limits of data protection legislation and the limits of this contract.

This concept shall also be followed within p-medicine. Accordingly the CDP shall only be the central data controller for data stored within the p-medicine data warehouse as well as for the transfer to the end users, whereas the p-medicine end users shall be responsible for the data processing within their own entities. Thus, the end users of p-medicine are to be considered data controllers with respect to the data accessed over the p-medicine infrastructure. Accordingly the end users have to be compliant with data protection legislation applicable for their entity.

From the patients´ perspective this approach bears the problem that it might be unclear which end user is actually processing their data. This is problematic, in particular, when the patient wants to execute a right deriving from data protection law (e.g. right of access), as these rights have to be executed against the data controller. A patient therefore would have to find out which end user in concreto processes his/her data before being able to execute his/her rights. Besides the patients would need to find out, in which country the actual data controller is established, as the national data protection legislation of that state is applicable for any dispute regarding data protection law. Therefore, at first sight, this approach suffers intransparency.

[176] See ACGT D10.4, pp. 17 ss., available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf..

[177] ACGT D10.4, pp. 17 ss., available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf..

[178] ACGT D10.4, p. 18, available at: http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf..

Granting additional rights to the patients had evened out these disadvantages. In order to ensure the transparency of the data processing and to facilitate the exercise of the patient´s rights as a data subject, it was provided in the contract between the CDP and the p-medicine end users that the patients can exercise their rights (e.g. right of access) not only against the actual data controller but also against the CDP as central data protection authority. To enable the CDP to fulfil these obligations the end users are obliged to provide the CDP with all necessary information. Moreover, the patient can also sue the CDP for any damage caused by any unlawful processing of his/her data by the end user. Of course the end user will then have to compensate the CDP for this damage. By introducing these provisions, the patient gets a central contact point for all data processing within the p-medicine framework.

7.3.7.2.2 Ensuring the de facto anonymity

As the end users within p-medicine will process data on their on behalf and responsibility it is important that they are bound to contractual rules guaranteeing the de facto anonymity of the data processed. Taking into account that these data are mostly genetic data that are unique and therefore can be linked to the person concerned if a reference data set is available, it is of crucial importance that neither the CDP nor the end users are allowed to undertake any measures in order to re-identify the patient by e.g. matching the data sets within the framework with other data sets. This is especially true for end users working in hospitals that also deliver data, as they of course have corresponding data sets with the clear name of the patient. As a consequence they could easily identify the patient concerned just by carrying out such matching procedures.

Another key element is, that the de facto anonymous data is only accessible to a closed user group, meaning all project participants (like the CDP, the p-medicine end users) and all people that have to be attributed to them (such as data processors in the hospitals). The data received over the p-medicine infrastructure must not be disclosed to anybody outside the framework who is not bound by contractual provisions and guarantees provided by this contract. Accordingly the p-medicine end users shall not be allowed to publish any data received from p-medicine to any third party.

7.3.7.2.3 Applicable law

Regarding the applicable law it has again to be distinguished between the national law that is applicable to data protection issues and national law that is applicable to law of obligations issues.

The applicable national data protection law is the law of the Member State in which the data controller is established. For p-medicine this means that the processing of data under the control of the CDP (established in Belgium), Belgian data protection law will be applicable, whereas for the processing of data under the control of the end users within p-medicine, the respective national data protection law of the country applies in which the end users are established. As this would in most causes problems for a patient wanting to access his/her personal data, the patient by way of contractual agreement should be given the opportunity to address the CDP in any question of access, rectification or violation of privacy beside his/her rights against the particular p-medicine end user. The applicable law in such cases where the CDP acts as contact point for the patient arises from the law of obligations.

Just as the data transfer agreement, also the end user agreement consists not only of data protection provisions, but also of provision in the field of the law of obligations. For the same reasons as explained above,[179] it was decided that German law should

[179] See 7.3.7.1.6.

govern this contract (except for the data protection provisions). Accordingly also the courts of Hannover/Germany shall have exclusive jurisdiction.

7.3.7.3 Contract with the TTP

As the TTP shall act as a trustee, it is of high importance that it acts independent from every other participant of the project. It has to be a data controller next to the central data controller of p-medicine, therefore not having to justify its decisions to the central data controller or anybody else. The role and the duties of the TTP shall however be clearly defined by a contractual agreement with the central data controller. This contract should include the conditions and the procedure for the de-anonymisation process as well as rules regarding the storage of the links, the access control to the database and other data security issues. Furthermore it must be ensured the data custodian must not use the data for other reasons. For the purposes of keeping proof, the parts of the contract or the legal act relating to data protection and the requirements relating to the technical and organisational measures shall be in writing or in another equivalent form.

7.3.8 Informed consent

Although the data protection framework of p-medicine is based on the de facto anonymisation of all data entering the p-medicine infrastructure, anonymisation being the best way to protect the patients´ privacy rights, the patients consent to the use of his/her data within the framework is an imperative precondition. Already from an ethical point of view it shall be provided that the data collected in the course of medical treatment shall only be used for the purposes of scientific research if the patient agreed to this. It is a vital part of the respect regarding the patient´s self determination that data shall not be processed, even in a de facto anonymous way, against the expressed wish or without the knowledge of the patient.

Furthermore the patient´s consent to the processing of his/her data for the purposes of scientific research is the first fall-back scenario of the data protection framework set up for p-medicine. In order to obtain the patient´s consent the patient has to be provided all the information necessary to understand what will be done with his/her data. Accordingly the requirement of informed consent will ensure that the use of their data is transparent to the patients. Therefore the requirement of informed consent will not only prevent legal uncertainty but also generate trust among the patients.

In this context it has to be taken into account that this empowerment to consent to the processing of his/her data cannot be seen as unlimited or under no control. Already the Directive 95/46/EC establishes certain requirements, which have to be met by the patients consent. First the consent has to be indubitable, indisputable, without any doubt and freely[180] given. Moreover the consent of the data subject has to be specific and informed. The patient must know exactly what he/she consents to. In the context of p-medicine, information inter alia concerns the processing of the de facto anonymisation of the data by the implication of a trusted third party, the use of the anonymised data and the rights of the patient as provided by the law as well as by the contractual agreement that build up the network of trust.[181]

[180] In this regard the consent has to be free of any vice, constraint or pressure.

[181] C.f. See also Forgó (ed.), http://eu-acgt.org/uploads/media/ACGT_D10.2_IRI_Final_01.pdf, pp. 15 ss..

7.3.9 National exception

For the unlikely event that for a specific patient the safety net build up by both the de facto anonymisation fails and the processing of the data is not covered by the consent of the patient because it too narrow, invalid, or does not exist, the national law may provide for exemptions according to Art. 8 para. 4 Directive 95/46/EC for the use of patient data for the purposes of scientific research that could serve as a second fall-back scenario. According to that provision the Member States may, for reasons of substantial public interest, lay down further exemptions from the general prohibition on processing of sensitive data, e.g. for scientific research.[182]

However, Member States are free to implement such exemptions and to remove them at any time. Hence the legal situation in the respective countries does not only vary but also underlies the possibility of changes over the time. Whether the Member State whose law is applicable for the processing of data in question has introduced such an exemption in its national law, would have to be analysed individually if needed.

7.3.10 Transfer to third countries

As the Consortium does not only consist of partner residing in a Member State of the European Union, the rules on the transfer of data to third countries have to be taken into account.

Under the Directive 95/46/EC a transfer of personal data to a non EU-Member State may – as a general rule – only take place if the country in question ensures an adequate level of protection. The European Commission issues decisions (binding for the Member States) recognizing the adequacy of the data protection level provided by third countries.

Where a third country does not ensure an adequate level of protection, Member States may authorize the transfer of personal data to that third country where the controller adduces adequate safeguards with respect to the protection of privacy and data protection standards. Such safeguards may result from appropriate contractual clauses. In this case, the transfer of data to a third country in general may only be effected after the competent data protection authorities have examined if the clauses guarantee an adequate level of data protection (authorisation required!). This examination and authorisation procedure is often complex and time-consuming.

In order to facilitate the transfer of data the European Commission has issued standard model clauses, which guarantee an adequate level of data protection by the recipient of transmitted data. In this case the transfer of data can be effected on the contractual clauses without prior examination by the competent data protection authorities. Therefore the conclusion of standard model clauses is the easiest and fastest way to contractually establish a basis for the transfer of data to a third country. However it has to be pointed out, that these standard contractual clauses may only be used without prior authorization, if they are not at all modified.

Currently there are two decisions laying down standard clauses for the transfer of personal data to controllers outside the EU (Decisions 2001/497/EC[183] and 2004/915/EC[184]). In addition there is a decision of the European Commission on standard

[182] See Recital 34 Directive 95/46/EC.

[183] Available under: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:181:0019:0031:EN:PDF.

[184] Available under: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:385:0074:0084:EN:PDF.

contractual clauses for the transfer of personal data to processors established in third countries (Decision 2010/87/EC[185]).

7.4 Data security Framework for p-medicine

The p-medicine architecture is, contrary to ACGT, not a GRID-based architecture. It will be designed around a REST[186] architecture with loosely coupled components. Therefore, although the experience gained in ACGT is very useful, the ACGT data security framework (with a GRID security middleware) itself will not be directly reused. Instead p-medicine will be designed around a lightweight dynamic architecture, allowing it to evolve over time according to newly arising requirements. It will consist of modular re-usable components dealing with authentication, authorisation, auditing, de-identification, etc. P-medicine can hereby more easily opt for widely used industry standards instead of being locked in by the technology choices made by the middleware solution.

7.4.1 Security Architecture

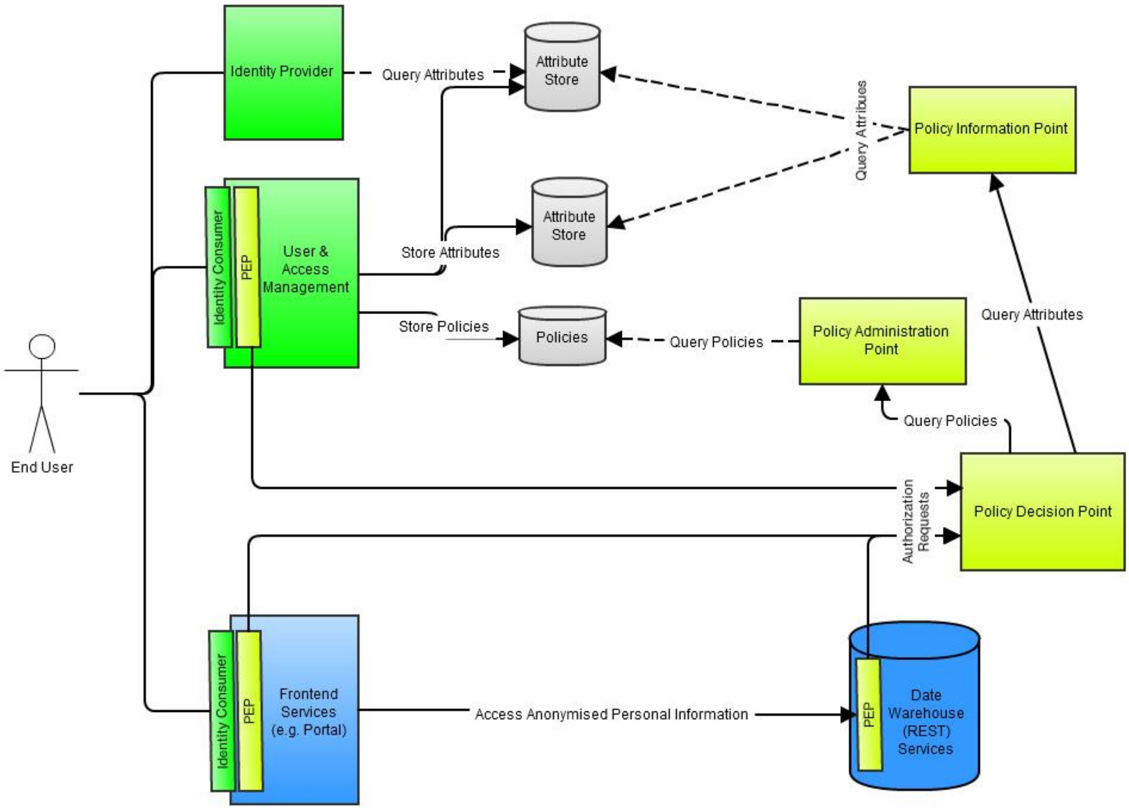

Figure 13: Security Architecture

Major components in the architecture are:

[185] Available under: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:039:0005:0018:EN:PDF.

[186] Representational state transfer (REST) is a style of software architecture which consists of clients and servers. Clients initiate requests to servers which process requests and return a response. Requests and responses are based around the transfer of representations of resources, typically the current or intended state of resources.

- The Identity Provider (IdP) is a service provider within a federation responsible for authentication. It provides identity assertions to other service providers.
- An Identity Consumer is a software component that is part of a service provider. It consumes the assertions provided by the Identity Provider. It will verify the received assertion and pass it to the service provider's application layer.
- A Policy Enforcement Point (PEP) is a software component which requests and enforces authorisation decisions.
- A Policy Decision Point (PDP) is an entity that makes authorisation decisions. A PDP accepts authorisation requests and will make a decision based on policies fetched from a Policy Administration Point (PAP).
- A Policy Information Point (PIP) is an endpoint which provides missing information to a PDP i.e. attribute information. For example if a policy requires information on a specific attribute which has not been provided with the authorisation request, a PDP might request a PIP for information on that attribute.
- A Policy Administration Point is an endpoint, which manages policies. It will provide a PDP with all policies required to produce an authorisation decision.
- The user & Access Management will probably consist of the following two mayor components:
  - A User Enrolment & Management Service where users can be enrolled, revoked, edited, etc.
  - An Authorisation Rule (Policy) Management Service where authorisation rules can be configured generating authorisation policies.

7.4.1.1 Authentication and Single Sign-on

When a client accesses a p-medicine service provider (SP) (e.g. the p-medicine portal) where he/she has no local active authenticated session, the client will be requested to pass a p-medicine identity assertion. The client should then request this assertion from the p-medicine identity provider (IdP). If he/she is already authenticated, the IdP will provide the identity assertion (Single Sign-on), if not, the client will first have to authenticate himself. The client will then pass that assertion to the SP he/she originally wished to access, which will verify the assertion and give the client access if it is evaluated as valid.

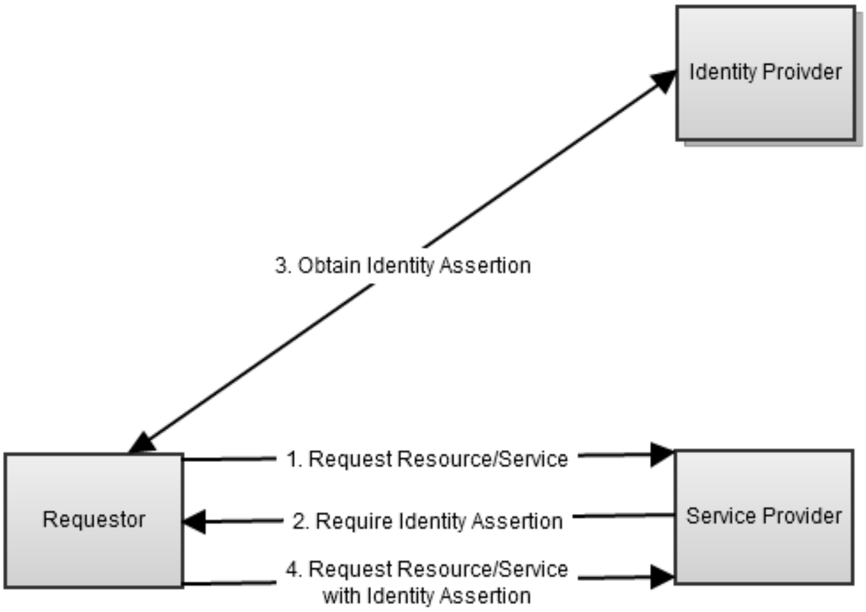
Identity Proivder
3. Obtain Identity Assertion
Requestor
1. Request Resource/Service
2. Require Identity Assertion
4. Request Resource/Service with Identity Assertion
Service Provider

Figure 14: Single Sing-on

The Security Assertion Markup Language (SAML)[187] is a commonly used standard to provide the SSO functionality. It defines an XML-based protocol, making it possible to exchange authorisation and authentication data between one or more security domains. SAML provides, amongst others, bindings for single sign-on (e.g. the HTTP Redirect Binding).

7.4.1.2 Authorisation

Each time a subject wishes to access (create, read, update, delete) anonymised sensitive patient information or results calculated from it, the p-medicine central Policy Decision Point (PDP) should be called. The PDP will then fetch the configured data protection policies from the Policy Administration Point (PAP) to decide whether the subject is allowed to access the requested patient information. To enforce patient empowerment the PDP can also fetch the consent policies, which were configured by the patient. This way the consent the patient gives, is automatically evaluated during authorisation by the PDP.

Policy Enforcement Points (PEP) will be made available which integrate easily with the service providers that need to make the authorisation calls (e.g. a PEP can be a servlet filter, spring annotation, etc..

Policies, decision requests and decision responses will be defined in the eXtensible Access Control Markup Language (XACML)[188], which is an XML-based language. XACML is the current industry standard for which multiple, both open as commercial, implementations are available. XACML is based on the attribute-based access control (ABAC) model in which the attributes that are associated with a user, action or resource serve as inputs to the authorisation decision. ABAC is inherently capable of meeting many of the "modern" access control demands (e.g. data or environment dependent access policies).

[187] p-medicine; 2011; D3.1 State-of-the-Art report on Standards, Chapter 8.1.1 SAML

[188] http://www.oasis-open.org/committees/xacml.

7.4.1.3 User & Access Management

For user enrolment and management a user friendly front-end to the attribute stores (i.e. LDAP containing the identity information) will be developed. The Authorisation Rule (Policy) Management Service is a front-end to the authorisation policy store or Policy Administration Point (PAP). In this way administrators will be presented with an intuitive user-friendly authorisation configuration site, which generates the policies, instead of having to edit cumbersome (i.e. XML) policies. This shields away the complexity of the policy files from the users (in this case p-medicine administrators). The policies generated are then uploaded to the Policy Administration Point so that when an authorisation decision needs to be made, the Policy Decision Point can retrieve the most up-to-date applicable policies.

7.4.1.4 Credential Forwarding and Delegation

Through credential forwarding a service A can access a service B on behalf of an end user U. As credential forwarding means that the identity assertion of U received by A is just forwarded as is to B, the service B will think that the user U directly contacted him/her without knowing that there is a service A lying between. Forwarding is not possible though if the identity assertion is limited in audience to A or encrypted for A. In the latter case service B would not be able to decrypt and interpret the assertion.

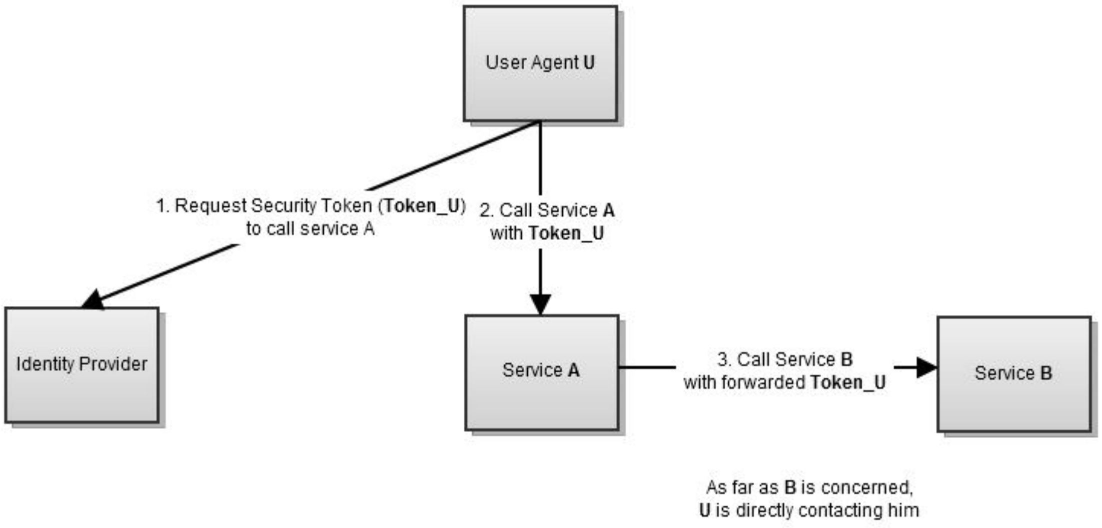

Figure 15 Credential Forwarding

Through delegation an end user U can allow a service A to access a service B on his/her behalf in a limited context (i.e. limited in time, limited in action, etc.). To support this a service A can request a delegation assertion from the IdP. This delegation assertion will then state the identity of the current client (service A) and the identity of the user on whose behalf the client is acting (end user U).

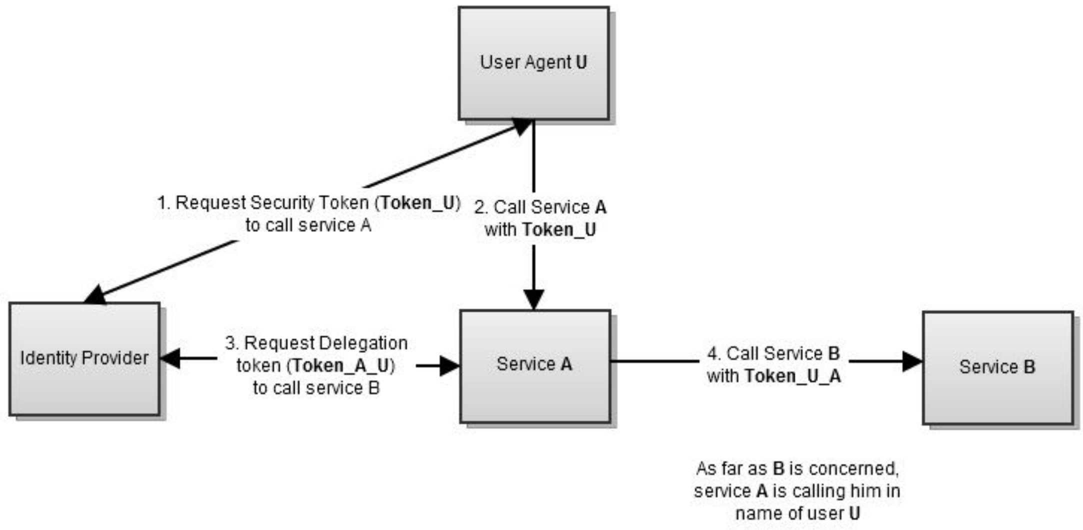
User Agent U
1. Request Security Token (Token_U) to call service A
2. Call Service A with Token_U
Identity Provider
3. Request Delegation token (Token_A_U) to call service B
Service A
4. Call Service B with Token_U_A
Service B
As far as B is concerned, service A is calling him in name of user U

Figure 16: Credential Delegation

7.4.2 De-Identification/Pseudonymisation

7.4.2.1 Data Import Requirements

As described in the legal requirements (see chapter 4.2.1.3 above) patient information should be anonymised when imported in p-medicine. As anonymisation (i.e. by using k-anonymity) removes or generalises the information, it is not useful for genetic data. Therefore the patient data will not be fully anonymised but rather be de-facto anonymised (5.2.2.1) by removing all direct identifiers.

7.4.2.2 PIMS/CATS

CAT (Custodix Anonymisation Tool) and its service-oriented evolution CATS (Custodix Anonymisation Tool Services) are responsible for the transformation of input files (plain text, CSV, XML ...) to output files. Based on a predefined set of transformation rules, called a privacy profile, CATS will process an input file and deliver it to the next component in chain (e.g. a database on a research platform). CATS supports multiple privacy profiles. Before processing a file, it will select the correct privacy profile based on the detected file type, and given content.

Important transformation are:

- Pseudonymisation: Based on person identifying information a pseudonym is added. CATS can easily be integrated with PIMS.
- De-identification: Person identifying information is cleared from the input.
- Encryption: Sensitive data can be encrypted with configurable public key.
- String replacement: Based on regular expression, string values can be replaced.

PIMS is a personal information system. Feeding it with personal identifying information coming from different sources, allows PIMS to issue pseudonyms to different domains. A pseudonym is a cryptographically strong identifier, created randomly as a GUID. A given person can be uniquely identified within an administrative domain by assigning domain specific pseudonyms, which are completely isolated. This means pseudonyms aren't derivable from one another. The re-identification module allows a user to translate an issued pseudonym back to its original personal identifying information. This is accomplished by keeping all information in a secured database.

PIMS is able to link person records that slightly defer but actually match the same person with each other. This process is called indexation. It matches records, based on personal identifying information, and adds them to an index kept in an index tree, the Master Patient Index (MPI). By indexing records the same domain specific pseudonym can be issued for multiple matching person records.

Background processes will scan the secured database for newly added person records and add them to the MPI. PIMS incorporates a probabilistic matching engine based on the well-known Fellegi-Sunter algorithm. Using fuzzy matching techniques (Jaro-Winkler) and the calculation of relative occurrences on record fields, a weight is assigned to comparison of two records. Based on that weight a match/non-match decision is made. The matching engine is fully configurable to reduce the number of false positives and false negatives to a minimum.

7.4.2.3 Initial De-Identification Architecture

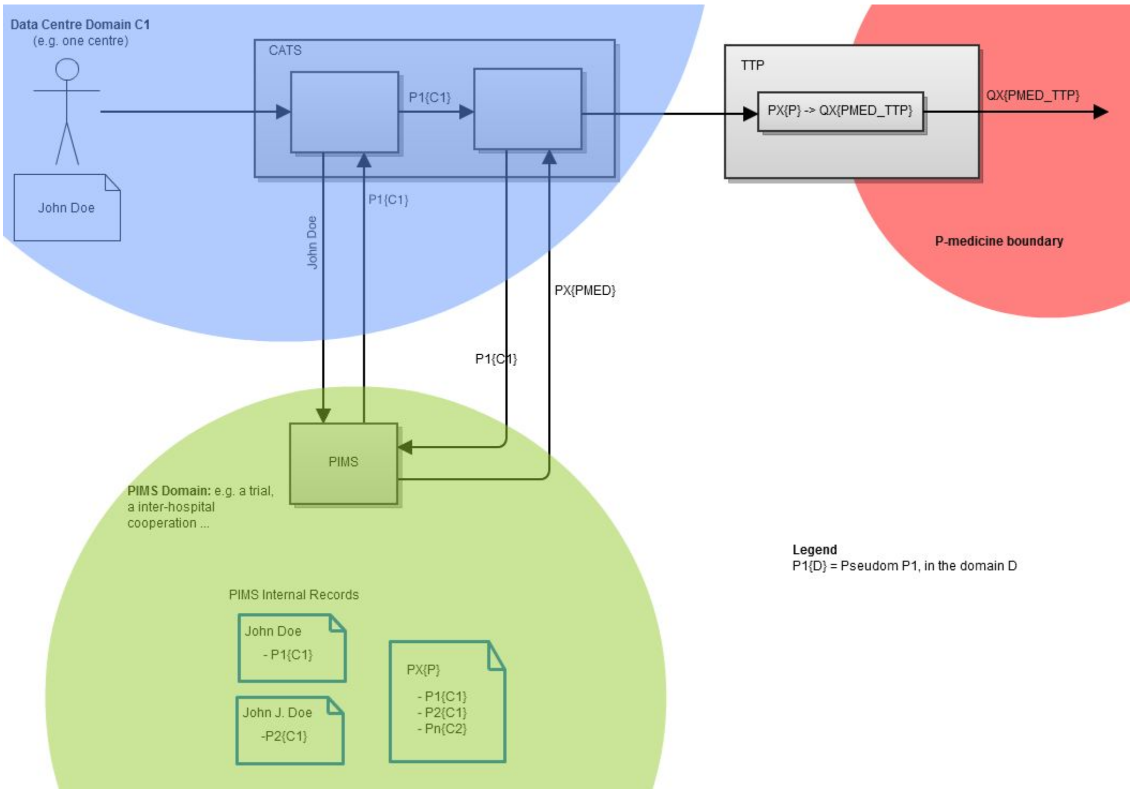

Figure 17: De-Identification Architecture

Data centres upload their data through CAT or CATS into p-medicine. CAT, which is an application or service (CATS) that runs locally within the data centre domain (domain C1), will remove all identifying data (attributes) of the to be uploaded patient information and replace them by pseudonyms retrieved from PIMS. To retrieve a pseudonym the identifying attributes are sent to PIMS. Therefore, as PIMS contains identifying information, it will typically be located within the domain of a trial or inter-hospital cooperation. CAT(S) will then use this pseudonym in the data centre's domain P1{C1} to request from PIMS a pseudonym for the p-medicine domain PX{PMED}. CAT(S) will finally upload the de-identified data (pseudonyms and non identifying attributes) to a Trusted Third Party (TTP) who's responsibility it is to transform the pseudonym which PIMS issued PX{PMED} into the actual pseudonym QX{PMED_TTP} which is deliver to the p-medicine platform. This extra step is required to ensure it is not possible to link the de facto anonymised data with the patient without going through the TTP.

7.4.2.4 Further Research

7.4.2.4.1 Encryption of Identifying Attribute before upload to PIMS

As explained in the previous section when a pseudonym is requested, identifying information is sent to PIMS. This identifying information is also stored on PIMS to allow for re-identification and linking of identities over different pseudonymisation requests. Ideally though identifying information should never leave the data centre's domain. Therefore the CAT client can encrypt individual identifying attributes (name, data of birth, address, etc.) before sending them to PIMS.

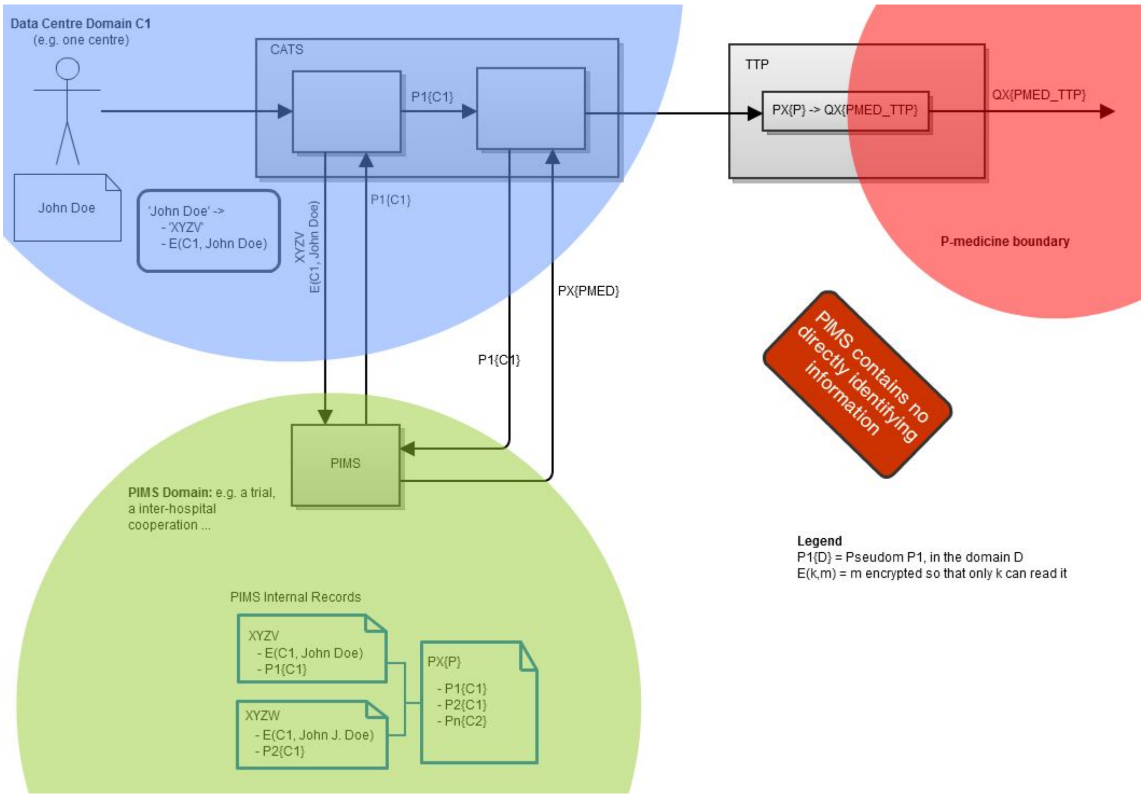

Figure 18: De-Identification architecture with encryption of identifying attributes

However, due to the nature of cryptographic algorithms, very similar attributes (e.g. typos) will be transformed to different encrypted values. Encryption doesn't sustain the similarity between records. For this reason, fault-tolerant matching, implemented by the matching engine in PIMS, on individual attributes therefore is not possible anymore. It is important to keep in mind that there is still a risk of re-identification when using encrypted attributes. Through statistical or frequency analysis techniques, re-identification of (parts of) encrypted attributes can still be achieved.

There are two main strategies in tackling the problem of matching encrypted record.

- The first strategy includes algorithms to match encrypted words.
- The second strategy is to match the records at client side, so called distributed probabilistic matching.

7.4.2.4.1.1 Matching Encrypted Data using Q-gram

A Q-gram is a word with length Q that is a substring of a given word. Q-grams are used in fault-tolerant matching of words: one can assume that for the most common errors, the number of Q-grams that match with two different words will be big, if the first word is a variant of the second. Usage of the Dice coefficient can determine the matching weight.

In this algorithm encrypted subsets of the collection of Q-grams[189] for a given word are sent to the matching service. The subsets are encrypted using a secret key shared by all sources. Based on the encrypted data the matching service can calculate a weight for matching attributes of records. Further the already existing matching algorithms can be used to take a match/non-match decision.

The biggest disadvantage of this algorithm is the massive amount of subsets that are generated and encrypted at the source. Each encrypted subset must be send to the matching service that will match with all previously received subsets. This implies an enormous consumption of resources (network, storage, memory, etc.).

7.4.2.4.1.2 Matching Encrypted Data using Bloom Filters

A bloom filter[190] is a compact data structure, which allows to check whether a given element is part of a collection, without the need to store the complete collection. Due to the compact representation false positives are possible. A bloom filter could indicate that an element is part of a collection but in reality it is not. On the other hand, false negatives can never occur. Bloom filters can be applied to investigate whether a given substring is part of a given word, without showing the complete word.

To overcome the disadvantages of the Q-gram technique, one can use bloom filters to match encrypted data. The algorithm itself is similar, but the collection of Q-grams for a given attribute is encrypted in a bloom filter instead of a collection of subsets.

7.4.2.4.1.3 Distributed Probabilistic Matching

Another approach is to allow source to execute a part of the record matching. Instead of providing the matching service with encrypted data, the source could match attributes to a list of reference values. The source will calculate a distance vector for a record against the reference values and supplies it to the matching service. The service will use these distance vectors in its algorithms to make a match/non-match decision.

7.4.2.4.2 Reduction of false positives and negatives

All studies on the subject of record matching aim for one goal: reduce the number of false negatives and positives.

The current implementation of PIMS makes a match/non-match decision based on a configurable threshold. If the calculated matching weight is below the threshold, two records are considered non-matching, otherwise the records are considers matching. To reduce the amount of false positives, PIMS could introduce a second threshold. Given two thresholds Ta and Tb, Ta < Tb. If the calculated weight is below Ta two records are considered non-matching, above Tb considered matching and between the two thresholds, considered possibly matching. For these records human intervention is necessary to make a final match/non-match decision. Two strategies are possible for human intervention. The pessimistic strategy requires human intervention when a pseudonym is requested for a subject in the grey area. The optimistic strategy issues a temporary pseudonym. Later on this temporary pseudonym can be invalidated if deemed necessary. All receivers of the pseudonym would have to be notified of its invalidity.

[189] Churches/Christen, Blind Data Linkage Using n-gram Similarity Comparisons (2004).

[190] Schnell/Bachteler/Reiher, Privacy-preserving record linkage using Bloom filters (2009).

8 Conclusion

Within p-medicine a research infrastructure for scientific medical research on patient health data is elaborated. As personal health data, in particular genetic data, contain very sensitive information concerning the patient as well his/her relatives, special emphasis has to be laid on the protection of patient´s privacy rights. Therefore a strong data protection and data security framework has to be set up for p-medicine, which has to be in line with the current legal requirements. The fulfilment of these conditions is a vital factor for the compliance of p-medicine with current data protection laws. Besides full compliance with data protection regulations generates trust and increases the acceptance by patients and thus is an indispensable precondition for the success of p-medicine.

The use of personal data is regulated by the Data Protection Directive 95/46/EC. The Directive sets out the rights of the data subject, establishes general rules on the lawfulness of the processing of personal data and states inter alia conditions for the transfer of personal data to third countries. Directive 95/46/EC, thus, introduces the rules applicable to every processing of personal data throughout the EU. The analysis showed that Directive 95/46/EC recognises the special sensitivity of certain types of data, including information on the health of a person. Accordingly the Directive generally prohibits the processing of such sensitive data. However, it introduces several exemptions to the general prohibition in Art. 8. For the processing of patient data within p-medicine two exemptions may be taken into account, which is the informed consent of the patient (see chapter 4.2.1.3.2) on the one hand or a national transposition of the exemption stated in Art. 8 para. 4 in case of “substantial public interest” (see chapter 4.2.1.3.2 above), which medical research is assumed to be.

The informed consent has to meet specific requirements in order to be valid. The patient has to be provided with all information necessary for his/her decision. The information has to be comprehensive and understandable and should at least include the main intentions of p-medicine and the range of possible uses of data, measures taken to protect patients’ personal rights, the possible risks and benefits, and further implications of participation. Further the consent has to be given voluntarily the patient has to be capable to take decisions.

However, in the context of the building up of a long term research database, in particular the information of the patient regarding the envisaged research projects may cause problems, e.g. when the data shall be used for a variety of research purposes in the future. In these cases the purposes of future research projects often cannot be specified in the respective consent forms. Hence it was decided that the setting up of such a data protection framework should not primarily be based on the patients´ consent.

An even more data protection friendly approach is to restrict the processing of health data in p-medicine to anonymised data, as this is the best way to ensure the respect of the patients´ privacy rights. Therefore it is recommended to base the data protection framework for p-medicine on anonymous data only, as such data do not fall under the scope of the data protection regime and can be used without the restrictions of Directive 95/46/EC. In addition the Directive states itself that personal data shall be anonymised before they are used for the purposes of scientific research, whenever this does not impede the purposes of research project.

However complete anonymisation is not an option for p-medicine as there shall be a possibility to contact the patient in the event that research findings could influence the treatment of a patient. In order to provide a solution ensuring the highest level of data protection while providing for the possibility to contact the patient in the event that the research leads to findings relevant for the patient we propose to base the framework on the use of de facto anonymous data. The analysis showed that the term anonymous data is not limited to completely anonymous data. Also pseudonymous data can be regarded as anonymous data within the meaning of the Directive 95/46/EC, when it is provided that the controller of the data has no access to the link back to the patient and no possibility is available within reasonable means to procure such link. Furthermore it has to be guaranteed that no other person that is allowed to process the data may re-identify the patient. All this shall be safeguarded by the data protection and data security framework set up for p-medicine. It shall ensure that the data transferred to the p-medicine infrastructure is only accessible for researchers that are contractually bound to the data protection rules set up for p-medicine (network of trust).

Technically this framework shall be designed as a lightweight dynamic architecture based on REST technology (cf. p-medicine technological requirements) with a focus on usability, ease of integration and the use of industry standards such as the Security Assertion Markup Language (SAML) and the eXtensible Access Control Markup Language (XACML). From a legal perspective, as outlined in chapter 7.3, the setting up of a network of trust requires the establishment of a central data protection authority within p-medicine, the introduction of a Trusted Third Party (TTP) and the conclusion of contracts between the participating hospitals (data exporters) and researchers (end users) with the authority. Accordingly the setting up of the “network of trust” mainly needs three components.

The task of the data protection authority shall be fulfilled by the Center for Data Protection (CDP), an association under Belgian law that had already successfully served as data protection authority within ACGT. The CDP shall however only serve as controller for data stored within the p-medicine data warehouse (including) the transfer to the end users, whereas the p-medicine end users shall be responsible for the data processing within their own entities. Thus, the end users of p-medicine are to be considered data controllers with respect to the data they research on and that they accessed via the p-medicine infrastructure. Accordingly the end users have to be compliant with data protection legislation applicable for their entity. Nevertheless the CDP will serve as a central contact point for the patients regarding all data processed within the p-medicine infrastructure, including data processed under the control of the end users.

Second, a Trusted Third Party is needed, which is responsible for the pseudonymisation of patients’ data. The TTP will also act as a trustful custodian for the pseudonymisation key to re-identify the patient concerned. The requirements for a TTP have been examined in chapter 7.3.2. Custodix that shall serve as TTP within the p-medicine framework meets all these requirements. Custodix is an expert in data security.

Finally contracts between all participating hospitals, researchers or other end users of p-medicine and the CDP must be concluded in order to ensure confidentiality, data security and compliance with data protection legislation. As pointed out in chapter 7.3.7 mainly three contracts are required to set up the framework. The first regards the transfer of patient data to the p-medicine infrastructure (Data Transfer Agreement). This agreement is to be concluded between the CDP and the healthcare organisation/hospital delivering patient data (data exporter). The second agreement concerns the data processing within the p-medicine framework (Contract on data protection and data security within p-medicine). This agreement will have to be concluded between the CDP and all end-users of p-medicine doing research on these data. This deliverable introduces the essential provisions of these two contracts such as the question of data control within p-medicine, obligations concerning the network of trust, third beneficiary rights and the applicable law. The different topics are discussed and it is explained for what reasons it has been decided to design the different provisions as they are. The contracts are included in the Annex to this document.

Furthermore in the course of p-medicine a Patient Identity Management System will be developed ensuring that the same patient gets the same pseudonym when his/her data has already been transferred to the data warehouse. This is done by verifying whether a specific patient has already been registered in the common PIMS-database (see chapter 7.3.4). Thus

PIMS avoids the creation of different pseudonyms for the same patient (synonyms) as well as the creation of the same pseudonyms for different patients (homonyms).

Anyhow for the unlikely event that the network of trust will fail and personal data are processed within p-medicine, we will need a legal basis. For this event as fall back and for ethical reasons anyway patients will be asked to consent to the data processing of their data within p-medicine. And last but not least if an informed consent should be invalid for an unforeseeable reason, the national exemption for data processing in medical research would serve as legal basis.

Finally, as the analysis in chapter 6 showed, this concept is not only in line with the current data protection legislation but also reflects and combines the strengths of data protection approaches developed in comparable research projects in the field of medical research.

Appendix 1 - Abbreviations and acronyms

CA	Certificate Authority
CAT	Custodix Anonymisation Tool
CATS	Custodix Anonymisation Tool Services
CDP	Center for Data Protection
CRL	Certificate Revocation List
CSV	Comma-separated values
DICOM	Digital Imaging and Communications in Medicine
EU	European Union
EECs	End entity certificates
GAS	Gridge Authorisation Service
GSI	Grid Security Infrastructure
GT4	Globus Toolkit 4
GUID	Globally unique identifier
IdP	Identity Provider
LDAP	Lightweight Directory Access Protocol
MPI	Master Patient Index
OASIS	The Organization for the Advancement of Structured Information Standards
OCSP	Online Certificate Status Protocol
PAP	Policy Administration Point
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PIMS	Patient Identity Management Service
PIP	Policy Information Point
PKI	Public Key Infrastructure
SAML	Security Assertion Markup Language
SP	Service provider

SOA	Service Oriented Architecture
SSO	Single Sign-on
TTP	Trusted Third Party
XACML	eXtensible Access Control Markup Language
XML	Extensible Markup Language

Annex 1: Data Transfer Agreement

(Version 1.0, January 2012)

between

P-MEDICINE Center for Data Protection

____

(address and country of establishment)

hereinafter "CDP"

and

____

____

(address and country of establishment)

hereinafter "data exporter"

each a "party", together "the parties".

Preamble

The project P-MEDICINE (**From data sharing and integration via VPH models to personalized medicine**) is a European financed project supported by partners from eleven European countries and Japan, coming from different backgrounds, including physicians, clinicians, genomic scientists, medical- and bio-informaticians, and legal and ethical experts. P-MEDICINE brings together international leaders in their fields to create an infrastructure that will facilitate the translation from current medical practice to personalised medicine. In achieving this objective P-MEDICINE has formulated a coherent, integrated workplan for the design, development, integration and validation of technologically challenging areas of today.

The emphasis of P-MEDICINE is on formulating an open, modular framework of tools and services, so that it can be adopted gradually, including efficient secure sharing and handling of large personalised data sets, building standards-compliant tools and models for research, and providing tools for large-scale, privacy-preserving data and literature mining. P-MEDICINE will ensure that privacy, non-discrimination, and access policies are aligned to maximize protection of and benefit to patients. The P-MEDICINE tools and technologies will be validated within the concrete setting of advanced clinical research. Pilot cancer trials have been selected based on clear research objectives, emphasizing the need to integrate multilevel datasets, in the domains of Wilms tumor, breast cancer and leukemia. To sustain a self-supporting infrastructure realistic use cases will be built that will demonstrate tangible results for clinicians. The project is clinically driven and promotes the principle of open source and open standards.

Sharing clinical and genomic expertise implies the transfer and exchange of patient data within the project. Therefore the infrastructure of P-MEDICINE is embedded in the P-MEDICINE Data Protection Framework, which guarantees compliance with current European data protection legislation, primarily by de facto anonymising the patient data. Due to the diverse participants it is of high importance to process patient data in accordance with the P-MEDICINE General Terms on data protection (ANNEX A) and keep the data exchange within the project under the control of P-MEDICINE. This is guaranteed by the implementation of the P-MEDICINE Center for Data Protection (CDP) as central data controller within P-MEDICINE. The P-MEDICINE end-users (scientific researchers) will process data received from the P-MEDICINE data warehouse as data controllers themselves.

Technically each data exporter, e.g. healthcare organisation/hospital, will transfer their patient data to the “P-MEDICINE data warehouse”, where the data will be kept in de facto anonymised form.

Clause 1: Definitions

For the purposes of the clauses:

1. **Personal data**, **process/processing**, **data controller/controller**, **processor**, **data subject, technical and organisational security measures** and **supervisory authority/authority** shall have the same meaning as in Directive 95/46/EC of 24 October 1995 (whereby **supervisory authority/authority** shall mean the competent data protection authority in the territory in which the data controller is established);
2. **Patient** shall mean a person treated in a healthcare organisation/ hospital, whose mainly health related (e.g. genetic) data are processed within the P-MEDICINE network. The patient therefore is a data subject, where his/her personal data are processed;
3. **Data exporter** shall mean the data controller who transfers the data to the P-MEDICINE data warehouse under control of the CDP. In most of the cases the data exporter will be a healthcare organisation/hospital, which is the intermediary between the patients willing to participate in P-MEDICINE and the P-MEDICINE project itself;
4. **P-MEDICINE Consortium** shall mean the group of partners who signed the P-MEDICINE Consortial Agreement;
5. **P-MEDICINE data warehouse** shall mean the database/databases where the patient data used within P-MEDICINE are stored (also referred to as P-MEDICINE database);
6. **Center for Data Protection (CDP)** shall mean the central data controller within P-MEDICINE, which agrees to receive from the healthcare organisations/hospitals data intended for processing in accordance with the P-MEDICINE General Terms (Annex A) and the terms of this contract. The CDP also serves as a central data protection authority within P-MEDICINE ensuring the compliance of all P-MEDICINE participants with the Data Protection Framework, particularly with regard to P-MEDICINE’s policies and procedures;
7. **P-MEDICINE end user** shall mean the entity or person, e.g. healthcare organisation/hospital or investigator, conducting scientific research and participating in P-MEDICINE after having signed the “Contract on data protection and data security within P-MEDICINE” (P-MEDICINE end user agreement).
8. **De-facto anonymous data** shall mean data that has been modified in such a way that the information concerning personal or material circumstances can be attributed to an identified or identifiable individual only with a disproportionate amount of time, expense and labour.

9. **Clauses** shall mean these contractual clauses.

Clause 2: Scope and responsibility

(1) The scope of this contract is solely the data transfer from the data exporter to the P-MEDICINE data warehouse.

(2) On the basis of this scope, the data exporter is responsible for and has control over data processing and storage performed in its organisation including the data transfer to the P-MEDICINE data warehouse. As soon as the data are transferred to the P-MEDICINE data warehouse their processing is under the sole responsibility and control of the CDP.

Clause 3: Obligations of the data exporter

The data exporter warrants and undertakes that:

1. all data transferred to the P-MEDICINE data warehouse have been collected, processed and transferred in accordance with the laws applicable to the data exporter.
2. it shall transfer patient data to the P-MEDICINE data warehouse only if the patients concerned have signed all corresponding P-MEDICINE consent forms listed in Annex B after having been sufficiently informed by the data exporter. The data exporter shall provide the CDP with one signed copy of all consent forms to be signed by patients, minor patients, their representatives, hospitals or investigators and physicians;
3. it shall assure the quality of pseudonymisation by implementing the Custodix Anonymisation Tool (CATS) or any other state of the art pseudonymisation tool recommended or accepted by the CDP according to clause 4.2.
4. it shall transfer data to the P-MEDICINE data warehouse only after initiating the pseudonymisation procedure selected according to nr. 3 of this clause, so that no personal data will enter the P-MEDICINE network.
5. it shall not run any matching procedures between data transferred to the P-MEDICINE data warehouse and data sources existing within its organisation and it will use no other means in order to identify the patient concerned.
6. it shall provide the CDP upon request with copies of relevant data protection laws or references to them (where relevant, and not including legal advice) of the country in which the data exporter is established.
7. it has no reason to believe, at the time of entering into these clauses, in the existence of any local laws that would have a substantial adverse effect on the guarantees provided for under these clauses, and it will inform the CDP (which will pass such notification on to the authority where required) if it becomes aware of any such laws.
8. it shall forward enquiries from patients and the authority concerning the processing of data within P-MEDICINE to the CDP. The data exporter will identify to the CDP a contact point within its organisation authorised to respond to enquiries concerning processing of the data, and will cooperate in good faith with the CDP, the data subject and the authority concerning all such enquiries within a reasonable time. The CDP will respond to that enquiry to the extent reasonably possible and within reasonable time directly to the data subject/patient or via the data exporter depending on the choice of the data subject/patient.
9. it shall make available, upon request, a copy of the clauses to patients who are third party beneficiaries. In case these clauses contain confidential information the data exporter may remove such information. Where information is removed, the data exporter shall inform the patient in writing about the reason for removal.
10. it shall support the CDP and provide the CDP with all necessary information and documents needed in case of prior checking by supervisory authorities.
11. it shall provide the CDP with contact details of the person responsible for data protection within its organisation.

Clause 4: Obligations of the CDP

The CDP warrants and undertakes that:

1. it shall have in place appropriate technical and organisational measures to protect patient data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, and which provide a level of security appropriate to the risk represented by the processing and the nature of the data to be protected.
2. it shall accept any pseudonymisation tool chosen by the data exporter, if it has confirmed that it ensures a state of the art pseudonymisation as provided for in clause 3.3.
3. it shall have in place procedures to ensure that any third party authorised to access the data, in particular each P-MEDICINE end-user, respects and maintains the confidentiality and security of the data.
4. it shall process patient data for research purposes only, and has the legal authority to give the warranties and fulfil the undertakings set out in these clauses.
5. it shall respond to enquiries from patients and the authority concerning the processing of the personal data within P-MEDICINE. Responses will be made within a reasonable time.
6. it shall process the data in accordance with:
    a. the data protection laws of Belgium,
    b. the provisions of Directive 95/46/EC, and
    c. the data processing principles set forth in this agreement.
7. it shall not disclose or transfer personal data to a third party data controller located outside the European Economic Area (EEA) unless it is empowered by the P-MEDICINE Consortium to do so and
    a. the third party data controller processes the personal data in accordance with the law of a third country that is recognized in a Commission Decision as providing adequate data protection, or
    b. the third party data controller becomes a signatory of the P-MEDICINE end-user contract and fulfils the conditions stated in Art. 25, 26 of Directive 95/46/EC.

Clause 5: PIMS

(1) The data exporter may use a common patient identity management system (PIMS) provided by P-MEDICINE together with other healthcare organisations/hospitals participating in P-MEDICINE. The use of this tool falls under the sole responsibility of the data exporter.

(2) The hospitals may use the pseudonym generated by the PIMS solely for the transfer of the data to the P-MEDICINE data warehouse. This pseudonym has to be stored safely and must not be used for the internal use of the data exporter (as identifier for the patient).

Clause 6: Confidentiality

(1) The data exporter and the persons employed in data processing of the data exporter and persons employed in the technical maintenance of the P-MEDICINE database must not disclose data or any other information stored in the P-MEDICINE data warehouse to any person not underlying the confidentiality agreement stated in para. 2 or the confidentiality agreement set down in Clause 6 of the P-MEDICINE end user agreement.

(2) On taking up their duties the persons employed in data processing are required to give an undertaking to maintain such confidentiality in the event that they have access to, or any other form of contact with, the P-MEDICINE data warehouse. This undertaking shall continue to be valid after termination of the employed persons activity and the termination of this agreement.

Clause 7: Liability and third party rights

(1) Each party shall be liable to the other parties for damages it causes by any breach of these clauses. Liability as between the parties is limited to actual damage suffered.

(2) The parties agree that if one party is held liable for a violation of the clauses committed by the other party, the latter will, to the extent to which it is liable, indemnify the first party for any cost, charge, damages, expenses or loss it has incurred, provided that:

1. the parties promptly notify each other of a relevant liability claim; and
2. each party is given the possibility to cooperate in the defence and settlement of the claim.

(3) The parties agree that a patient shall have the right to enforce as a third party beneficiary this clause and clauses 3.5, 3.9, 3.11, 4.1, 4.3, 4.4, 4.5, 4.6, 4.7 and clause 6 against the CDP or the data exporter, for their respective breach of their contractual obligations, with regard to his/her data.

(4) The parties accept the exclusive jurisdiction of the competent courts in Hanover, Germany and the application of the laws of Germany for determining this.

Clause 8: Resolution of disputes, mediation and arbitration

(1) The parties agree with regard to disputes between each other in connection with these clauses:

1. that, subject to further agreement with each other, such disputes can be referred to mediation by an independent person or, where applicable, by the supervisory authority;
2. that, subject to further agreement with each other, the resolution of a specific dispute can be referred to an arbitration body if the data exporter is established in a country which has ratified the New York Convention on enforcement of arbitration awards.

(2) The parties agree with regard to patients' rights:

1. that if the patient invokes third-party beneficiary rights and/or claims compensation for damages under these clauses, the parties will accept the decision of the patient:
   (a) to refer the dispute to mediation, by an independent person or, where applicable, by the supervisory authority; or
   (b) to refer the dispute to the courts of Hanover, Germany.
2. that, subject to agreement by the patient, the resolution of a specific dispute can be referred to an arbitration body if the data exporter is established in a country which has ratified the New York Convention on enforcement of arbitration awards.
3. to inform each other about any disputes or claims brought by a patient concerning the processing of his/her data. The parties will cooperate with a view to settling any such claim amicably in a timely fashion.

Clause 9: Penalty

(1) The parties agree to pay to the other party a penalty of ____________EUR for any negligent or intentional breach of clause_______ caused by itself, its employees or any of its subcontractors.

(2) This shall be without prejudice to the parties' right to terminate the contract, to seek compensation for damages or to enforce any claims under this contract.

Clause 10: Term of the agreement, termination and obligations of the parties after the termination

(1) This agreement shall come into force upon signature of both parties and remain effective until_______.

(2) In case of violation of clauses ___________ by one of the parties to this agreement the other party is entitled to terminate this contract immediately.

(3) Without prejudice to the foregoing provisions, any party may terminate this contract for good cause.

(4) Each party must inform the other party in written form in case of termination of the contract.

(5) The parties agree that on the termination of the provision of data processing services, the data exporter shall, at the request of the CDP, provide all signed informed consent forms of its participating patients to the CDP.

(6) The data exporter must inform the concerned patients of the termination of the data processing services, of any handover of the consent forms to the CDP, and of the right of each of them to withdraw consent to further processing of their data.

Clause 11: Governing law and Jurisdiction, miscellaneous

(1) This contract shall be governed by German Law. The courts of Hanover, Germany shall have exclusive jurisdiction. This shall also apply to disputes on the validity of this clause.

(2) Changes and amendments to this contract and all of its components, including any assurances by the CDP, require written agreement and an explicit statement that they represent a change or amendment to these conditions. The same applies to the waiving of this formal requirement.

(3) If any provision of this contract shall be entirely or partly invalid or unenforceable, this shall not affect the validity and enforceability of all other provisions of this contract. An invalid or unenforceable provision shall be regarded as replaced by such a valid and enforceable provision that as closely as possible reflects the privacy, security and/or economical purpose that the parties hereto had purposed with the invalid or unenforceable provision.

Made in……….signed copies, each party having received its own signed copy.

Dated: ____________________________________

____

FOR the CDP

____

FOR the DATA EXPORTER

Annex A: General Terms

GENERAL TERMS

(Version 1.0, January 2012)

Preamble:

The project P-MEDICINE (**From data sharing and integration via VPH models to personalized medicine**) in this present document, aims at creating clinico-genomic databases on cancer. The P-MEDICINE project will start up by collecting data on breast cancer (BRCA), Acute Lymphoblastic Leukemia (ALL) and nephroblastoma (PN), but it is projected to involve further cancer types in the future. The final purpose of such scientific research is to improve cure and management of future cancer patients by putting together the results of several researches running in Europe.

Therefore a data warehouse will be set up within P-MEDICINE to enable the project's participants to exchange patient data. This P-MEDICINE data warehouse will contain patient data transferred by participating hospitals/investigators, upon the patients´ informed consent to use their data within P-MEDICINE. The data warehouse can consists of different databases. The databases will not host anything else than data (thus excluding, for example, biomaterial).

All data transferred to the P-MEDICINE data warehouse will be pseudonymised through dedicated state of the art pseudonymisation software. In order to be able to re-identify a given patient, for example in the event that a new treatment for him/her is developed, the name of the patient is replaced by a pseudonym during this procedure. The pseudonymisation key needed to link the pseudonymised data set to the patient concerned will be kept only by a Trusted Third Party (TTP). The TTP's independence from hospitals/investigators will be guaranteed. That means that the user (researcher) using the data will be unable to identify the patient to whom the data relates. In addition contracts are concluded between the participating hospitals/investigators and P-MEDICINE guaranteeing that patient data are not transferred to any party outside the project and no matching of data set takes place in order to identify the patients concerned. In interaction with strong technical and organisational security measures patient data in P-MEDICINE is to be seen as de-facto anonymous. Such data can only be de-anonymised by the TTP and with permission of the P-MEDICINE Center for Data Protection (CDP) if the de-anonymisation is needed in the interest of the patient concerned.

The data will be stored for a length no longer than the P-MEDICINE project. During the whole term of storage it will always be provided that the data remain de-facto anonymous for the end users. For a longer storage of patient data the explicit informed consent of the patient will be required.

The patient data remain under the control of the respective hospital/investigator (data exporter) where the data are collected until the data have been transferred to the P-MEDICINE data warehouse. From that point on the CDP will be responsible for the data processing within the P-MEDICINE data warehouse. The CDP concludes contracts with the P-MEDICINE end users that guarantee the protection and security of the data received from the data warehouse for the purposes of scientific research. Further the CDP controls the compliance to these contractual agreements. It, thus, serves as a central data protection authority for the P-MEDICINE framework. The P-MEDICINE end users process the data received from P-MEDICINE on their own behalf, so that they are to be regarded as data controllers according to the law.

The users (researchers) are not allowed to publish the data or to transmit or disclose data received via P-MEDICINE to any third person outside of P-MEDICINE.

These General Terms will apply to the CDP (as a legal person), the patients, physicians and end users.

Explanatory Glossary:

Anonymous data / Rendering anonymous

Rendering data anonymous means to modify personal data in a way that the information concerning personal or material circumstances can no longer be identified or identification is only possible with a disproportionate amount of time, expense and labour be attributed to an identified or identifiable individual. Data that have been anonymised are no longer "personal data" in the legal sense. It will be an aim to have as much anonymised data within P-MEDICINE as possible and reasonable.

Center for Data Protection (CDP)

The CDP shall mean the central data protection authority of the P-MEDICINE infrastructure, which agrees to receive from the healthcare organisations/hospitals (data exporters) data intended for processing in accordance with the terms of the data exporter agreement. The CDP guarantees privacy within the P-MEDICINE data warehouse.

Confidentiality

Persons employed in data processing shall not collect, process or use personal data without authorisation (confidentiality). On taking up their duties such persons shall be required to give an undertaking to maintain such confidentiality. This undertaking shall continue to be valid after termination of their activity. Any person acting under the authority of the CDP who has access to P-MEDICINE patient data must not process them except on instructions from the controller, unless he/she is required to do so by law.

Consent

Informed consent means any express indication of patient´s wish expressing his/her agreement to data relating to him/her being processed, provided that he/she has sufficient information about the purposes of the processing, the data or categories of data concerned, the recipient of the data, and the name and address of the controller and of his/her legal representative if any. The patient's consent must be freely given and specific, and may be withdrawn by the patient at any time. If the patient is incapable of a free decision or domestic laws do not permit the patient to act on his/her own behalf, consent is required of the person recognised as legally entitled to act in the interest of the patient or of an authority or any person or body provided for by law (legal representative).

Data controller

The data controller/controller is, according to the Data Protection Directive 95/46/EC, the natural or legal person who alone, or jointly with others, determines the purposes and means of the processing of personal data. The data controller is liable for the legality of the processing and the fulfilment of the obligations towards the national data protection authority and the patients. The hospitals/investigators (data exporters) are data controllers with regard to the collection of patient data and theirs transmission to P-MEDICINE, whereas CDP is the data controller with regard to the data stored in the P-MEDICINE data warehouse. Finally the

P-MEDICINE end users are data controllers regarding the data received from the P-MEDICINE data warehouse.

Data processor

Data processor shall mean a natural or legal person, public authority, agency or any other body which processes patient data on behalf of the controller, such controller being liable for the legality of the processing and the fulfilment of the obligations towards the national data protection authority and the patients.

Data subject

The data subject is the subject of personal data, meaning an identified or identifiable person the data refers to. An identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity. As a rule the patient, whose data are collected and processed for P-MEDICINE will be the data subject, when his/her personal data are processed.

Disclosing

Disclosure is a processing operation in which patient data are provided by a controller to a third party. The data controller must only disclose data to third parties if permitted by law or by the data subject´s consent. In P-MEDICINE data are only disclosed to P-MEDICINE end users who have signed a special agreement that forbids any disclosure of data received via P-MEDICINE to any other third party.

Hospital

Hospitals are health institutions where patients are treated and their personal data are collected for the purpose of the P-MEDICINE project.

Investigator

The legal or natural person who gathers and manages the patient's data from the hospitals, laboratories etc. and maintains and controls the trial/study database.

Legal representative of the patient ("legal representative"):

The legal representative(s) of the patient is/are the person(s) who has/have the power by law or legal decision to decide for a minor patient (or equivalent status such as mentally disabled patients).

Necessary processing

When deciding which data will be collected and further processed, the controller must limit these data to the extent necessary to achieve the purpose of processing. This means that personal data will only be processed when it is necessary for the project.

Patient:

Patient means the person treated in a hospital. Certain data collected in the hospitals will upon the patient´s consent be transferred to the p-medicine framework where they will be used for the purposes of scientific research in de facto anonymous form.

Personal data

Personal data means any information relating to an identified or identifiable natural person ('data subject'). An identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity. Therefore a set of data collected under a certain number or sign “patient xxx”, “tissue YYY” can be personal data, if the patient concerned can still be identified by other means than his/her name.

Physician

The physician is the natural person who is in charge of the patient's treatment.

Pseudonymisation

To pseudonymise a data set means to replace the patient's name and other identifying characteristics with a coded label in order to preclude direct identification of the patient or to render such identification substantially difficult. Within P-MEDICINE only pseudonymised data are processed.

Publish

The controller and the processors will refrain from publishing personal data or otherwise making them public, unless consent from the patient concerned is obtained.

Purpose

The purposes for processing of personal data must be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed. The purposes must be specified, explicit and legitimate. Personal data must be not further processed in a way incompatible with those purposes. The purpose for the collection, transfer and use of the data within P-MEDICINE is to conduct scientific research.

Sensitive (personal data)/Special categories of data

Sensitive personal data are personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and data concerning health (genomic data) or sex life. The processing of sensitive data is only allowed in case of certain exceptions explicitly stated by the national laws of the Member State.

Storage

Storage of personal data is allowed by the Data Protection Directive 95/46/EC. But when the purpose of processing is achieved and the data are not required any more for that particular purpose, personal data must be rendered anonymous or must be destroyed. Most national laws allow personal data to be stored for a longer term, provided that this is in order to use the data exclusively to carry out scientific research or statistics. Nevertheless, some national laws impose supplementary conditions or formalities in order to allow longer storage.

Technical and organisational measures

Organisational measures, together with technical measures, must ensure an appropriate level of security of the data processing, taking into account the state of the art and the costs of their implementation relative to the risks inherent in the processing and the nature of the data to be protected. Appropriate organisational measures shall be taken by the controller against accidental loss, destruction or alteration of, or damage to, personal data and against unauthorized or unlawful processing of personal data in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing. The controller must, where processing is carried out on his/her behalf, choose a processor providing sufficient guarantees in respect of the technical security measures and organisational measures governing the processing to be carried out, and must ensure compliance with those measures. Such appropriate organisational measures to ensure the confidentiality, integrity and accuracy of processed data should include for example:

- control of the entrance to installations
- control of data media
- memory control
- control of utilization
- access control

- control of communication
- control of data introduction
- control of transport
- availability control

Such technical and organisational measures have to be taken by all the P-MEDICINE-participants processing patient data.

Third Party

A third party is a natural or legal person, public authority, agency or any other body other than the patient, the controller, the processor or persons who, under the direct authority of the controller or the processor, are authorised to process the data. With regard to P-MEDICINE, third parties will be all the other persons and bodies who have no authorisation of P-MEDICINE to process the data.

Transfer

Transfer of data means the transmission of patient data from one data controller to another.

Trusted Third Party

The Trusted Third Party is an independent security authority, which has no interest in the content of the processed data and can therefore be trusted by all participants of the P-MEDICINE project. Within P-MEDICINE the Trusted Third Party will hold the pseudonymisation key needed to link data sets to the patient they belong to. The involvement of the TTP guarantees that a data set will only be re-identified, if the patient must be identified for medical or scientific reasons and if he/she wishes to be informed.

Article 1: Patient's rights

1.1. Information

The patient or his/her legal representative must be informed, in an intelligible form and fully, before giving his/her consent.

The information given to the patient or his/her legal representative will consist in describing and explaining:

- the identity of the data controller;
- the purpose of the processing of his/her data;
- his/her rights;
- the security of the data processed;
- the categories of data concerned;
- the recipients or categories of recipients;
- such other matters as may be specified from time to time by applicable data protection laws (including professionally enforceable codes of practice);

The information will be provided by the physician, who will act here as the representative of the hospital towards the patient.

1.2. Access

The patient has, directly or through his/her legal representative, a right of access to his/her data processed by or on behalf of the data controller.

This right of access includes inter alia the right to be informed:

- in an intelligible form of the data undergoing processing and of any available information as to their source;
- of the identities of the persons who have had access to his/her data and the moment of this access (log file).

The demand of access shall be addressed to the hospital through the patient´s physician.

This access is free of charge.

1.3. Object and withdrawal of consent

The patient or his/her legal representative has the right to object to the processing and to withdraw his/her former consent without giving reasons and at any time. In case the patient or his/her representative withdraws the given consent all data, which have not been pseudonymised yet by the TTP, can't be used anymore and the data already pseudonymised by the TTP have to be completely anonymised (the TTP must erase the key used for the pseudonymisation).

The exercise of these rights is free of charge.

1.4. Right to rectify

In case of inaccuracy the patient or his/her legal representative has the right to demand the rectification of his/her data from the data controller.

The exercise of this right is free of charge.

1.5. In case of death

If the TTP is informed (by the treating physician) that a patient died, his/her data can be used for research, but the TTP must erase the pseudonymisation key.

1.6. Feedback

The patient or his/her legal representative will receive information generated as a result of the research involving his/her personal data if:

- the information is likely to be directly useful to his/her therapy/treatment and
- (where he/she has no legal representative) he/she is physically and psychologically able to receive the information.

He/she may refuse this information by letter or secured Email sent to his/her physician, who for this purpose acts as the representative of the hospital or the CDP.

Article 2: Patient's obligation to provide information

At the time of providing consent to the processing of his/her data for the purposes of P-MEDICINE, the patient or his/her legal representative shall answer to the best of his/her knowledge and belief a questionnaire given to him/her by the physician. This questionnaire is designed to provide the physician with information regarding the patient’s health, medical history, etc., which is necessary for the physician to perform his/her role in the context of the P-MEDICINE network.

Article 3: Rights of P-MEDICINE, hospitals/investigators and physicians

P-MEDICINE and the participating hospitals/investigators have the right to withdraw from the project any patient (non exhaustive list):

- who has given false information in the questionnaire mentioned in article 2;
- whose consent is, in their judgement, impaired by reason of pressure from relatives or other third persons;
- in respect of whom there are other grounds for believing further participation is not in his/her interest or in that of the project

In this case, all the data from the patient concerned will, at the discretion of P-MEDICINE, either be erased or completely anonymised in the same way as provided in 1.3.

P-MEDICINE has the right to stop the project at any time without giving reasons or any explanation. In this case, all the data within the P-MEDICINE network will be erased and no liability of any form will be incurred by P-MEDICINE.

P-MEDICINE has the right to exclude a physician, hospital/investigator or user (researcher) from the project in case of violation of these general terms, the contract or agreements with P-MEDICINE, the patient´s right to informed consent, national or international laws, or relevant codes of professional practice.

The physician, hospital/investigator or user has the right to withdraw from the program at any time without giving reasons or any explanation and without any payment of a penalty. In this case, his/her grant will be stopped and claimed back, if paid in advance.

If the physician withdraws or is required to withdraw from the project the patient, whose data are stored in the hospital, can remain in the project by choosing another physician of the hospital taking part in the project. If the hospital itself withdraws from the project, the patient’s data will remain in the project until the patient withdraws his/her consent. In that case the respective data will be completely anonymised in the same way as provided in 1.3.

Article 4: P-MEDICINE's, hospital/investigator, end user, physician and Trusted Third Party obligations

Access to the P-MEDICINE network must be granted by the CDP. The grant can only be given by the CDP.

The CDP is responsible for the security of any data processing within P-MEDICINE whereas the participating healthcare organisations/hospitals (data exporters) as well as the P-MEDICINE end users are responsible for the processing of patient data within their own organisation and are, thus, obliged to ensure the confidentiality and protection of the patient data processed. These obligations are defined by the contractual agreements concluded with the CDP.

In order to be in accordance with the Directive 95/46/EC, the hospital/investigator will implement a pseudonymisation tool guaranteeing a state of the art pseudonymisation of patient data transferred to the P-MEDICINE data warehouse. Having regard to the state of the art and the cost of their implementation, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the sensitive data to be protected.

The Trusted Third Party holds the pseudonymisation key to re-identify the patient. Re-identification is permitted only if the scientific research reveals findings that are likely to be useful for the patient´s treatment and the patient wishes to be informed.

Article 5: Applicable legislation and jurisdiction

In respect of any claim against the CDP involving an alleged breach of data protection legislation, Belgian data protection legislation shall be applicable.

Subject to the provision below in relation to third party beneficiary rights, any other dispute between the patient and the CDP, under this contract or otherwise, shall be determined by the national law of the country where the patient resides.

Third beneficiary rights (not involving an alleged breach of data protection) shall be governed by the laws applicable to the contract from which those rights derive. Any dispute or claim in relation to the same shall be within the exclusive jurisdiction of the courts specified in the relevant contract.

Annex B: Consent and information forms

1. Patient

Patient information sheet

(Version 1.0, January 2012)

Explanation of the research project P-MEDICINE

(**From data sharing and integration via VPH models to personalized medicine**)

1. Invitation

You are being invited to take part in a clinical research project involving patients with cancer. The study is called P-MEDICINE for short. The full name is: "**From data sharing and integration via VPH models to personalized medicine**". It is a European Union funded project in Framework Programme 7 (Grant agreement number: 270089).

P-MEDICINE aims to improve treatment of cancer and to promote research on cancer by building up a technical trans-European infrastructure to connect trials on cancer running in different hospitals, healthcare and research organisations all over Europe. Researchers participating in P-MEDICINE will be able to work more efficiently and successfully in developing new treatments for cancer and in personalising the treatment, so that patients suffering from cancer can be treated better. In particular, P-MEDICINE aims to identify individualised molecular profiles correlated with sensitivity or resistance to therapy. Each patient suffering from cancer has a particular molecular profile. In many cases the success of a specific treatment depends on the molecular profile of a patient. In this regard P-MEDICINE aims to develop a software tool linking a particular molecular profile to the best available treatment for someone with this profile. Your data will be used to build up this new technical framework and enable researchers to conduct scientific research by using this new infrastructure.

If you decide to take part, you will be asked to provide your data anonymously to this project. Before you make this decision, it is important for you to understand further why the research is being done and what it will involve.

This document describes the project in order to help you to make your decision. Please read the information provided carefully and discuss it with others if you wish. Feel free to ask your physician and other members of your healthcare team if anything is unclear or if you would like more information.

Please 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 at any time without any disadvantages. Also, if you decide not to volunteer for the project, this will not affect your treatment in any way.

2. Purpose of the project

It is the aim of P-MEDICINE to create a collaborative environment for the development, validation and sharing of VPH models and simulators and their transformation into decision making tools for clinical routine. Three exemplary multiscale simulation models of clinical tumour response to treatment will be developed: one for nephroblastoma, one for breast cancer and one for acute lymphoblastic leukaemia (ALL), based on the principles that have been shown to be most appropriate for the clinical trial context. These three models will constitute the simulation core of the "P-MEDICINE Oncosimulator" which will provide an integrated platform for simulating, investigating, better understanding and exploring the natural phenomenon of cancer and - after successful validation and transformation - an integrated treatment support tool. Modelling and validation will require access to large sets of data, which will be made available through a data warehouse. The resulting models of tumour response to treatment and the Oncosimulator will be made available to the users in the P-MEDICINE workbench for sharing, usage and exploitation. The tools and scientific knowledge produced by this project will be publicly available. The project itself will not necessarily develop new individual treatments, but the knowledge gained within the project should eventually improve our understanding of which patient is most likely to benefit from an individual treatment. It will also help to predict cancer prognosis and the chance of the cancer recurring or not.

P-MEDICINE is sponsored by the European Union. The P-MEDICINE research consortium consists of 19 (as at 2012) cancer hospitals and institutions located in different European countries and Japan working in a variety of disciplines.

3. Why have you been chosen?

You have been chosen because your cancer is one of the types that are being considered in the P-MEDICINE project.

P-MEDICINE will involve several hundred cancer patients. Currently, it is focusing on three types of cancer: (1) breast cancer in women, (2) Acute Lymphoblastic Leukaemia (ALL) in children, and (3) kidney cancer called nephroblastoma or Wilms Tumour in children. The data are currently collected at hospitals in Germany, Italy, and the United Kingdom. It is envisaged that more cancer diseases will be considered and that more hospitals will become involved over the next few years.

4. Do you have to take part?

Your participation to this study is entirely voluntary. If you decide to take part you will be asked to sign a consent form. By signing the consent form, you will confirm that you were properly informed about this project and that any questions you had have been answered. A copy of the patient information sheet and of the consent form will be given to you to keep.

If you decide to take part, you are free to withdraw your consent at any time and leave the study without giving any reason. This, or the decision not to take part, will not affect your medical care or the relationship with your medical doctor or medical staff.

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

If you have decided to take part your tumour and/or blood will be analysed by your physician or medical staff with respect to different characteristics, such as cell types present, characteristics of proteins, and patterns of gene function or malfunction (e.g. gene activity).

These data will be sent to P-MEDICINE in an anonymous way for computer analysis and experiments. They will be compared with data of other patients, which are also processed in an anonymous way (see below).

Since research is being done on blood and tumour samples, which have already been collected from you in connection with the clinical trial or study in which you are enrolled, your participation in this project will not require extra visits to the hospital, nor extra examinations.

6. How is your data protected?

The data that will be sent to P-MEDICINE consist of socio-demographic data (your 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), biological data (characteristics of cells and proteins, etc.), and genomic data (for example, data on the genes which are typical in your cancer).

Personal information such as your name or address will never be disclosed to P-MEDICINE. Before any data are sent to P-MEDICINE, all personal identifiers (linking the data to you) will be removed by a special anonymisation tool and replaced by a coded pseudonym. The code can only be unlocked with a special key, which will be guarded by an independent Trusted Third Party. The only circumstance in which the key may later be used to unlock the code, and thus re-identify you, is where this is in your own interests (e.g. an improved treatment becomes available) and you have decided you wish to be informed of this (see point 7).

Your data will be stored as described until the present project is completed, which will be in 2015 or later, or until a follow-up project with the same purpose and objective is completed. We will not retain your data beyond this point without your further express agreement.

If you decide to withdraw your consent to this project, no more data will be sent to P-MEDICINE. Moreover, the data, which have already been sent to P-MEDICINE, can only be further used if they are strictly anonymous (including for your physician or medical staff). To achieve this, the pseudonymisation key held by the Trusted Third Party will be destroyed so that the data can no longer be linked back to your person.

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

P-MEDICINE is aiming to generate a support scaffold or infrastructure for conducting basic and clinical scientific research in cancer. In the long term there is good reason to hope this will lead to improved treatment for patients with cancer. However, the results generated by P-MEDICINE are at this stage unlikely to be relevant for the treatment of any individual single patient. Thus, in general you will not be personally informed about the results of the research conducted on you in the context of P-MEDICINE.

It remains possible that research conducted in the context of P-MEDICINE may also yield results, which are of direct relevance for your own treatment or for the prevention of future ailments. If you consent to participating in this project, you may choose whether or not you wish to be informed of such results by treating physician.

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.

8. Risks and benefits of this project

The data transmitted to P-MEDICINE will be extracted from your patient file and the blood and tissue samples, which have been collected by the treating physician or medical staff before diagnosis and during treatment. Therefore, you will not be subject to any extra medical procedures, examinations or visits involving any risk.

As described under point 6, before your data are used for research, personal information (e.g. name, address, etc.) will be removed with the intention that you cannot be identified from the data that remains. This will occur by means of a technically advanced process, which complies with data security standards prescribed by law. There is a residual, albeit extremely small risk that these data might be linked back to your person.

As noted under point 7, it is unlikely you will receive a direct therapeutic benefit from participating in the study. Nor is any other form of benefit, e.g. a financial reward, contemplated. However, by making available your data for research use, you will contribute in an important way to the advancement of knowledge about cancer and the development of new and better treatments for others.

9. Information of relatives

Your health data, in particular your genetic data, might also concern your relatives. In the event that information regarding your close relatives (e.g. your parents, siblings, or children) would be beneficial to the project, you should be aware that your physician is allowed to disclose information about you to your relatives exclusively for the purpose of requesting consent from your relatives to the processing of their data in the purpose of the P-MEDICINE project.

10. Costs

There will not be any additional costs for you if you decide to participate in the P-MEDICINE project.

PATIENT CONSENT FORM

(Version 1.0, January 2012)

I, the undersigned ……………………, born on the………, in …………………., and living in………………………………….., agree by signing this consent form to take part in the project: "**From data sharing and integration via VPH models to personalized medicine**" (EU Grant agreement number: 270089), called P-MEDICINE in this document.

(Patient to initial box)

□ I confirm that I have read and that I understand the patient information sheet (version dated……….).

□ I confirm that I was given the opportunity to ask my attending physician and the medical staff any questions regarding the P-MEDICINE project, the general terms, the information sheet, and the present consent form; and I confirm that I am satisfied with their answers.

□ I understand that the data controllers are the Center for Data Protection (CDP) and the P-MEDICINE end-users having signed the end-user agreement, which provides data protection and data security policies, and that the supervisory authority reference is……………………………

□ I understand that my attending physician, Dr…………………., is my contact person for all questions that I might have regarding the P-MEDICINE project and the rights I enjoy against the data controller.

□ I understand that I am free to decide whether or not to participate in the P-MEDICINE project and that refusing to participate will not affect the quality of my medical care or my legal rights.

□ I understand that I am free to withdraw my consent at any time without giving any reasons and that this will not affect the quality of my medical care or my legal rights.

□ I understand that I have all the rights described in the general terms form above to access my processed data, correct my processed data, and object to their processing. My requests concerning these rights will be transmitted to P-MEDICINE via my attending physician by letter or secured Email.

□ I understand that this consent form refers to the general terms, which are an integral part of the present form, and to European law (namely the Directive 95/46/EC of the 24.10.1995) on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

□ I understand and I agree that samples from my tumour, blood, tissue and other biological samples will be analysed.

□ I understand that information about me, about my disease, about genetic tests performed on my tissue samples, and information contained in my medical records will be transferred to the P-MEDICINE network that consists of databases which are located in Member States of the European Union, and that it will be used for the purposes of the P-MEDICINE research project only. I understand that to guarantee anonymity of my data a state of the art pseudonymisation procedure will be undertaken and the key for re-identification held by an independent Trusted Third Party (TTP). I understand that the TTP is……………………………

□ I understand and agree that following such pseudonymisation, the processing of those data may be carried out by a scientific researcher (end user) who is not necessarily a health professional subject under national law or rules established by national competent bodies to the obligation of professional secrecy or a person subject to an equivalent obligation of secrecy.

□ I understand that my health data, in particular my genetic data, may not be disconnected from information concerning close relatives (parents, siblings and/or children). Therefore, I allow the attending physician to disclose information about me to them for the sole purpose of requesting consent from my close relatives for the processing of their data in the P-MEDICINE project. I understand that this will be done only if it is allowed by, and in compliance with, the law of this country.

□ I **do/do not** agree that DNA material from my tumour, blood, tissue and other biological samples can be analysed using genetic and other tests (delete as appropriate).

□ I understand that it may be possible, although unlikely, that results from the research conducted in the context of P-MEDICINE may be of direct relevance for my treatment or for the prevention of future illness. I **do/do not** wish to be informed of such results by my attending physician (delete as appropriate).

□ I understand that three original copies of this consent form will be produced and will be retained by me, the CDP and the hospital respectively.

Name of the patient:………………..

Signature of the patient:………………

Date (please date your own signature):…………

2. Minor or patient under mental disability

Legal representative information sheet (where patient is a minor or mentally disabled)

(Version 1.0, January 2012)

Explanation of the research project P-MEDICINE

(**From data sharing and integration via VPH models to personalized medicine**)

1. Invitation:

You are being invited, as legal representative of the minor or the person under disability mentioned below, to take part in a clinical research project involving patients with cancer. The study is called P-MEDICINE for short. The full name is: "**From data sharing and integration via VPH models to personalized medicine**". It is a European Union funded project (EU Grant agreement number: 270089).

P-MEDICINE aims to improve treatment of cancer and to promote research on cancer by building up a technical trans-European infrastructure to connect trials on cancer running in different hospitals, healthcare and research organisations all over Europe. Researchers participating in P-MEDICINE will be able to work more efficiently and successfully in developing new treatments for cancer and in personalising the treatment, so that patients suffering from cancer can be treated better. In particular, P-MEDICINE aims to identify individualised molecular profiles correlated with sensitivity or resistance to therapy. Each patient suffering from cancer has a particular molecular profile. In many cases the success of a specific treatment depends on the molecular profile of a patient. In this regard P-MEDICINE aims to develop a software tool linking a particular molecular profile to the best available treatment for someone with this profile. Your data will be used to build up this new technical framework and enable researchers to conduct scientific research by using this new infrastructure.

If you decide to take part, you will be asked to provide the data of the patient anonymously to this project. Before you make this decision, 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 make your decision. Please read the information provided carefully and discuss it with others if you wish. Feel free to ask your physician and other members of your healthcare team if there is anything unclear or if you would like more information.

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 at any time without any disadvantages. Also, if you decide not to volunteer for the project, it will not affect the treatment of the patient in any way.

2. Purpose of the project

It is the aim of P-MEDICINE to create a collaborative environment for the development, validation and sharing of VPH models and simulators and their transformation into decision making tools for clinical routine. Three exemplary multiscale simulation models of clinical tumour response to treatment will be developed: one for nephroblastoma, one for breast cancer and one for acute lymphoblastic leukaemia (ALL), based on the principles that have been shown to be most appropriate for the clinical trial context. These three models will constitute the simulation core of the "P-MEDICINE Oncosimulator" which will provide an integrated platform for simulating, investigating, better understanding and exploring the natural phenomenon of cancer and - after successful validation and transformation - an integrated treatment support tool. Modelling and validation will require access to large sets of data, which will be made available through a data warehouse. The resulting models of tumour response to treatment and the Oncosimulator will be made available to the users in the P-MEDICINE workbench for sharing, usage and exploitation. The tools and scientific knowledge produced by this project will be publicly available. The project itself will not necessarily develop new individual treatments, but the knowledge gained within the project should eventually improve our understanding of which patient is most likely to benefit from an individual treatment. It will also help to predict cancer prognosis and the chance of the cancer recurring or not.

P-MEDICINE is sponsored by the European Union. The P-MEDICINE research consortium consists of 19 (as at 2012) cancer hospitals and institutions located in different European countries and Japan working in a variety of disciplines.

3. Why has the patient been chosen?

The patient you are representing has been chosen because his/her cancer is one of the types that are being considered in the P-MEDICINE project.

P-MEDICINE will involve several hundred cancer patients. Currently, it is focusing on three types of cancer: (1) breast cancer in women, (2) Acute Lymphoblastic Leukaemia (ALL) in children, and (3) kidney cancer called nephroblastoma or Wilms Tumour in children. The data are currently collected at hospitals in Germany, Italy, and the United Kingdom. It is envisaged that more cancer diseases will be considered and that more hospitals will become involved over the next few years.

4. Does the patient have to take part?

Your decision to enter the patient in this study is, just as with any other decision regarding the patient, for you to take in accordance with your assessment of the patient's presumed will and/or his/her best interests. If you decide to enter the patient you will be asked to sign a consent form on behalf of the patient. By signing the 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.

If you decide to consent to the patient's participation, you are free to withdraw your consent at any time and withdraw the patient from the study without giving any reason. This, or the decision not to take part, will not affect the medical care provided to the patient or the relationship between you and/or the patient and the patient's medical doctor or medical staff.

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

If you have decided to take part, the tumour and/or blood will be analysed by your physician or medical staff with respect to different characteristics, such as cell types present, characteristics of proteins, and patterns of gene function or malfunction (e.g. gene activity).

This data will be sent to P-MEDICINE in an anonymous way for computer analysis and experiments. It will be compared with data of other patients, which are also processed in an anonymous way (see below).

Since research is being done on blood and tumour samples, which have already been collected from the patient in relation to the clinical trial or study in which he/she is enrolling, participation in this project does not imply extra visits to the hospital, nor extra examinations.

6. How is the data protected?

The data that will be transmitted to P-MEDICINE is 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), biological data (characteristics of cells and proteins, etc.), and genomic data (for example, data on the genes which are typical of your cancer).

Personal information such as name or address will never be disclosed to P-MEDICINE. Before any data are sent to P-MEDICINE, all personal identifiers (linking the data to the patient) will be removed by a special anonymisation tool and replaced by a coded pseudonym. The code can only be unlocked with a special key, which will be guarded by an independent Trusted Third Party. The only circumstance in which the key may later be used to unlock the code, and thus re-identify the patient, is where this is in the interest of the patient (e.g. an improved treatment becomes available) and where you have decided you wish the patient to be informed of this (see point 7).

The data will be stored as described until the present project is completed, which will be 2015 or later, or until a follow-up project with the same purpose and objective is completed. We will not retain such data beyond this point without your further express agreement.

If you decide to withdraw the consent to this project, no more data will be sent to P-MEDICINE. Moreover, the data, which have already been sent to P-MEDICINE, can only be further used if they are strictly anonymous (including for the patient´s physician or medical staff). To achieve this, the key held by the Trusted Third Party will be destroyed so that the data can no longer be linked back to the patient.

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

P-MEDICINE is aiming to generate a support scaffold or infrastructure for conducting basic and clinical scientific research in cancer. In the long term there is good reason to hope this will lead to improved treatment for patients with cancer. However, the results generated by P-MEDICINE are at this stage unlikely to be relevant for the treatment of any individual single patient. Thus, in general neither the patient nor you as the patient´s legal representative will be personally informed about the results of the research conducted on the patient in the context of P-MEDICINE.

It remains possible that research conducted in the context of P-MEDICINE may also yield results, which are of direct relevance for the patient´s own treatment or for the prevention of future ailments. If you consent to participating in this project, you may choose whether or not you wish to be informed of such results by the medical doctor.

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.

8. Risks and benefits of this project

The data transmitted to P-MEDICINE will be extracted from the patient file and the blood and tissue samples, which have already been collected by the treating physician or medical staff before diagnosis and during treatment. Therefore, there will not be any extra medical procedures, examinations or visits involving any risk for the patient.

As described under point 6, before the patient’s data are used for research, personal information (e.g. name, address, etc.) will be removed with the intention that he/she cannot be identified from the data that remains. This will occur by means of a technically advanced process, which complies with data security standards prescribed by law. There is a residual, albeit extremely small risk that these data might be linked back to the patient.

As noted under point 7, it is unlikely that the patient will receive a direct therapeutic benefit from participating in the study. Nor is any other form of benefit, e.g. a financial reward, contemplated. However, by making available the data for research use, you will contribute in an important way to the advancement of knowledge about cancer and the development of new and better treatments for others.

9. Information of relatives

The health data, in particular the genetic data of the patient might also concern his/her relatives. In the event that information regarding the patient´s close relatives (e.g. parents, siblings, or children) would be beneficial to the project, you should be aware that the patient´s physician is allowed to disclose information about the patient to his/her relatives exclusively for the purpose of requesting consent from his/her relatives to the processing of their data in the purpose of the P-MEDICINE project.

10. Costs

There will not be any additional costs for you if you decide to participate in the P-MEDICINE project.

LEGAL REPRESENTATIVE CONSENT FORM

(Version 1.0, January 2012)

I, the undersigned ………………………., born on the………., in ……………., and living in………………………., as legal representative of ……………………………, born on the……………,in………………… and living in…………………….., agree by signing this consent form to the patient's participation in the project: "**From data sharing and integration via VPH models to personalized medicine**" (EU Grant agreement number: 270089), called P-MEDICINE in this document.

(Legal representative of the patient to initial box)

□ I confirm that I have read and that I understand the legal representative information sheet (version dated……….).

□ I confirm that I was given the opportunity to ask the patient's attending physician and the medical staff any questions regarding the P-MEDICINE project, the general terms, the information sheet, and the consent form; and I confirm that I am satisfied with their answers.

□ I understand that the data controllers are the Center for Data Protection (CDP) and the P-MEDICINE end-users that have signed the end-user agreement providing data protection and data security policies and that the supervisory authority reference is……………………………

□ I understand that the patient's attending physician, Dr.…………………, is my contact person for all questions that I might have regarding the P-MEDICINE project and the rights and protections enjoyed by the patient.

□ I understand that I am free to decide whether or not to participate the patient in the P-MEDICINE project and that refusing to participate will not affect the quality of the medical care for the patient or my legal rights.

□ I understand that I am free to withdraw my consent at any time without giving any reasons and that this will not affect the quality of the medical care for the patient or my legal rights.

□ I understand that I have all the rights described in the general terms form above to access the processed data, correct the processed data, and object to their processing. My requests concerning these rights will be transmitted to the CDP via the physician attending the patient by letter or secured Email.

□ I understand that this consent form refers to the general terms form, and to European law (namely the Directive 95/46/EC of the 24.10.1995) on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

□ I understand and I agree that samples from the patient's tumour, blood, tissue and other biological samples will be analysed.

□ I understand and I agree that information about the patient, about his/her disease, about genetic and other tests performed on his/her tissue samples, and information contained in his/her medical records will be transferred to the P-MEDICINE network consisting of databases which are located in Member States of the European Union, and that it will be used for the purposes of the P-MEDICINE research project only. I understand that to guarantee anonymity of his/her data a state of the art pseudonymisation procedures will be undertaken; key for re-identification is hold by an independent Trusted Third Party (TTP). I understand that the TTP chosen is……………………………

□ I understand and agree that, from the moment the TTP receives pseudonymised data, the processing of those data won't be made by a health professional subject under national law or rules established by national competent bodies to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy.

□ I understand that genetic data in respect of the patient cannot be disconnected from information concerning his/her close relatives (parents, siblings, and/or children). Therefore, I allow the patient’s physician to disclose information about him/her for the sole purpose of requesting consent from his/her close relatives for the processing of their data in the P-MEDICINE project. I understand that this will be done only if it is allowed by, and in compliance with, the law of this country.

□ I **do/do not** agree that DNA material from the patient's tumour, blood, tissue and other biological samples can be analysed using genetic and other tests (delete as appropriate).

□ I understand that it may be possible, although unlikely, that results from the research conducted in the context of P-MEDICINE may be of direct relevance for the treatment of the patient or for the prevention of future ailment. I **do/do not** want to be informed of such results by the attending physician (delete as appropriate).

□ I understand that three original copies of this consent form will be produced and will be kept by myself as the legal representative of the patient, the CDP and the hospital/investigator respectively.

Name of the legal representative of the patient:………………..

Signature of the legal representative of the patient:………………

Date (please date your own signature):…………

3. Minor's agreement

AGREEMENT

(Version 1.0, January 2012)

I, the undersigned …………………………… born on the………, in……………. and living in…………………………, declare my agreement to take part in the project: "**From data sharing and integration via VPH models to personalized medicine**" (EU Grant agreement number: 270089), called P-MEDICINE in this document.

I understand that, at the present time, I am represented by my legal representative who is………………………………………………

I have read/my legal representative has read to me (delete as appropriate) the information sheet and the General Terms, which are part of the present document. My legal representative has explained to me what these terms mean, and how they may affect me as a project participant. I accept them.

I understand that when I reach the age of majority I will be asked again to decide if I wish to continue to take part in the project. I will then be legally able to exercise all my rights described in the general terms and in the patient consent form.

I understand that four original copies of this agreement will be produced and will be kept by me, my legal representative, the CDP and the hospital/investigator respectively.

Name of the minor patient:………………..

Signature of the minor patient:………………

Date (please date your own signature):…………

4. Hospital's/investigator's agreement

AGREEMENT

(Version 1.0, January 2012)

I, the undersigned …………………………, born on the…………, in ………………… and living in……………………………, declare that I act as the legal representative of the hospital/investigator ………………………………………………………… (the statute giving this power must be annexed).

I agree by signing this form to the hospital's/investigator's participation in the project: "**From data sharing and integration via VPH models to personalized medicine**" (EU Grant agreement number: 270089), called P-MEDICINE in this document.

I have read, I understand and I agree with the general terms - which are part of this document (version dated………………..).

In order to comply with the Directive 95/46/EC, the hospital/investigator will code or pseudonymise all the data of the patients who take part in the P-MEDICINE project. The TTP is……………………………

The TTP will hold the pseudonymisation key for the de-identification of the patients. Having regard to the state of the art and the cost of the implementation of the pseudonymisation tool, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the sensitive data to be protected.

I accept that the hospital/investigator will retain full legal and ethical responsibility for the security of its own database(s) and the processing of the patients' data, including the initiation of the pseudonymisation procedure before transferring the data to an organisationally separated "P-MEDICINE database" possibly located within the hospital/investigator.

I understand that two original copies of this agreement will be produced and they will be kept by the CDP and the hospital/investigator respectively.

Name of the hospital/investigator legal representative:………………..

Signature of the hospital/investigator legal representative:………………

Date (please date your own signature):…………

5. Physician's agreement

AGREEMENT

(Version 1.0, January 2012)

I, the undersigned Dr. …………………………………, born on the………., in …………………………… and living in …………………………………, declare by the present consent form to subscribe to the project: "**From data sharing and integration via VPH models to personalized medicine**" (EU Grant agreement number: 270089), called P-MEDICINE in this document.

I have read, I understand and I agree to the general terms - which are part of this document (version dated………..).

I agree to be the representative of the hospital/investigator towards the patient or his/her legal representative. I will provide the latter with all the information requested and needed. I agree to receive and respond to his/her requests relating to his/her rights described in the general terms or the Directive 95/46/EC.

I understand that two original copies of this agreement will be produced and will be kept by me and the CDP respectively.

Name of the physician:………………..

Signature of the physician:………………

Date (please date your own signature):…………

Annex 2:
Contract on data protection and data security within P-MEDICINE

(Version 1.0, January 2012)

between

the Center for Data Protection ("CDP")

Rempart de la Vierge, 5, Namur, Belgium 5000

____

(address and country of establishment)

and

____

("P-MEDICINE-end user")

____

(address and country of establishment)

Individually referred to as a "Party" or collectively referred to as the "Parties".

Preamble

The project P-MEDICINE (**From data sharing and integration via VPH models to personalized medicine**) is a European financed project supported by partners from eleven European countries and Japan, coming from different backgrounds, including physicians, clinicians, genomic scientists, medical- and bio-informaticians, and legal and ethical experts. P-MEDICINE brings together international leaders in their fields to create an infrastructure that will facilitate the translation from current medical practice to personalised medicine. In achieving this objective P-MEDICINE has formulated a coherent, integrated workplan for the design, development, integration and validation of technologically challenging areas of today.

The emphasis of P-MEDICINE is on formulating an open, modular framework of tools and services, so that it can be adopted gradually, including efficient secure sharing and handling of large personalised data sets, building standards-compliant tools and models for research, and providing tools for large-scale, privacy-preserving data and literature mining. P-MEDICINE will ensure that privacy, non-discrimination, and access policies are aligned to maximize protection of and benefit to patients. The P-MEDICINE tools and technologies will be validated within the concrete setting of advanced clinical research. Pilot cancer trials have been selected based on clear research objectives, emphasizing the need to integrate multilevel datasets, in the domains of Wilms tumor, breast cancer and leukemia. To sustain a self-supporting infrastructure realistic use cases will be built that will demonstrate tangible results for clinicians. The project is clinically driven and promotes the principle of open source and open standards.

Sharing clinical and genomic expertise implies the transfer and exchange of patient data within the project. Therefore the Infrastructure of P-MEDICINE is embedded in the P-MEDICINE Data Protection Framework, which guarantees compliance with current European data protection legislation, primarily by de facto anonymising the patient data. Due to the diverse participants it is of high importance to process patient data in accordance with the P-Medicine General Terms on data protection (ANNEX 1) and keep the data exchange within the project under the control of P-MEDICINE. This is guaranteed by the implementation of the P-MEDICINE Center for Data Protection (CDP) as central data controller for the P-MEDICINE data warehouse. The P-MEDICINE end-users (scientific researchers) will process data received form the P-MEDICINE data warehouse as data controllers themselves.

This contract is needed to state the conditions and obligations under which a P-MEDICINE end-user is allowed to get data from the CDP and to conduct research. It also describes the procedure in case of a re-identification request, as well as the obligations of the CDP with regard to P-MEDICINE end-users.

Clause 1: Definitions

For the purposes of the Clauses:

1. **Personal data, process/processing, data controller/controller, processor data subject, technical and organisational security measures and supervisory authority/authority** shall have the same meaning as in Directive 95/46/EC of 24 October 1995 (whereby **supervisory authority/authority** shall mean the competent data protection authority in the territory in which the data controller is established);
2. **Patient** shall mean a person treated in a healthcare organisation/ hospital, whose mainly health related (e.g. genetic) data are processed within the P-MEDICINE network;
3. **Healthcare organisation/hospital** shall mean a data controller who transfers patient data to P-MEDICINE. The healthcare organisations/hospitals are the connection between the patients willing to participate in P-MEDICINE and P-MEDICINE itself. They are responsible for safeguarding the patient's privacy rights and data security issues in their own organisation and for the transmission of the data to P-MEDICINE.
4. **Center for Data Protection (CDP)** shall mean the central data controller of the P-MEDICINE data warehouse, who agrees to receive from the healthcare organisations/hospitals data intended for processing in accordance with the P-MEDICINE General Terms (Annex A) and the terms of this contract. The CDP also serves as a central data protection authority within P-MEDICINE ensuring the compliance of all P-MEDICINE participants with the Data Protection Framework, particularly with regard to P-MEDICINE's policies and procedures;
5. **P-MEDICINE data warehouse** shall mean the database/databases where the patient data used within P-MEDICINE are stored (also referred to as P-MEDICINE database);
6. **Trusted Third Party** shall mean an independent security authority, which has no interest in the content of the processed data and can therefore be trusted by all participants of the P-MEDICINE project. It is a data controller itself.
7. **P-MEDICINE end user** shall mean the entity or person, e.g. healthcare organisation/hospital or investigator, conducting scientific research and participating in P-MEDICINE after having signed this "Contract on data protection and data security within P-MEDICINE" (P-MEDICINE end user agreement).
8. **De-facto anonymous data** shall mean data that has been modified in such a way that the information concerning personal or material circumstances can be attributed to an identified or identifiable individual only with a disproportionate amount of time, expense and labour.
9. **Clauses** shall mean these contractual clauses.

Clause 2: Scope and responsibility

(1) The scope of this contract are all issues related to data protection and data security deriving from the storage of data and transfer of data from the CDP to the P-MEDICINE end user and the use of such data by the P-MEDICINE end user.

(2) The CDP is responsible as data controller for the storage in the P-MEDICINE data warehouse and the transfer of data to the P-MEDICINE end user, whereas the later processes the data received on its own responsibility within its own organisation as data controller. The P-MEDICINE end user is solely responsible for complying with current data protection legislation once the data are transferred to his/her organisation.

Clause 3: Obligations of the CDP

The **CDP** agrees and warrants:

1. that the processing of the personal data within the P-MEDICINE data warehouse, including the transfer to the P-MEDICINE end user, has been and will continue to be carried out in accordance with the relevant provisions of the applicable data protection law and the P-MEDICINE General Terms (Annex A);
2. that it will transfer or disclose data only to P-MEDICINE end users on the basis of the Contract on data protection and data security within P-MEDICINE;
3. that it will not match any data of patients participating in P-MEDICINE with any other data in order to re-identify the patient;
4. that it will forward enquiries from patients and the authority concerning the processing of data to all P-MEDICINE end users concerned;
5. that it will provide adequate data security measures;
6. to make available to the patients upon request a copy of the Clauses set out in this contract.

Clause 4: Obligations of the P-MEDICINE end user

The **P-MEDICINE end user** agrees and warrants:

1. to process data in compliance with:
   (a) his applicable national data protection law,
   (b) the Data Protection Directive 95/46/EC and
   (c) the provisions stated in this contract;

   if he/she cannot provide such compliance for whatever reasons, he/she agrees to inform promptly the CDP of his/her inability to comply, in which case the CDP is entitled to suspend his/her processing of data and/or terminate the contract;
2. that he/she has implemented and follows appropriate technical and organisational security measures to protect the data against misuse and loss, in accordance with the requirements, in particular as stated in Annex B, before beginning the data processing. The P-MEDICINE end user ensures that the internal organisation of his/her enterprise meets the specific requirements of data protection. The P-MEDICINE end user will bind any other person in his/her sphere of influence with possible access to data provided by P-MEDICINE to obey the obligations and duties set by this contract;
3. that he/she ensures that when P-MEDICINE data are stored by him/her, it is technically and organisationally separated from other data;
4. that he/she does no matching between data received via P-MEDICINE and any other data and that he/she uses no other means in order to identify the patient concerned;
5. that he/she conducts only scientific research that is in line with P-MEDICINE policies, in particular with the P-MEDICINE General Terms (Annex A). For this reason the P-MEDICINE end user shall be allowed to store and use the data received via P-MEDICINE. The P-MEDICINE end user shall not be allowed to transmit or disclose any data received via P-MEDICINE to any third party nor to disclose or publish such data;
6. that he/she has no reason to believe that the legislation applicable to him/her prevents him/her from fulfilling the obligations under these Clauses. In the event of a change in this legislation which is likely to have a substantial adverse effect on the warranties and obligations provided by these clauses, he/she will promptly notify the change to the CDP as soon as he/she is aware, in which case the CDP is entitled to suspend his/her processing of data and/or terminate the contract;
7. that in case of a de-anonymisation request he/she contacts the CDP only, so that the patient can be identified by the CDP with the help of the Trusted Third Party that holds the key to link the pseudonymised data set to the patient concerned;

8. he/she shall provide the CDP upon request with copies of relevant data protection laws or references to them of the country in which he/she is established;
9. that he/she has the legal authority to give the warranties and fulfil the undertakings set out in these clauses;
10. that he/she shall promptly notify the CDP about:

(a) any legally binding request for disclosure of the data by a law enforcement authority unless otherwise prohibited, such as a prohibition under criminal law to preserve the confidentiality of a law enforcement investigation;
(b) any accidental or unauthorised access; and
(c) any request received directly from patients without responding to that request, unless he/she has been otherwise authorised to do so;

11. to deal promptly and properly with all inquiries from the CDP relating to his/her data processing and his/her data security measures (including technical and organisational measures referred to in Annex B);
12. to assist the CDP to provide information about the collection, processing or use of data, if the CDP has an obligation in current data protection law or due to contractual obligations with regard to the participating patients. The CDP has to request the P-MEDICINE end user in writing to do so. The P-MEDICINE end user will respond to the enquiry to the extend reasonably possible and within reasonable time.
13. upon reasonable request of the CDP he/she will submit his/her data processing facilities, data files and documentation needed for reviewing, auditing and/or certifying by the CDP (or any independent or impartial inspection agents or auditors, selected by the CDP and not reasonably objected to by the P-MEDICINE end user) to ascertain compliance with the warranties and undertakings in these clauses, with reasonable notice and during regular business hours. The same obligations apply in case a supervisory authority demands auditing;
14. to provide the CDP with contact details of the person responsible for data protection;
15. he/she will support the CDP and provide the CDP with all necessary information and documents needed in case of prior checking by supervisory authorities;
16. to inform the CDP immediately, should the data while in the hands of the P-MEDICINE end user be threatened with seizure or confiscation through bankruptcy or settlement proceedings, or through other circumstances or the actions of a third party.

Clause 5: Cooperation with supervisory authorities

(1) The CDP agrees to deposit a copy of this contract with the supervisory authority if it so requests or if such deposit is required under the applicable data protection law.

(2) The parties agree that the supervisory authority has the right to conduct an audit of the P-MEDICINE end user which has the same scope and is subject to the same conditions as would apply to an audit of the CDP under the applicable data protection law.

Clause 6: Confidentiality

(1) The P-MEDICINE end user must not disclose data or any other information acquired from CDP to any other person unless in pursuit of their duties as detailed in this contract.

(2) The P-MEDICINE end user is required to give an undertaking to maintain confidentiality, when he/she comes to any of the P-MEDICINE end user sites where he/she may see or have access to data delivered via P-MEDICINE. This undertaking shall continue to be valid after termination of the end user's activity and the termination of this contract.

(3) On request by the CDP the P-MEDICINE end user has to submit these undertakings to audit.

Clause 7: Right to audit/inspection

(1) With appropriate notice, the CDP can inspect the P-MEDICINE end user working premises during normal working hours and without disturbing work in progress to satisfy itself that adequate measures are being taken to meet the technical and organisational requirements according to the data protection laws and the rules set out by these Clauses.

(2) On written application, the P-MEDICINE end user undertakes to provide the CDP, within a period of one month, with all the information necessary or overall monitoring of performance of the contract.

Clause 8: Subcontractors

(1) The P-MEDICINE end user may issue commissions to subcontractors only with the written permission of the CDP.

(2) If subcontractors are engaged by the P-MEDICINE end user, the contractual agreements will be designed so that they conform with the requirements for confidentiality, data protection and data security between the partners to these Clauses. Rights to audit and control according to Clause 7 must be reserved to the CDP. The CDP is similarly entitled, on written demand, to be informed of the main content of the contract and the implementation of the data protection obligations of the subcontractor, including, if necessary, through access to the relevant contractual documentation.

Clause 9: Third-party beneficiary clause

(1) The patient can enforce against the CDP this Clause and clause 3 no (2) to (6), clause 4 no (4) and (5), clause 6 para. 1 and 2, clause 10 para. 1 and 2 and clause 12 para. 2 no (1) and (2) as third party beneficiaries.

(2) The patient is granted the right of access to data referring to him/her processed by the P-MEDICINE end user. The patient will address the information request to the CDP.

(3) The parties do not object to a patient being represented by an association or other body if the patient so expressly wishes and if permitted by national law.

Clause 10: Liability

(1) Each party shall be liable to the other party for damages it causes by any breach of these clauses. The parties agree that if one party is held liable for a violation of the clauses committed by the other party, the latter will, to the extent to which it is liable, indemnify the first party for any cost, charge, damages, expenses or loss it has incurred. Indemnification is contingent upon:

(a) the parties promptly notifying each other of a claim; and
(b) each party is given the possibility to cooperate in the defence and settlement of the claim.

(2) The parties agree that each party shall be liable for patient’s damages it caused by any negligent violation of data protection legislation or any other provisions of national or international law.

Clause 11: Penalty

(1) The parties agree that the P-MEDICINE end user pays a penalty of 10.000 (ten thousand) EUR for any negligent breach of clause _______________ of the contract caused by itself or any of its subcontractors.

(2) This shall be without prejudice to the parties’ right to terminate the contract, to seek compensation for damages or to enforce any claims under this contract.

Clause 12: Mediation and arbitration

(1) The parties agree with regard to disputes between each other in connection with these clauses:

1. that, subject to further agreement with each other, such disputes in conjunction with this contract can be referred to mediation by an independent person or, where applicable, by the supervisory authority;
2. that , subject to further agreement with each other, the resolution of a specific dispute in conjunction with these clauses can be referred to an arbitration body if the P-MEDICINE end user is established in a country which has ratified the New York Convention on enforcement of arbitration awards.

(2) The P-MEDICINE end user agrees with regard to patients' rights:

1. that if the patient invokes against him/her third-party beneficiary rights and/or claims compensation for damages under these clauses, the P-MEDICINE end user will accept the decision of the patient:
(a) to refer the dispute to mediation, by an independent person or, where applicable, by the supervisory authority; or
(b) to refer the dispute to the courts of Hanover, Germany.
2. that, subject to further agreement by the patient, the resolution of a specific dispute can be referred to an arbitration body if the P-MEDICINE end user is established in a country which has ratified the New York Convention on enforcement of arbitration awards.
3. to inform the CDP (and vice versa) about any disputes or claims brought by a patient concerning the processing of its data. The parties will cooperate with a view to settling any such claim amicably in a timely fashion.

Clause 13: Termination and obligations of the parties after the termination

(1) This contract will be terminated by 31st January 2015, if the P-MEDICINE project is not extended over such data and an agreement intervenes between the parties to continue the contract.

(2) In case of violation of clauses 3, 4, 6 by one of the parties, the other party is entitled to terminate this contract immediately

(3) Without prejudice to the foregoing provisions, any party may terminate this contract for good cause, giving the reason of such termination.

(4) Each party has to inform the other party by prior written notice in case of termination of the contract.

(5) The parties agree that on the termination of the provision of data processing services, the P-MEDICINE end user shall, at the choice of the CDP, return all the data and the copies thereof to the CDP or shall destroy all the data and certify to the CDP that he/she has done so, unless legislation imposed upon the P-MEDICINE end user prevents him/her from returning or destroying all or part of the data transferred. In that case, the P-MEDICINE end user warrants that he/she will guarantee the confidentiality of the data transferred and will not actively process the data transferred anymore.

Clause 14: Governing law and Jurisdiction, miscellaneous

(1) The Clauses shall be governed by German Law. The courts of Hanover/Germany shall have exclusive jurisdiction. This shall also apply to disputes on the validity of this clause.

(2) The Agreement and all other documents exchanged between the parties constitute the whole undertaking of the parties. All appendices attached hereto shall be deemed to be incorporated herein. Changes and amendments to this contract and all of its components, including any assurances by the CDP, require written agreement signed by the parties and an explicit statement that they represent a change or amendment to these conditions. The same applies to the waiving of this formal requirement.

(3) If any provision of this contract shall be entirely or partly invalid or unenforceable, this shall not affect the validity and enforceability of all other provisions of this contract. An invalid or unenforceable provision shall be regarded as replaced by such a valid and enforceable provision that as closely as possible reflects the privacy/security and/or economical purpose that the parties hereto had purposed with the invalid or unenforceable provision.

(4) Each person signing below and each party on whose behalf such person executes this Agreement warrants that he/she, as the case may be, has the authority and the legal capacity to enter into this Agreement and perform the obligation herein.

(5) This Agreement is being entered into force on the effective Date, i.e. the date of the last binding signature of this Agreement, All of the parties' rights and obligations contained in Clauses 4.2, 4.3, 4.4, 4.5, 4.7, 4.10, 4.12, 6., 9, 10, 11 shall survive termination.

Made in two signed copies, each party having received its own signed copy.

________________________________ ____

(Place, Date) (Signature of Nikolaus Forgó (president of the CDP)

________________________________ ____

(Place, Date) (Signature [P-MEDICINE end user])

Annex:

A. General Terms - version .01 (Jan. 2012)
B. Technical and organisational measures
C. Agreement on the Participation in the EU-project “Advancing Clinico-Genomic Trials on Cancer” (P-MEDICINE)

Annex A

GENERAL TERMS

(Version 1.0, January 2012)

Preamble:

The project P-MEDICINE (**From data sharing and integration via VPH models to personalized medicine**) in this present document, aims at creating clinico-genomic databases on cancer. The P-MEDICINE project will start up by collecting data on breast cancer (BRCA), Acute Lymphoblastic Leukemia (ALL) and nephroblastoma (PN), but it is projected to involve further cancer types in the future. The final purpose of such scientific research is to improve cure and management of future cancer patients by putting together the results of several researches running in Europe.

Therefore a data warehouse will be set up within P-MEDICINE to enable the project's participants to exchange patient data. This P-MEDICINE data warehouse will contain patient data transferred by participating hospitals/investigators, upon the patients´ informed consent to use their data within P-MEDICINE. The databases will not host anything else than data (thus excluding, for example, biomaterial).

All data transferred to the P-MEDICINE data warehouse will be pseudonymised through a dedicated state of the art pseudonymisation software. In order to be able to re-identify a given patient, for example in the event that a new treatment for him/her is developed, the name of the patient is replaced by a pseudonym during this procedure. The pseudonymisation key needed to link the pseudonymised data set to the patient concerned will be kept only by a Trusted Third Party (TTP). The TTP's independence from hospitals/investigators will be guaranteed. That means that the user (researcher) using the data will be unable to identify the patient to whom the data relates. In addition contracts are concluded between the participating hospitals/investigators and P-MEDICINE guaranteeing that patient data are not transferred to any party outside the project and no matching of data set takes place in order to identify the patients concerned. In interaction with strong technical and organisational security measures patient data in P-MEDICINE is to be seen as de-facto anonymous. Such data can only be de-anonymised by the TTP and with permission of the P-MEDICINE Center for Data Protection (CDP) if the de-anonymisation is needed in the interest of the patient concerned.

The data will be stored for a length no longer than the P-MEDICINE project. During the whole term of storage it will always be provided that the data remain de-facto anonymous for the end users. For a longer storage of patient data the explicit informed consent of the patient will be required.

The patient data remain under the control of the respective hospital/investigator (data exporter) where the data are collected until the data have been transferred to the P-MEDICINE data warehouse. From that point on the CDP will be responsible for the data processing within the P-MEDICINE data warehouse. The CDP concludes contracts with the P-MEDICINE end users that guarantee the protection and security of the data received from the data warehouse for the purposes of scientific research. Further the CDP controls the compliance to these contractual agreements. It, thus, serves as a central data protection authority for the P-MEDICINE framework. The P-MEDICINE end users process the data received from P-MEDICINE on their own behalf, so that they are to be regarded as data controllers according to the law.

The users (researchers) are not allowed to publish the data or to transmit or disclose data received via P-MEDICINE to any third person outside of P-MEDICINE.

These General Terms will apply to the CDP (as a legal person), the patients, physicians and end users.

Explanatory Glossary:

Anonymous data / Rendering anonymous

Rendering data anonymous means to modify personal data in a way that the information concerning personal or material circumstances can no longer be identified or identification is only possible with a disproportionate amount of time, expense and labour be attributed to an identified or identifiable individual. Data that have been anonymised are no longer "personal data" in the legal sense. It will be an aim to have as much anonymised data within P-MEDICINE as possible and reasonable.

Center for Data Protection (CDP)

The CDP shall mean the central data protection authority of the P-MEDICINE infrastructure, which agrees to receive from the healthcare organisations/hospitals (data exporters) data intended for processing in accordance with the terms of the data exporter agreement. The CDP guarantees privacy within the P-MEDICINE data warehouse.

Confidentiality

Persons employed in data processing shall not collect, process or use personal data without authorisation (confidentiality). On taking up their duties such persons shall be required to give an undertaking to maintain such confidentiality. This undertaking shall continue to be valid after termination of their activity. Any person acting under the authority of the CDP who has access to P-MEDICINE patient data must not process them except on instructions from the controller, unless he/she is required to do so by law.

Consent

Informed consent means any express indication of patient´s wish expressing his/her agreement to data relating to him/her being processed, provided that he/she has sufficient information about the purposes of the processing, the data or categories of data concerned, the recipient of the data, and the name and address of the controller and of his/her legal representative if any. The patient's consent must be freely given and specific, and may be withdrawn by the patient at any time. If the patient is incapable of a free decision or domestic laws do not permit the patient to act on his/her own behalf, consent is required of the person recognised as legally entitled to act in the interest of the patient or of an authority or any person or body provided for by law (legal representative).

Data controller

The data controller/controller is, according to the Data Protection Directive 95/46/EC, the natural or legal person who alone, or jointly with others, determines the purposes and means of the processing of personal data. The data controller is liable for the legality of the processing and the fulfilment of the obligations towards the national data protection authority and the patients. The hospitals/investigators (data exporters) are data controllers with regard to the collection of patient data and theirs transmission to P-MEDICINE, whereas CDP is the data controller with regard to the data stored in the P-MEDICINE data warehouse. Finally the P-MEDICINE end users are data controllers regarding the data received from the P-MEDICINE data warehouse.

Data processor

Data processor shall mean a natural or legal person, public authority, agency or any other body which processes patient data on behalf of the controller, such controller being liable for the legality of the processing and the fulfilment of the obligations towards the national data protection authority and the patients.

Data subject

The data subject is the subject of personal data, meaning an identified or identifiable person the data refers to. An identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity. As a rule the patient, whose data are collected and processed for P-MEDICINE will be the data subject, when his/her personal data are processed.

Disclosing

Disclosure is a processing operation in which patient data are provided by a controller to a third party. The data controller must only disclose data to third parties if permitted by law or by the data subject´s consent. In P-MEDICINE data are only disclosed to P-MEDICINE end users who have signed a special agreement that forbids any disclosure of data received via P-MEDICINE to any other third party.

Hospital

Hospitals are health institutions where patients are treated and their personal data are collected for the purpose of the P-MEDICINE project.

Investigator

The legal or natural person who gathers and manages the patient's data from the hospitals, laboratories etc. and maintains and controls the trial/study database.

Legal representative of the patient ("legal representative"):

The legal representative(s) of the patient is/are the person(s) who has/have the power by law or legal decision to decide for a minor patient (or equivalent status such as mentally disabled patients).

Necessary processing

When deciding which data will be collected and further processed, the controller must limit these data to the extent necessary to achieve the purpose of processing. This means that personal data will only be processed when it is necessary for the project.

Patient:

Patient means the person treated in a hospital. Certain data collected in the hospitals will upon the patient´s consent be transferred to the p-medicine framework where they will be used for the purposes of scientific research in de facto anonymous form.

Personal data

Personal data means any information relating to an identified or identifiable natural person ('data subject'). An identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity. Therefore a set of data collected under a certain number or sign “patient xxx”, “tissue YYY” can be personal data, if the patient concerned can still be identified by other means than his/her name.

Physician

The physician is the natural person who is in charge of the patient’s treatment.

Pseudonymisation

To pseudonymise a data set means to replace the patient's name and other identifying characteristics with a coded label in order to preclude direct identification of the patient or to render such identification substantially difficult. Within P-MEDICINE only pseudonymised data are processed.

Publish

The controller and the processors will refrain from publishing personal data or otherwise making them public, unless consent from the patient concerned is obtained.

Purpose

The purposes for processing of personal data must be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed. The purposes must be specified, explicit and legitimate. Personal data must be not further processed in a way incompatible with those purposes. The purpose for the collection, transfer and use of the data within P-MEDICINE is to conduct scientific research.

Sensitive (personal data)/Special categories of data

Sensitive personal data are personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and data concerning health (genomic data) or sex life. The processing of sensitive data is only allowed in case of certain exceptions explicitly stated by the national laws of the Member State.

Storage

Storage of personal data is allowed by the Data Protection Directive 95/46/EC. But when the purpose of processing is achieved and the data are not required any more for that particular purpose, personal data must be rendered anonymous or must be destroyed. Most national laws allow personal data to be stored for a longer term, provided that this is in order to use the data exclusively to carry out scientific research or statistics. Nevertheless, some national laws impose supplementary conditions or formalities in order to allow longer storage.

Technical and organisational measures

Organisational measures, together with technical measures, must ensure an appropriate level of security of the data processing, taking into account the state of the art and the costs of their implementation relative to the risks inherent in the processing and the nature of the data to be protected. Appropriate organisational measures shall be taken by the controller against accidental loss, destruction or alteration of, or damage to, personal data and against unauthorized or unlawful processing of personal data in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing. The controller must, where processing is carried out on his/her behalf, choose a processor providing sufficient guarantees in respect of the technical security measures and organisational measures governing the processing to be carried out, and must ensure compliance with those measures. Such appropriate organisational measures to ensure the confidentiality, integrity and accuracy of processed data should include for example:

- control of the entrance to installations
- control of data media
- memory control
- control of utilization
- access control
- control of communication
- control of data introduction
- control of transport
- availability control

Such technical and organisational measures have to be taken by all the P-MEDICINE-participants processing patient data.

Third Party

A third party is a natural or legal person, public authority, agency or any other body other than the patient, the controller, the processor or persons who, under the direct authority of the controller or the processor, are authorised to process the data. With regard to P-MEDICINE, third parties will be all the other persons and bodies who have no authorisation of P-MEDICINE to process the data.

Transfer

Transfer of data means the transmission of patient data from one data controller to another.

Trusted Third Party

The Trusted Third Party is an independent security authority, which has no interest in the content of the processed data and can therefore be trusted by all participants of the P-MEDICINE project. Within P-MEDICINE the Trusted Third Party will hold the pseudonymisation key needed to link data sets to the patient they belong to. The involvement of the TTP guarantees that a data set will only be re-identified, if the patient must be identified for medical or scientific reasons and if he/she wishes to be informed.

Article 1: Patient's rights

1.1. Information

The patient or his/her legal representative must be informed, in an intelligible form and fully, before giving his/her consent.

The information given to the patient or his/her legal representative will consist in describing and explaining:

- the identity of the data controller;
- the purpose of the processing of his/her data;
- his/her rights;
- the security of the data processed;
- the categories of data concerned;
- the recipients or categories of recipients;
- such other matters as may be specified from time to time by applicable data protection laws (including professionally enforceable codes of practice);

The information will be provided by the physician, who will act here as the representative of the hospital towards the patient.

1.2. Access

The patient has, directly or through his/her legal representative, a right of access to his/her data processed by or on behalf of the data controller.

This right of access includes inter alia the right to be informed:

- in an intelligible form of the data undergoing processing and of any available information as to their source;
- of the identities of the persons who have had access to his/her data and the moment of this access (log file).

The demand of access shall be addressed to the hospital through the patient´s physician.

This access is free of charge.

1.3. Object and withdrawal of consent

The patient or his/her legal representative has the right to object to the processing and to withdraw his/her former consent without giving reasons and at any time. In case the patient or his/her representative withdraws the given consent all data, which have not been pseudonymised yet by the TTP, can't be used anymore and the data already pseudonymised by the TTP have to be completely anonymised (the TTP must erase the key used for the pseudonymisation).

The exercise of these rights is free of charge.

1.4. Right to rectify

In case of inaccuracy the patient or his/her legal representative has the right to demand the rectification of his/her data from the data controller.

The exercise of this right is free of charge.

1.5. In case of death

If the TTP is informed (by the treating physician) that a patient died, his/her data can be used for research, but the TTP must erase the pseudonymisation key.

1.6. Feedback

The patient or his/her legal representative will receive information generated as a result of the research involving his/her personal data if:

- the information is likely to be directly useful to his/her therapy/treatment and
- (where he/she has no legal representative) he/she is physically and psychologically able to receive the information.

He/she may refuse this information by letter or secured Email sent to his/her physician, who for this purpose acts as the representative of the hospital or the CDP.

Article 2: Patient's obligation to provide information

At the time of providing consent to the processing of his/her data for the purposes of P-MEDICINE, the patient or his/her legal representative shall answer to the best of his/her knowledge and belief a questionnaire given to him/her by the physician. This questionnaire is designed to provide the physician with information regarding the patient's health, medical history, etc., which is necessary for the physician to perform his/her role in the context of the P-MEDICINE network.

Article 3: Rights of P-MEDICINE, hospitals/investigators and physicians

P-MEDICINE and the participating hospitals/investigators have the right to withdraw from the project any patient (non exhaustive list):

- who has given false information in the questionnaire mentioned in article 2;

- whose consent is, in their judgement, impaired by reason of pressure from relatives or other third persons;
- in respect of whom there are other grounds for believing further participation is not in his/her interest or in that of the project

In this case, all the data from the patient concerned will, at the discretion of P-MEDICINE, either be erased or completely anonymised in the same way as provided in 1.3.

P-MEDICINE has the right to stop the project at any time without giving reasons or any explanation. In this case, all the data within the P-MEDICINE network will be erased and no liability of any form will be incurred by P-MEDICINE.

P-MEDICINE has the right to exclude a physician, hospital/investigator or user (researcher) from the project in case of violation of these general terms, the contract or agreements with P-MEDICINE, the patient´s right to informed consent, national or international laws, or relevant codes of professional practice.

The physician, hospital/investigator or user has the right to withdraw from the program at any time without giving reasons or any explanation and without any payment of a penalty. In this case, his/her grant will be stopped and claimed back, if paid in advance.

If the physician withdraws or is required to withdraw from the project the patient, whose data are stored in the hospital, can remain in the project by choosing another physician of the hospital taking part in the project. If the hospital itself withdraws from the project, the patient’s data will remain in the project until the patient withdraws his/her consent. In that case the respective data will be completely anonymised in the same way as provided in 1.3.

Article 4: P-MEDICINE's, hospital/investigator, end user, physician and Trusted Third Party obligations

Access to the P-MEDICINE network must be granted by the CDP. The grant can only be given by the CDP.

The CDP is responsible for the security of any data processing within P-MEDICINE whereas the participating healthcare organisations/hospitals (data exporters) as well as the P-MEDICINE end users are responsible for the processing of patient data within their own organisation and are, thus, obliged to ensure the confidentiality and protection of the patient data processed. These obligations are defined by the contractual agreements concluded with the CDP.

In order to be in accordance with the Directive 95/46/EC, the hospital/investigator will implement a pseudonymisation tool guaranteeing a state of the art pseudonymisation of patient data transferred to the P-MEDICINE data warehouse. Having regard to the state of the art and the cost of their implementation, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the sensitive data to be protected.

The Trusted Third Party holds the pseudonymisation key to re-identify the patient. Re-identification is permitted only if the scientific research reveals findings that are likely to be useful for the patient´s treatment and the patient wishes to be informed.

Article 5: Applicable legislation and jurisdiction

In respect of any claim against the CDP involving an alleged breach of data protection legislation, Belgian data protection legislation shall be applicable.

Subject to the provision below in relation to third party beneficiary rights, any other dispute between the patient and the CDP, under this contract or otherwise, shall be determined by the national law of the country where the patient resides.

Third beneficiary rights (not involving an alleged breach of data protection) shall be governed by the laws applicable to the contract from which those rights derive. Any dispute or claim in relation to the same shall be within the exclusive jurisdiction of the courts specified in the relevant contract.

Annex B

Technical and organisational measures

(Version 1.0, January 2012)

The P-MEDICINE end user will take appropriate technical and organisational measures to protect the data received from the data warehouse against misuse and loss, in accordance with the requirements, in particular:

- to prevent unauthorised persons from gaining access to data processing systems with which patient data are processed or used (physical access control),
- to prevent data processing systems from being used without authorisation (denial of use control),
- to ensure that persons entitled to use a data processing system can gain access only to the data to which they have a right of access, and that patient data cannot be read, copied, modified or removed without authorisation in the course of processing or use and after storage (data access control),
- to ensure that patient data cannot be read, copied, modified or removed without authorisation during electronic transmission, transport or storage and that it is possible to examine and establish to which bodies the transfer of personal data by means of data transmission facilities is envisaged (data transmission control),
- to ensure that it is possible retrospectively to examine and establish whether and by whom patient data have been input into data processing systems, modified or removed (input control),
- to ensure that patient data being processed on commission are processed solely in accordance with the directions of the controller (contractual control),
- to ensure that patient data are protected against accidental destruction or loss (availability control),
- to ensure that data collected for different purposes can be processed separately (separation rule).

Annex C
Participation in the EU-project
P-MEDICINE

AGREEMENT

(Version 1.0, January 2012)

I, undersigned ……………………….............................., (title) born on the……………................, in…………………..................... and working in …………………………………/on behalf of ………………………………..(please cancel if not applicable) declare by the present consent form to subscribe to the project: “From data sharing and integration via VPH models to personalized medicine” (Grant agreement number 270089), called p-medicine in this document.

I have read, I understand and I agree to subscribe to the General Terms - which form a part of this document (Version 1.0, January 2012)

I understand that two original copies of this agreement will be produced and will be kept by me and p-medicine respectively.

Name of the user: ……………………………….

Name of the representative: ……………………………….

Signature of the user/ its representative: ……………………………….

Date (please date your own signature): ……………………………….