Deliverable No. 2.6.2

Regular update of the user needs and requirements based on evaluation and validation

Grant Agreement No.:	270089
Deliverable No.:	D2.6.2
Deliverable Name:	Regular update of the user needs and requirements based on evaluation and validation
Contractual Submission Date:	31/01/2014
Actual Submission Date:	31/01/2014

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

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	***p-medicine***
Project Full Name:	From data sharing and integration via VPH models to personalized medicine
Deliverable No.:	D 2.6.2
Document name:	Regular update of the user needs and requirements based on evaluation and validation
Nature (R, P, D, O)[1]	R
Dissemination Level (PU, PP, RE, CO)[2]	PU
Version:	1.0
Actual Submission Date:	31/01/2014
Editor: Institution: E-Mail:	Norbert Graf USAAR graf@uks.eu

ABSTRACT:

The main objective of this document is to present the activities focused on collection and description of the available updates of the user needs and requirements underlined in the frames of the proposed for implementation p-medicine project scenarios.

This report is a part of an iterative process of requirements analysis, elicitation, documentation and validation process. One of the main goals of this document is to present the linkage to the project's deliverables which describe in details the implementation activities of the elaborated, updated or new p-medicine's project scenarios (presented in the frames of D 2.6 and D 2.6.1).

KEYWORD LIST: user needs, requirements, use cases, evaluation, validation, update

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

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

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

MODIFICATION CONTROL			
Version	Date	Status	Author
0.1	29/11/2013	Draft	Ruslan David
0.2	12/12/2013	Pre-Final Draft	Ruslan David, Holger Stenzhorn, Norbert Graf
0.3	14/01/2014	Pre-Final Version	Ruslan David, Holger Stenzhorn, Norbert Graf
1.0	31/01/2014	Final Version	Ruslan David, Holger Stenzhorn, Norbert Graf

List of contributors

- Norbert Graf, USAAR
- Holger Stenzhorn, USAAR
- Ruslan David, USAAR

Contents

1 Executive Summary

The main objective of this document is to present the activities focused on collection and description of the available updates of the user needs and requirements underlined in the frames of the proposed for implementation p-medicine project scenarios.

This report is a part of an iterative process of requirements analysis, elicitation, documentation and validation process. One of the main goals of this document is to present the linkage to the project's deliverables which describe in details the implementation activities of the elaborated, updated or new p-medicine's project scenarios (presented in the frames of D 2.6 and D 2.6.1).

The main structure of this document is similar to the previous deliverable version (D 2.6.1), the difference is the alignment of the project's scenarios to the priority and implementation status. The collected and presented updates are presented in the Table 1.

Table 1: Priority table with implementation status

Nr	Scenario Name	Priority / Implementation Status
1. VPH Scenario		
1.1	VPH Toolbox Scenario	High / Implemented and reported (D 8.1.1)
2. Security Scenarios		
2.1	Single Sign-on Scenario	High / Implemented and reported (D 8.6.2 and D 3.4)
2.2	Single Sign-out Scenario	
2.3	Access Rights Scenario	
2.4	User Enrolment Scenario	
3. Clinical Scenarios - Nephroblastoma		
3.1	Pathway Scenario	Medium / Under development, evaluation or validation
3.2	Imaging Scenario	Medium / Under development, evaluation or validation
3.3	(Severe) Adverse Event ((S)AE) Prediction Scenario	Medium / Under development, evaluation or validation
3.4	Tumour Marker Scenario	Medium / Under development, evaluation or validation
3.5	Oncosimulator Scenario	Medium / Under development, evaluation or validation
4. Clinical Scenarios - Breast Cancer		
4.1	Breast Cancer Scenarios	High / Implemented and reported (D 9.7)

4.2	Oncosimulator Scenario	Medium / Under development, evaluation or validation
5. Clinical Scenarios - Acute Lymphoblastic Leukaemia		
5.1	Oncosimulator Scenario	Medium / Under development, evaluation or validation
6. Clinical Trial Scenarios		
6.1	Statistical analysis of cancer samples with associated gene expression data and clinical data	Medium / Under development, evaluation or validation
6.2	Data management in international clinical trials by ECRIN	High / Implemented and reported (D 8.1.1)
6.3	Use of data mining to improve study feasibility	Medium / Under development, evaluation or validation
6.4	Improved patient recruitment in oncological clinical trials	Medium / Under development, evaluation or validation
6.5	Use of p-medicine platform with decision support to conduct oncological clinical trials by ECRIN	Medium / Under development, evaluation or validation
6.6	Increased subject retention rates in oncological clinical trials	Medium / Under development, evaluation or validation
7. Education and Patient Empowerment Scenarios - Education and Training		
7.1	Help patients understand the patient empowerment tool	High / Implemented and reported (D 16.2)
7.2	Teach health care professionals when to use the p-medicine tools	
7.3	Impart understanding of the p-medicine environment	
7.4	Scenario for Education and Training	
8. Education and Patient Empowerment Scenarios - Patient Empowerment		
8.1	Search for running clinical trials in Europe	High / Implemented and reported (D 14.4)
8.2	Consent and Re-consent Scenario	
8.3	Informed Consent (Patient's Perspective)	
8.4	Own Data Scenario	
8.5	Summarize the history of the disease in an understandable way and increase patient-doctor understanding	

9. Biobank Access, Portal and IT Scenarios		
9.1	Access to Biobanks Scenario	High / Implemented and reported (D 10.2)
9.2	Linking the own biomaterial data repository to the p-medicine biobank access framework for the collaboration with specific user groups	
9.3	Managing patient's biomaterial and related data in clinical trials with ObTiMA	
9.4	Offering human biomaterial to a closed and/or open clinical research community for research	
9.5	Requesting specific human biomaterial within a closed and/or an open clinical research community for research purposes	
9.6	p-medicine portal scenario	High / Implemented and reported (D 8.1.2)
9.7	Sync and Push services	High / Implemented and reported (D 8.5)
9.8	Data translation for PUSH services	High / Implemented and reported (D 4.2)
9.9	Ontology annotation of external databases	High / Implemented and reported (D 4.2)
9.10	Ontology-Based Semantic Search Framework (this scenario could be renamed to "Data Mining Framework")	High / Implemented and reported (D 11.2)
10. ObTiMA Scenarios		
10.1	Pseudonymization Scenario	High / Implemented and reported (D 8.3)
10.2	Data Entry of Prospective Clinical Trial Data	High / Partially implemented and reported (D 8.4) (some scenarios are under development, evaluation or validation)
10.3	Data Manager of Prospective Clinical Trials	
10.4	eCRF Developer for Prospective Clinical Trials	
10.5	Data Synchronization with HIS during running trial in ObTiMA	
10.6	SAE/SUSAR Scenario	
10.7	Drug interaction Scenario	
10.8	DICOM Scenario	
10.9	Consultation Scenario	
10.10	Trial Development Scenario	

10.11	Trial Outline Builder Scenarios	
10.12	Participating Centres Scenario	
10.13	Patient Access to his/her trial data and Diary Scenario	
10.14	Repository Scenario	
10.15	Semantic Interoperability Scenario	
10.16	Reporting Scenario	
11. DoctorEye Scenarios		
11.1	Segmenting Nephroblastoma MRI	High / Implemented and reported (D 9.4)
11.2	Signal Intensity Scenario	High / Implemented and reported (D 8.2)
12. New Scenario		
12.1	Annotation of existing data set	High / Implemented and reported (D 4.5)

2 VPH Scenario

2.1.1 VPH Toolbox Scenario

The D 8.1.1 (Specification of the interaction with the VPH Toolkit) has described the general resources and specific tools developed under the VPH Network of Excellence which might be used in *p-medicine*, and *p-medicine* tools which might be contributed to the Network of Excellence toolkit. Because of the limited remaining lifespan and many changes of priority in the NoE, contributions to *p-medicine* are limited. The key output is the Toolkit Guidelines which specify how we might describe and improve our tools so they are deemed suitable for submission to the Toolkit portal. Likewise, our key contribution to the NoE will be in the form of tools submitted to the Toolkit portal. A small number of tools have been identified which might be of general interest to the VPH modelling community. However, due to the limited opportunities for interaction, and the unique challenges associated with exchanging tools and resources between the two projects, it is not regarded as worthwhile or even possible to specify a precise mechanism for this exchange, as suggested by the *p-medicine* Description of Work. *p-medicine* collaborators will attend many of the remaining VPH meetings to keep up-to-date on developments in the NoE, particularly with regard to sustainability of the work of the NoE once its funding finished in October 2012.

3 Security Scenarios

According to the D 8.6.2 (Initial version of the *p-medicine* integrated platform) the *p-medicine* security framework is designed around the SAML standard. A *p-medicine* Identity Provider (IdP) provides identity assertions to all services within *p-medicine*.

Web sites integrate with the IdP by providing a SAML compliant Identity Consumer presented in the frames of D 3.4 (Service Integration Guidelines). An Identity Consumer consumes and validates the assertions provided by the IdP.

REST Web Services integrate by accepting a SAML Identity Assertion through the HTTP Authorization header.

The general, summary description is when a user is redirected to the IdP when he visits the *p-medicine* portal. The user authenticates himself on the IdP and is then redirected back to the portal passing through an identity assertion. This assertion identifies the end user. Through a portlet the user then wishes to call a backend service. For this the portal requests a delegation token from the Secure Token Service. This delegation token contains the identity of the portal, which acts as intermediate service, and the end user on whose behalf the portal is acting. The delegation token is then passed through a HTTP Authorization header of the REST call to the backend service. This backend server in turn also requests a delegation token to call the data warehouse.

The related security scenarios are:

- Single Sign-on
- Single Sign-out
- Access Rights
- User Enrolment

4 Clinical Scenarios

4.1 Nephroblastoma

The D 9.2 (Report on the planning and management of the SIOP Wilms Tumour trial) gives details about the nephroblastoma study that will be run in Europe under the umbrella of SIOP-RTSG (International Society of Paediatric Oncology - Renal Tumour Study Group). It shows the results of SIOP 2001 as a basis for the SIOP Wilms Tumor trial running within Europe in 2012 to 2014. Tools of *p-medicine* are addressed that will be used in running this trial. Most important are ObTiMA, the data warehouse and DoctorEye.

4.1.1 Pathway Scenario

Scenario is under development (no updates are available).

4.1.2 Imaging Scenario

Scenario is under development (no updates are available).

4.1.3 (Severe) Adverse Event ((S)AE) Prediction Scenario

Scenario is under development (no updates are available).

4.1.4 Tumour Marker Scenario

Scenario is under development (no updates are available).

4.1.5 Oncosimulator Scenario

Scenario is under development (no updates are available).

4.2 Breast Cancer

4.2.1 Breast Cancer Scenarios

The D 9.7 (Report on the use cases of the breast cancer trials) describes breast cancer trials with metronomic and antiangiogenic chemotherapy ongoing at European Institute of Oncology within the framework of *p-medicine*. The primary aim of our studies is to maximize efficacy of therapy while minimizing side effects. Electronic patient records interfaced with bio-banks, genetic databases, and medical imaging systems will be available for new methodologies of data analysis.

Patients' characteristics and clinical/biological data collected are homogeneous.

4.2.2 Oncosimulator Scenario

Scenario is under development (no updates are available).

4.3 Acute Lymphoblastic Leukaemia

The main topic of the D 9.9 (Report on ALL integrated data analysis environment) was the description of patient data on childhood ALL in *p-medicine*. The data were collected in trial ALL-BFM 2000, conducted from July 2000 to the end of June 2006. The disease and its treatment are briefly introduced in the second chapter. Specific treatment information on the trial are also described in the second chapter by including a shortened version of the protocol. A detailed description of the patient data made use in the *p-medicine* community has been provided. Here, mainly the preparation of the data was explained. A detailed listing of the data that were disseminated to the partners in *p-medicine* was attached in the appendix. Furthermore, the collaboration with *p-medicine* partners was summarized and the work with our data and the benefits for the community were outlined. Finally the work with the

data at CAU was introduced and some first approaches to model patient-specific outcomes were presented.

4.3.1 Oncosimulator Scenario

Scenario is under development (no updates are available).

5 Clinical Trial Scenarios

Introduction

Clinical drug trials are investigations with humans intended to discover or verify the effects of one or more medicinal products. General requirements for the conduct of these clinical trials in the EU are provided for in EU Directive 2001/20/EC, which is called the “Clinical Trials Directive”, 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. The Clinical Trials Directive is concretised by Directive 2005/28/EC, the "the GCP Directive", 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.

There are a number of guidelines specifying various aspects of clinical trials, and in particular: (1) the information to be submitted to the competent authorities and to the ethics committees, (2) requirements on safety monitoring and the reporting of adverse reactions, (3) requirements regarding Good Clinical Practice, including the documentation, of the clinical trials (e.g. Trial Master File), (4) specific requirements regarding the products and the clinical trials, (5) inspections of competent authorities and the applicable procedures and (6) requirements relating to the quality, safety and efficacy of products. These guidelines have been published in Volume 10 of "EudraLex - The rules governing medicinal products in the European Union". The Clinical Trials Directive harmonises the rules in the EU for the approval of clinical trials conducted in a member state. For clinical trials in the US regulations are laid down in a number of Codes of Federal Regulations which are concretised in FDA regulations and guidances.

The European Commission is planning in 2012 to propose a revision of the Clinical Trials Directive 2001/20/EC. This update was supported by public consultation. Here ECRIN (European Clinical Research Infrastructures Network) has provided, together with other organisations and pharma companies, suggestion for improvements.

Because in *p-medicine* developed software will have direct implications on patient safety, e.g. by providing decision support and decisions on therapies, software must be validated according to GCP (System Validation) and according to medical device laws. Clinical trials on medical devices are regulated by EN ISO 14155:2011 "Clinical investigation of medical devices for human subjects - Good Clinical Practice" and in Europe by EU Directive 2007/47/EG. In addition, guidances on information security, electronic source data and data privacy play an important role for *p-medicine* software and data processing, for example ISO 27001: Information Security and in Europe EU Directive 95/46/EC, the GCP Inspectors Group’s Reflection paper on expectations for electronic source documents used in clinical trials and GMP Annex 11 and in the US HIPAA (Health Insurance Portability and Accountability Act of 1996).

p-medicine will use biobanks, advanced therapies in cancer research and innovative products and will therefore be subject to the corresponding regulations and guidelines, often still in the state of a draft. Especially this regulatory area, covering requirements for biomarkers, genetic tests and use of tissues, is in development.

The application of these regulations and guidances for several of the software applications developed in *p-medicine* is presented in detail in the frames of D 9.1 (Report on regulatory aspects of clinical trials).

5.1.1 Statistical analysis of cancer samples with associated gene expression data and clinical data

Scenario is under development (no updates are available).

5.1.2 Data management in international clinical trials by ECRIN

The D 6.1 (Report on use cases, scenarios, user needs, tools, and interoperability issues for the ECRIN community) has presented basic principles for the integration of *p-medicine* tools into ECRIN clinical trial infrastructure.

All partners in the ECRIN community are individuals and have their own idea how the tools have to be used and handled with. The measurement of the intended purpose and its interoperability is difficult to achieve. With user feedback, interviews and workshops it is at least possible to measure the general acceptance or refusal of the integrated tools.

A lot of interfaces to different recipients as well as different actors are concerned in the integration process. This can produce some distributed problems as well as provide new opportunities. The patient empowerment to interact with clinical trial data can motivate patients to join and to encourage other patients for clinical trials. On the other side it offers the chance to falsify a study by knowing its own data.

For all participating developers a workshop will be scheduled. There will be a discussion about the requirement process, testing process and validation process. Furthermore new findings could be gathered and the resulted modifications and improvements get incorporate into new versions.

The “validation simulation” is the process considering the generics aspects to prepare the validation of developed/envisaged tools, whereas the validation in terms of GCP is the more specific and dedicated instance of the process for one system to demonstrate GCP compliance.

5.1.3 Use of data mining to improve study feasibility

Scenario is under development (no updates are available).

5.1.4 Improved patient recruitment in oncological clinical trials

Scenario is under development (no updates are available).

5.1.5 Use of p-medicine platform with decision support to conduct oncological clinical trials by ECRIN

Scenario is under development (no updates are available).

5.1.6 Increased subject retention rates in oncological clinical trials

Scenario is under development (no updates are available).

6 Education and Patient Empowerment Scenarios

6.1 Education and Training

The D 16.2 (First demonstration of developed flash tutorials and e-learning tools) is a demonstration of the e-learning tools produced by the *p-medicine* project. The educational modules produced show the format and structure of that the final modules will adhere to, with video based, structured interviews giving the partners the opportunity to share their expertise in their specialist area. The modules produced as a demonstration show Prof. Gabriella Pravettoni discussing shared decision making, the modules' target audience are doctors.

This document could be used as a reference for elaborated education scenarios:

- Help patients understand the patient empowerment tool
- Teach health care professionals when to use the *p-medicine* tools
- Impart understanding of the *p-medicine* environment
- Scenario for Education and Training

6.2 Patient Empowerment

The D 14.4 (Implementation of Interactive Empowerment Service (IEmS)) has described the development and validation of the ALGA questionnaire, and its subsequent adoption on the IEmS, The validation was carried out in healthy people and breast cancer patients. The questionnaire is optimised on several smartphone Apps, Pc and iPad and tested on five browsers. An ALGA profiler has been developed in multiple visual forms to enable physicians to rapidly inspect each patient's individual cognitive profile and see at a glance which areas are of concern. With this tool (s)he can modulate the language and vocabulary and content of subsequent discussions with the patient, thus enabling easier understanding by the patient, which in turn helps the patient formulate questions and participate on an equal footing in the decision making processes. Finally a preview is given on the techniques under consideration for optimal customisation systems to be used within the Personal Health Record on the IEmS platform.

The previous scenarios related to patient empowerment services will be updated according the preliminary results of the D 14.4, and in special the scenarios:

- Search for running clinical trials in Europe
- Consent and Re-consent Scenario
- Informed Consent (Patient's Perspective)
- Own Data Scenario
- Summarize the history of the disease in an understandable way and increase patient-doctor understanding

7 Biobank Access, Portal and IT Scenarios

7.1.1 Biobanks Scenarios

Biobanks, sometimes called biorepositories or biomaterial banks represent key resources for clinico-genomic research and advances in personalized medicine. Through the increasing use of molecular and genetic factors in the study of disease causes and targeted individualised therapies biobanks have gained growing importance as part of the scientific infrastructure. The discovery of critical genes and pathways as well as the analysis of their impact and significance will critically depend on sufficient access to biomaterial and related information. In consequence, there is a growing interest in integrating biomaterial repositories into larger infrastructures in order to satisfy research communities' need to specific, quality-assessed samples and up-to-date sample related data with integrated (or at least the possibility to link to) clinical and epidemiological information. A prominent initiative in this direction represents the European Research Infrastructure project BBMRI (Biobanking and Biomolecular Resources Research Infrastructure) whose preparatory phase came to its end in January 2011.

With its ICT infrastructure and contractual framework, *p-medicine* will support researchers' growing demand to access and share high quality biomaterial and related data for their research projects. For this purpose a Biobank Access Framework called p-BioSPRE has been developed within the *p-medicine* platform to enable and simplify access to existing biobanks but also to offer own biomaterial collections to research communities and manage biobank specimens over the ObTiMA Trial Biomaterial Manager. p-BioSPRE is being developed along two underlying user scenarios. The D 10.2 (Initial Implementation of the Biobank Access Framework) describes these user scenarios, and its software modules implemented so far, and special the following scenarios:

- Access to Biobanks Scenario
- Linking the own biomaterial data repository to the *p-medicine* biobank access framework for the collaboration with specific user groups
- Managing patient's biomaterial and related data in clinical trials with ObTiMA
- Offering human biomaterial to a closed and/or open clinical research community for research
- Requesting specific human biomaterial within a closed and/or an open clinical research community for research purposes

7.1.2 p-medicine portal scenario

This scenario has been successfully implemented and reported. The D 8.1.2 (Design and prototype implementation of the *p-medicine* portal) presented why we need a portal solution for the *p-medicine* project and described how the portal for *p-medicine* has been developed and how the users involved into the project can use the portal. The document analyses user requirements for the portal and defines the portal user groups and their roles in the portal. Furthermore, several technical solutions that seem appropriate to fulfil the complex user requirements for the *p-medicine* portal are evaluated and the decision for using the Liferay framework for the *p-medicine* project is explained. Also information is provided about the *p-medicine* services and tools that will be directly or indirectly integrated into the portal as well as the community based architecture of the portal. In addition, the deliverable contains a description of the integration of the *p-medicine* security framework into the portal. Finally, this document presents the current state of the portal with the tools and services that are already accessible through it and how they can be used.

7.1.3 Sync and Push services

The secondary use of data from hospital information systems (HIS) in *p-medicine* are enabled by the so called “HIS Interaction Services”. These services can be divided into “sync” and “push” services.

The sync services allow the reuse of the data that are kept in hospital information systems in running clinical trials in ObTiMA. The push services support data owners to retrieve data from heterogeneous clinical information systems and other biomedical databases and integrate them semantically for further reuse.

A detailed overview of the push and sync services, the relevant data flow and the description of these services is given in the frames of D 8.5 (Release of services for the interaction with Hospital Information Systems)

7.1.4 Data translation for PUSH services

The state of the art review described in the last chapters of the D 4.2 (Requirements for Semantic Access to Clinical Trial Data and HIS) has shown that currently, a variety of semantic resources that describe different aspects of medicine, health care and clinical trials exist for this purpose and are utilized in different medical information systems. However, in our investigation several problems became clear. Firstly, a lot of standards which provide semantics and could improve semantic interoperability are only rarely applied. The few standards which are commonly in use, on the other hand, mostly provide no deep specifications of terms in a machine-readable format. For example, the widely used system ICD-10 is not suitable as a semantic resource for automated reasoning which would be a desideratum for machine-readable semantics. Furthermore, semantic standards for clinical care and clinical research are often developing independently, resulting in a lack of harmonized standards.

The approach in *p-medicine* to solve these problems is to develop a new semantic reference for *p-medicine*’s purposes, the Health Data Ontology Trunk (HDOT). HDOT will not be developed from scratch, but harmonize and link existing semantic resources, which are described in the last chapters, under one umbrella. It especially takes advantage of the fact that it integrates resources, which are widely applied and well-known by medical staff, as e.g. ICD-10. The provision for allowing the use of such terms and codes makes HDOT user-friendly. In particular, HDOT will provide a machine-readable semantic background for these resources. That means that as opposed to many existing resources, HDOT has underlying machine-readable axioms and supports automated reasoning.

HDOT will provide the semantic basis for the data integration and sharing in *p-medicine*. In order to enable the integration of data from different data sources based on the HDOT ontology, the *p-medicine* semantic layer will be developed, that is responsible to resolve semantic interoperability issues. This approach will foster the secondary usage of HIS and CTMS data for various use cases important to realize personalized medicine as e.g. simulations, data mining and the reuse of HIS data in the ontology-based clinical trial management system ObTiMA. To enable such scenarios push and sync services will be developed based on the *p-medicine* semantic layer to support the reuse of the data.

In the following we will give an initial idea of the *p-medicine* semantic layer and describe the push and the sync services. Details of the *p-medicine* semantic layer will be described in deliverable D4.3 taking into consideration the state of the art review presented in this deliverable.

7.1.5 Ontology annotation of external databases

The D 4.3 (Initial release of data integration technology) describes the initial release of the data integration technology in *p-medicine*. The two main components of this technology are the Ontology Annotator tool (a web-based tool for creating database annotations) and the

Data Translation service (a Java-based API for translating external databases into a common format. These two components, together with the HDOT ontology, form the backbone of the data integration schema in *p-medicine*, and are supported by a data warehouse dedicated to the storage of the integrated data. The approach adopted for the data integration technology in *p-medicine* pursues the integration of highly heterogeneous data from many different institutions. The tools developed aim to support this requirement and are the result of extensive studies in the areas of RDF representation, mapping representation and existing data integration technologies.

7.1.6 Ontology-Based Semantic Search Framework

This scenario could be renamed to “Data Mining Framework”.

The D 11.2 (Initial version of data mining tools in the *p-medicine* architecture) has presented the overall data mining framework developed in the context of *p-medicine* and has described the tools it provides. Authors have described how to support large-scale, high performance data mining by means of distributed computation within R and the subgroup discovery service. Moreover, the document suggests that the execution of data mining workflows is integrated in the *p-medicine* security framework. While the integration with the security framework is completed, the question of how to deal with privacy constraints when rules or models found in *p-medicine* are to be published to a wider audience is still open. While the different approaches offered by privacy-preserving data mining, it is still an open question which concrete tools are to be used and which specific privacy-preserving data mining patterns are to be realized in *p-medicine*. These issues will be addressed in the forthcoming deliverables, in particular in D 11.3.

8 ObTiMA Scenarios

Introduction

ObTiMA is developed in a modular fashion with a core module for data management of clinical trials. Some scenarios bellow are under active development, the activities related to ObTiMA's scenarios were documented and presented in the frames of D8.4 (Provision of new modules for clinical trial management).

8.1.1 Pseudonymization Scenario

The goal of the privacy framework is to pseudonymise the data, before delivering it to the *p-medicine* research domain. All data that is passed to the *p-medicine* platform must therefore pass through a state of the art pseudonymisation platform, the data anonymisation tools. The components that make up the data anonymisation tools are described in this software deliverable.

According to the D 8.3 (Release/Demonstration on data anonymisation tools) there are two rounds of pseudonymisation. A first round is performed within the boundaries of the treatment domain. This round consists of two steps. As first step a different pseudonym is assigned to each patient uploaded by the hospital. If available, identifying information on the patient can be uploaded, in encrypted form, to the patient identity management system (PIMS). This allows patient information coming from a different hospital, but belonging to the same real life person, to be linked with each other (second step). This second step results in one unique patient pseudonym over different sources.

Once the first round is finished, the data is passed through a Trusted Third Party (TTP) that performs a second round of pseudonymisation before delivering it to the *p-medicine* research domain. This second round will transform all pseudonyms to a *p-medicine* specific pseudonym. Re-identification of this pseudonym is only possible by passing through the TTP again. Files are delivered to *p-medicine* by uploading them to the *p-medicine* data warehouse.

8.1.2 Data Entry of Prospective Clinical Trial Data

Data Entry of Prospective Clinical Trial Data scenario has been successfully implemented and is reported in the frames of D 8.4 (Provision of new modules for clinical trial management), chapter 5. The scenario described here is the direct "follow-up" of the scenario "eCRF Developer for Prospective Clinical Trials" where the design of electronic CRFs was presented. The module in this scenario now constitutes the implemented solution where CRFs can be filled in right on the screen and sent to the trial center directly over the internet.

The interface of this scenario is directly deduced from how the corresponding question (and answer possibilities) is/are defined during the CRF creation step described there with the look of a CRF when being filled-in for a patient corresponding directly to the look of the CRF when it was (graphically) defined. The difference is that the buttons for adding, deleting or editing questions are replaced with buttons relevant for editing answers for the patient, such as for viewing the answer history of a CRF's question.

8.1.3 Data Manager of Prospective Clinical Trials

Scenario is under development (no updates are available).

8.1.4 eCRF Developer for Prospective Clinical Trials

eCRF Developer for Prospective Clinical Trials scenario has been successfully implemented and reported in the frames of D 8.4 (Provision of new modules for clinical trial management), chapter 3. The task of setting-up or creating eCRFs, is the task described in the base

scenario of this chapter. The core ObTiMA module which allows the creation of CRFs during the phase of setting up and designing a clinical trial and which can subsequently be used for collecting the patient data “on the screen”, as described in the scenario above (Data Entry of Prospective Clinical Trial Data).

8.1.5 Data Synchronization with HIS during running trial in ObTiMA

Scenario is under development (no updates are available).

8.1.6 SAE/SUSAR Scenario

Scenario is under development (no updates are available).

8.1.7 Drug interaction Scenario

Scenario is under development (no updates are available).

8.1.8 DICOM Scenario

Scenario is under development (no updates are available).

8.1.9 Consultation Scenario

Scenario is under development (no updates are available).

8.1.10 Trial Development Scenario

Trial Development scenario has been implemented and reported in the frames of D 8.4 (Provision of new modules for clinical trial management), where the chapter 2 is presenting the elaborated end user interfaces and the basic workflows for trial development and the trial outline builder scenario.

8.1.11 Trial Outline Builder Scenarios

When a trial chairman wants to create a new trial, he/she has to specify first the initial, basic data of the trial, as presented in D 8.4 (Provision of new modules for clinical trial management). Those data comprise, among others, the trial’s acronym and name, its EudraCT number, its start and end date.

After specifying the basic data of the trial, then the Master Protocol Creator guides the user, i.e. in this case the chairman of the trial to be created, through the various steps necessary to create the master protocol of the trial. This comprises a comprehensive description of, among many other topics, the various parts of the trial, such as a documentation of the goals of the trial, the actual procedures and treatments tested within the course of the trial or the legal documentation necessary.

The Master Protocol Creator opens with the start screen where the end-user can select:

- a tour of the editor which highlights its various capabilities and how those capabilities can be used efficiently by the user;
- a protocol wizard which provides the user with an easy-to-use guide to quickly set-up the most common items and the structure of the master protocol; and
- an entry point to the actual Master Protocol Creator where all of the full capabilities of the editor is offered directly to the user.

8.1.12 Participating Centres Scenario

Scenario is under development (no updates are available).

8.1.13 Patient Access to his/her trial data and Diary Scenario

Scenario is under development (no updates are available).

8.1.14 Repository Scenario

Repository scenario has been successfully implemented and reported in chapter 4 (CRF Repository) of the D 8.4 (Provision of new modules for clinical trial management).

Access to the CRF Repository is needed during the creation of a trial: From within the aforementioned CRF Creator, it is possible to transfer and store CRFs from the CRF Creator in the repository as well as retrieving CRFs stored in the repository to possible further elaboration of them in the creator in order to use them in the currently edited trial.

The interface to transfer and store CRFs from the CRF Creator into the Repository is simple and intuitive. When the user clicks on the small basket button which can be found in the general information section for the CRF, then the current state, i.e. the current version, of the CRF is transferred to the repository and stored within there.

To retrieve and then use/edit CRFs stored within the repository, the user clicks on the button “Add CRF from Repository” which can be found below the list of CRFs which have already been added to the given trial.

8.1.15 Semantic Interoperability Scenario

Scenario is under development (no updates are available).

8.1.16 Reporting Scenario

Scenario is under development (no updates are available).

9 DoctorEye Scenarios

Introduction

The DrEye imaging platform - developed in the Contra Cancrum project[3] - is an open access, flexible and easy to use clinical image analysis and simulation platform, for intuitive annotation and segmentation of tumor regions. Its clinically driven design and development is coupled with an open modular architecture focusing on plug-in components. DrEye's main advantage is that the user can quickly and accurately delineate complex areas in multi-modal medical images that can be loaded simultaneously, while multiple labels can be set to allow the user to annotate and manage many different areas of interest in each selected slide. The close collaboration with clinicians in designing the platform has ensured that it can be effectively used in the clinical setting. Another reported feature that adds value to the platform is that it allows computational "in-silico" models of cancer growth and simulation of therapy response to be easily plugged in, in order to provide a future integrated platform for modeling-assisted therapy decision making. The platform also offers comparison tools for assessing the accuracy of simulations via comparison with the actual therapy outcome. DrEye platform is based on the .NET framework architecture and can be used in any Windows-based computer. The graphical interface is based on well-known Microsoft Office applications to ensure a user-friendly environment.

Features List:

- Support for multiple users (with roles and access management)
- View a single DICOM image or a whole series of DICOM images
- Tab interface that allows for multiple series to be opened at once.
- Configuration of DICOM Level and Width for a selected image only or the whole series
- Intuitive navigation/viewing capabilities
- Support for multiple annotations per DICOM image that feature: Label, Color, Types, Opacity, Support for Annotation management (merge, sort, …) and batch editing (rename all, …)
- Powerful annotation tools: Pen, Eraser, Rectangular Marquee, Elliptical Marquee, Boolean operations among ROIs, Magic Wand, Active contours using Greedy algorithm, Active contours using Snakes algorithm, Semi-automatic selection of outer boundaries, and more…
- Metrics (Ruler, Surface estimation and Volume estimation of a selected ROI)
- Histogram generation for multiple ROIs
- More features/functionalities can be added with 3rd party plugins that can be embedded in the platform seamlessly. SDK and guides are available.
- Import/Export in various common formats (comma separated csv files, excel files, text files, xml, …)
- Embedded Viewer for DICOM tags
- 3D visualization of a selected series and of its annotations.

[3] Contra Cancrum website, http://contracancrum.eu [December 2013]

9.1.1 Segmenting Nephroblastoma MRI

The last chapters of the D 9.4 (Report on the segmentation of Wilms tumours using DoctorEye) are presenting in examples the Wilms' tumour segmentation/annotation techniques where the achieved results are presented and reported in details. Additionally this document is presenting the changes after chemotherapy in tumour volume estimated with DoctorEye platform. Authors are comparing the histopathological tumour type to these changes, hypothesizing that the histograms of Wilms' tumour images (MR) might have prognostic and histopathological diagnostic information.

9.1.2 Signal Intensity Scenario

The D 8.2 (Release of tools for multi-modal cancer image analysis and annotation) provided a comprehensive analysis of the technologies that have been developed within the context of the project's task (Task 8.2, Specific image analysis tools development for nephroblastoma and breast cancer applications).

The document aimed to provide the clinician or IT specialist with the basics regarding segmentation of Contrast-Enhanced MRI, quantitative Analysis using Diffusion Weighted Imaging, qualitative Analysis using Perfusion Weighted Imaging and CEMRI pharmacokinetic modelling and image biomarker toolkit. This is essential for the user of these tools in order to have a good idea of their physiological basis.

The deliverable first described the DrEye platform, which was first developed in the Contra Cancrum project and has been therefore enhanced within *p-medicine*. Then it summarises the Segmentation of Wilms' Tumour with DrEye (summary of D 9.4). In the last chapter (CEMRI pharmacokinetic modelling and image biomarker toolkit), authors have described the modules for Functional segmentation of Contrast-Enhanced MRI and Quantitative Analysis using Diffusion Weighted and Perfusion Weighted Imaging.

10 New Scenarios

10.1.1 Annotation of existing data set

The D 4.5 (Ontology Aggregator Prototype) describes the strategy behind and the implementation of the Ontology Aggregator Tool (OAT) up to the release and testing of a prototype version. This web-based tool has been developed in a highly modularized manner to keep it flexible and easily adjustable. It is developed for users who need quick, yet high-quality tailor-made semantic or terminological solutions in cases in which *p-medicine*'s semantic framework is not expressive enough, for instance for providing metadata, data annotations (in conjunction with the Ontology Annotator Tool developed by UPM in task 4.3) or other semantic standardisation purposes. The OAT enables the user to extend *p-medicine*'s semantic framework and compile her own dedicated semantic module. For this a large pool of pre-existing external semantic resources is re-arranged and categorized. Parts of the external resources are taken under the overall semantic structure of *p-medicine*'s ontological framework that is defined by the Health Data Ontology Trunk (HDOT). The OAT can semi-automatically enrich HDOT and its modules by adding concepts, classes or terms from other semantic resources under appropriate HDOT classes on-the-fly. The developers opt for a relatively high degree of automation in this process to make this tool useful for users with very little or no experience in the area of semantic standardisation and ontology. The tool displays necessary information to the user in a concise and understandable manner so that she can make a well-founded ultimate decision on the appropriateness of the generated recommended solutions.

The deliverable describes the development cycles of the tool (in line with 'Annotation of existing data set' scenario), presents the initial testing results and shows how the tool can generate data for its own improvement in use. In addition, the OAT features a user-profile-specific rights management system and an update service to keep HDOT and its modules always aware of the latest developments of important semantic resources and new versions or releases of those.

11 Conclusions

Introduction

The main objective of this deliverable was to identify and to describe the available updates of the user needs and requirements for the proposed for implementation *p-medicine* project scenarios.

This report has been elaborated to fulfil the necessity of a spiral process of requirements analysis, elicitation, documentation and validation described in the related deliverables. The main technique was the continuous monitoring of the elaborated scenarios by underlining the available updates and the scenario related project's activities. This report provided detailed information on end user needs and end user requirements (in the frames of the updated or implemented scenarios).

11.1.1 Concluding Table

The concluding table is similar to the table from the Executive Summary chapter, it presents the summary of actual state of all project's scenarios in line with the available updates of the user needs and requirements.

Table 1: Priority table with implementation status

Nr	Scenario Name	Priority / Implementation Status
1. VPH Scenario		
1.1	VPH Toolbox Scenario	High / Implemented and reported (D 8.1.1)
2. Security Scenarios		
2.1	Single Sign-on Scenario	High / Implemented and reported (D 8.6.2 and D 3.4)
2.2	Single Sign-out Scenario	
2.3	Access Rights Scenario	
2.4	User Enrolment Scenario	
3. Clinical Scenarios - Nephroblastoma		
3.1	Pathway Scenario	Medium / Under development, evaluation or validation
3.2	Imaging Scenario	Medium / Under development, evaluation or validation
3.3	(Severe) Adverse Event ((S)AE) Prediction Scenario	Medium / Under development, evaluation or validation
3.4	Tumour Marker Scenario	Medium / Under development, evaluation or validation
3.5	Oncosimulator Scenario	Medium / Under development, evaluation or validation

4. Clinical Scenarios - Breast Cancer		
4.1	Breast Cancer Scenarios	High / Implemented and reported (D 9.7)
4.2	Oncosimulator Scenario	Medium / Under development, evaluation or validation
5. Clinical Scenarios - Acute Lymphoblastic Leukaemia		
5.1	Oncosimulator Scenario	Medium / Under development, evaluation or validation
6. Clinical Trial Scenarios		
6.1	Statistical analysis of cancer samples with associated gene expression data and clinical data	Medium / Under development, evaluation or validation
6.2	Data management in international clinical trials by ECRIN	High / Implemented and reported (D 8.1.1)
6.3	Use of data mining to improve study feasibility	Medium / Under development, evaluation or validation
6.4	Improved patient recruitment in oncological clinical trials	Medium / Under development, evaluation or validation
6.5	Use of *p-medicine* platform with decision support to conduct oncological clinical trials by ECRIN	Medium / Under development, evaluation or validation
6.6	Increased subject retention rates in oncological clinical trials	Medium / Under development, evaluation or validation
7. Education and Patient Empowerment Scenarios - Education and Training		
7.1	Help patients understand the patient empowerment tool	High / Implemented and reported (D 16.2)
7.2	Teach health care professionals when to use the *p-medicine* tools	
7.3	Impart understanding of the *p-medicine* environment	
7.4	Scenario for Education and Training	
8. Education and Patient Empowerment Scenarios - Patient Empowerment		
8.1	Search for running clinical trials in Europe	High / Implemented and reported (D 14.4)
8.2	Consent and Re-consent Scenario	
8.3	Informed Consent (Patient's Perspective)	
8.4	Own Data Scenario	

8.5	Summarize the history of the disease in an understandable way and increase patient-doctor understanding	
9. Biobank Access, Portal and IT Scenarios		
9.1	Access to Biobanks Scenario	High / Implemented and reported (D 10.2)
9.2	Linking the own biomaterial data repository to the *p-medicine* biobank access framework for the collaboration with specific user groups	
9.3	Managing patient's biomaterial and related data in clinical trials with ObTiMA	
9.4	Offering human biomaterial to a closed and/or open clinical research community for research	
9.5	Requesting specific human biomaterial within a closed and/or an open clinical research community for research purposes	
9.6	*p-medicine* portal scenario	High / Implemented and reported (D 8.1.2)
9.7	Sync and Push services	High / Implemented and reported (D 8.5)
9.8	Data translation for PUSH services	High / Implemented and reported (D 4.2)
9.9	Ontology annotation of external databases	High / Implemented and reported (D 4.2)
9.10	Ontology-Based Semantic Search Framework (this scenario could be renamed to "Data Mining Framework")	High / Implemented and reported (D 11.2)
10. ObTiMA Scenarios		
10.1	Pseudonymization Scenario	High / Implemented and reported (D 8.3)
10.2	Data Entry of Prospective Clinical Trial Data	High / Partially implemented and reported (D 8.4) (some scenarios are under development, evaluation or validation)
10.3	Data Manager of Prospective Clinical Trials	
10.4	eCRF Developer for Prospective Clinical Trials	
10.5	Data Synchronization with HIS during running trial in ObTiMA	
10.6	SAE/SUSAR Scenario	
10.7	Drug interaction Scenario	
10.8	DICOM Scenario	

10.9	Consultation Scenario	
10.10	Trial Development Scenario	
10.11	Trial Outline Builder Scenarios	
10.12	Participating Centres Scenario	
10.13	Patient Access to his/her trial data and Diary Scenario	
10.14	Repository Scenario	
10.15	Semantic Interoperability Scenario	
10.16	Reporting Scenario	
11. DoctorEye Scenarios		
11.1	Segmenting Nephroblastoma MRI	High / Implemented and reported (D 9.4)
11.2	Signal Intensity Scenario	High / Implemented and reported (D 8.2)
12. New Scenario		
12.1	Annotation of existing data set	High / Implemented and reported (D 4.5)

Appendix 1 - Abbreviations and acronyms

AE	Adverse Event
ALL	Acute Lymphoblastic Leukaemia
CDA	Clinical Document Architecture
CDISC	Clinical Data Interchange Standards Consortium
CRO	Clinical Research Organisation
DICOM	Digital Imaging and Communications in Medicine
DSS	Decision Support Service
ECRIN	European Clinical Infrastructure Network
EDC	Electronic Data Capture
EMA	European Medicines Agency
EMR	Electronic Medical Records
ENCCA	European Network for Cancer in Children and Adolescents
EHR	Electronic Health Record
EMA	European Medicines Agency
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GRID	Distributed parallel computing
GUI	Graphical User Interface
HDOT	Health Data Ontology Trunk
HIS	Hospital Information System
LIMS	Laboratory Information Management System
OAT	Ontology Aggregator Tool
PHR	Personal Health Record
SaaS	Software as a service
SAE	Severe Adverse Event
SOA	Service Oriented Architecture

SUSAR	Suspected Unexpected Severe Adverse Reaction
TOB	Trial Outline Builder in ObTiMA