Journal of Biomedical Informatics 62 (2016) 32–47

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

The INTEGRATE project: Delivering solutions for efficient multi-centric clinical research and trials

Haridimos Kondylakis [a,*], Brecht Claerhout [b], Mehta Keyur [c], Lefteris Koumakis [a], Jasper van Leeuwen [f], Kostas Marias [a], David Perez-Rey [d], Kristof De Schepper [b], Manolis Tsiknakis [a,e], Anca Bucur [f]

[a] Computational BioMedicine Laboratory, FORTH-ICS, N. Plastira 100, Heraklion, Greece
[b] Custodix NV, Kortrijksesteenweg 214b3, Sint-Martens-Latem, Belgium
[c] German Breast Group, GBG Forschungs GmbH, Geschaeftsfuehrer: Prof. Dr. med. Gunter von Minckwitz, Handelsregister: Amtsgericht Offenbach, HRB 40477 Sitz der Gesellschaft ist Neu-Isenburg, Germany
[d] Biomedical Informatics Group, DLSIIS & DIA, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo S/N, 28660 Boadilla del Monte, Madrid, Spain
[e] Department of Informatics Engineering, Technological Educational Institute of Crete, Estavromenos 71004, Hearklion, Crete, Greece
[f] PHILIPS Research Europe, High Tech Campus 34, Eindhoven, Netherlands

ARTICLE INFO

Article history:
Received 14 December 2015
Revised 5 May 2016
Accepted 17 May 2016
Available online 17 May 2016

Keywords:
Innovative biomedical applications
Semantic integration
Post-genomic clinical trials
Collaborative environment
Cohort selection
Predictive models
Quantitative analysis

ABSTRACT

The objective of the INTEGRATE project (http://www.fp7-integrate.eu/) that has recently concluded successfully was the development of innovative biomedical applications focused on streamlining the execution of clinical research, on enabling multidisciplinary collaboration, on management and large-scale sharing of multi-level heterogeneous datasets, and on the development of new methodologies and of predictive multi-scale models in cancer.

In this paper, we present the way the INTEGRATE consortium has approached important challenges such as the integration of multi-scale biomedical data in the context of post-genomic clinical trials, the development of predictive models and the implementation of tools to facilitate the efficient execution of postgenomic multi-centric clinical trials in breast cancer.

Furthermore, we provide a number of key "lessons learned" during the process and give directions for further future research and development.

1. Introduction

There is a strong need in medical research, especially in complex, heterogeneous diseases such as cancer, to achieve an all-comprising harmonization of efforts across disciplines, organizations and industries. Wide-scale utilization of clinical data for research is often hindered by fragmentation of methodologies [1], by limited interoperability and adoption of standards [2], and by lack of efficient and easy to use solutions to support collaboration [3]. This results in high financial burden for health consumers and slow transfer of new knowledge and technology into care.

The INTEGRATE project focused on overcoming several key obstacles in oncology research by delivering an environment and solutions to facilitate efficient clinical research and clinical trials in particular in a multi-centric setting, building on experiences of prior efforts and research results in the domain [4]. The project was guided during the development by scenarios and requirements stemming from breast cancer research, provided by the clinical partners in the project. We developed innovative infrastructures to enable data and knowledge sharing, to foster large-scale collaboration, to bring together heterogeneous multi-scale biomedical data generated through standard and novel technologies within post-genomic clinical trials and seamlessly link to

Abbreviations: AC, Access Control; AP, Analysis Platform; BC, Breast Cancer; BRIDG, Biomedical Research Integrated Domain Group; CDM, Common Data Model; CDP, Centre for Data Protection; CIM, Common Information Model; CNS, Central Nervous System; CRP, Central Review for Pathology; eCRF, electronic Clinical Report Form; EHR, Electronic Health Record; ER, Estrogen Receptor; FUH, Frankfurt University Hospital; GBG, German Breast Group; HGNC, database of human gene names; HL7, health level 7; INTEGRATE, project acronym for the project with the full title "Driving Excellence in Integrative Cancer Research through Innovative Biomedical Infrastructures"; IJB, Institute Jules Bordet; ISO, International Organization for Standardization; IT, Information Technology; IdPs, Identity Providers; LOINC, Logical Observation Identifiers Names and Codes; MEDDra, Medical Dictionary for Regulatory Activities; PDP, Policy Decision Point; RIM, Reference Information Model; SIL, Semantic Integration Layer; SUS, System Usability Scale; SSO, Single Sign On; STS, Security Token Service; TTP, Trusted Third Party; XACML, eXtensible Access Control Markup Language.

* Corresponding author.
E-mail address: kondylak@ics.forth.gr (H. Kondylakis).

http://dx.doi.org/10.1016/j.jbi.2016.05.006
1532-0464/

existing research and clinical infrastructures (clinical trial systems, eCRFs, hospital EHRs) to enable a range of innovative applications with high added-value. Essential to the INTEGRATE environment, our semantic interoperability solution leverages existing and widely-adopted standards. The semantic interoperability layer has been implemented based on the HL7 v3 standard and on existing medical ontologies/terminologies: SNOMED-CT, MedDRA and LOINC. The BRIDG standard has been used to represent the clinical trial information in our environment. To be able to reuse previous efforts in data sharing, modeling and knowledge generation and to access relevant external sources of data and knowledge it is beneficial to adhere whenever possible to existing standards. The use of standards will also support the adoption of our solutions.

To facilitate efficient execution of post-genomic multi-centric clinical trials, we focused on the study of issues related to the automatic identification of eligible patients for inclusion into a clinical study, which, based on available scientific evidence [5] remains a key problem in contemporary clinical trials. We proposed a tool, named DECIMA, to support recruitment through the automatic evaluation of the eligibility of patients for trials based on matching the characteristics of the patient population required by the trial to the patient data available for instance in the hospital EHR.

We also developed a tool, named Central Review for Pathology (CRP), focusing on efficient central review of pathology images in order to enhance the collaboration among expert pathologists and enable high-quality decision making.

In addition, to facilitate the use of the datasets available in the environment for future research, we have built NONA, an expressive and intuitive cohort selection application that enables users to define, select, share and collaborate on cohorts of patient datasets that suit their research questions.

Finally, to enable both statistical and prediction analysis within the INTEGRATE environment the Analysis Platform (AP) was implemented supporting multiple project analysis scenarios.

The INTEGRATE technologies can be used by a large and multidisciplinary biomedical community, ranging from basic, translational and clinical researchers to the pharmaceutical industry, to share data and knowledge, and to collaborate on creating new knowledge with the end goal of improving patient outcome. This paper initially describes our methods in Section 2. Then Section 3, presents the key solutions delivered by the INTEGRATE project: Reconfigurable infrastructure components; applications supporting collaboration and the efficient execution of clinical research; a semantic interoperability solution and a standards-based data model focused addressing the needs of breast cancer research and care. We also detail, in Section 4 the technical validation of our technologies and the clinical evaluation of the solutions carried out with the clinical organizations in the project who are vanguards in research and care in oncology. In Section 5 we present related work and finally Section 6 concludes this paper and presents directions for future research.

2. Methods

2.1. Legal requirements

The first objective that needs to be supported by the INTEGRATE computational platform is to enable efficient and effective data sharing across institutional boundaries. Both health and genetic data are collected and stored by the INTEGRATE platform and as such distributed collaboration and research should be enabled and promoted. Such secondary use of health data has a vital role in improving and advancing medical knowledge [6], but it remains essential that steps are taken to respect wishes of the patient regarding secondary usage, and to ensure the privacy of the patient during such use scenarios. Informed consent [7], together with depersonalization and its related concepts of anonymization, pseudonymization, and data minimization are key methods used to provide this protection [8].

However, in principle, the processing of health and genetic data is prohibited, except if the data subject (the patient) has given his/her explicit consent to such processing (Data Protection Directive [9], Article 8, §2, lit. a) or if one of the exemptions provided for in article 8, §§2–5 is met (not applicable to our case). The consent to the processing of sensitive personal data shall be given in writing and in addition, the data controller cannot legitimately allow the processing of sensitive personal data in the absence of explicit written consent of the patient.

The easiest way to run such a platform from a legal perspective would be to use only anonymous data, as the processing of anonymous data does not need a legal basis or an informed consent of the patient, as anonymous data is not personal data and would therefore not fall under the scope of the Data Protection Directive. Unfortunately, most of the data cannot be processed anonymously. In addition, as the identification of each patient has to be guaranteed in order to give the best therapy, most of the data needed for such a platform has to be processed in a pseudonymous way (meaning 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). The INTEGRATE platform should carefully conform to the aforementioned requirements and as such an ethical framework should be constructed.

2.2. High-level architecture

The rationale for the technological developments of the INTEGRATE project lie in the actual IT needs of multicenter clinical trials. Setting up, managing and implementing such trials is an arduous and costly process [10]. There is a strong need to integrate the available data and knowledge in comprehensive models supported by interoperable infrastructures and tools, to standardize methodologies, and to achieve wide-scale data sharing and reuse, and multidisciplinary collaboration. It is therefore becoming critical that several IT solutions need to be in place, while novel concepts are also necessary for facilitating security, recruitment and execution of such trials including quality control [11]. The INTEGRATE project proposed and implemented a technological environment based on the following methodological principles:

- A security central authentication framework with Single-Sign-On (SSO) functionality allowing the user to log in once to all the services of INTEGRATE.
- Advanced tools and services for ensuring privacy protection of patients.
- A semantic interoperability solution for the homogenous representation and assessment of clinical data in order to reduce manual operations and render trial data seamlessly interoperable with the specialized tools of INTEGRATE.
- Advanced, specialized tools for automating patient screening processes, enhancing trial recruitment, facilitating the selection of patient cohorts, allowing the remote collaboration of pathologists for reviewing digital pathology data, perform quality control between participating pathology centers and analyzing data securely on the same platform for performing statistical analyses and predictive modeling.

In addition, it follows the principles of the View – Viewpoint model, as formalized in ANSI/IEEE 1471-2000, ISO/IEC 42010:2007 [12] and leverages several design principles:

- Architectural components have been designed according to the principles of a loosely coupled, open and scalable Service Oriented Architecture (SOA) [13], focusing on the interoperability and interfacing between the different components.
- Architectural components have adopted international standards as much as possible.

According to those principles a reference architecture was designed and implemented. The INTEGRATE reference architecture is result of an iterative interaction process between the clinical end-users, legal representatives and software architects. The architecture is designed as a multi-layered architecture, with responsibilities assigned to the various architectural layers. Every component designed within INTEGRATE can be mapped to/or spanned over one of the following layers:

- The presentation layer is the top layer. The components situated in this layer can be seen as endpoints to the end users of the system, presenting the underlying back-end functionality in an intuitive and user-friendly way.
- The business layer provides the core functionality of the INTEGRATE services as it houses a variety of application services. The components contain the functional algorithms that handle information exchange between the semantic integration layer and the presentation layer.
- The semantic interoperability layer utilizes an ontology based information model and translates or maps the underlying data and information sources to this model, as we shall see in the sequel. The semantic integration layer abstracts the underlying data sources for the upper application layers.
- The data layer contains the various data repositories (e.g. clinical trial protocol information, predictive models information, etc.) and their corresponding metadata (e.g. databases provide their own schema/metadata). Components on this layer are responsible for the actual data access.
- The security layer is typically a vertical layer as security impacts the architecture at all levels.

In Fig. 1 the main components identified in the INTEGRATE architecture are mapped to these layers. In addition the figure gives an overview of how these main components interact.

To help improve the execution of clinical research, the INTEGRATE environment facilitates the collection, preservation, management and reuse of data collected within multi-centric clinical trials by providing an advanced and flexible semantic interoperability solution that leverages widely-adopted standards in healthcare [14]. The datasets are exposed through uniform interfaces to enable efficient information sharing and collaborative knowledge generation. The semantically-interoperable data layer is leveraged by tools that support researchers across domains, institutions and industries to carry out clinical research and jointly contribute to research objectives, develop common methodologies and complex analyses, and efficiently make use of each other's expertise and results.

In the following subsections, a detailed description of the key layers and components of technical infrastructure developed and validated is provided.

3. Results

3.1. The ethical framework overview

In conformity with the legal considerations, within INTEGRATE a Data Protection scheme is created, as shown in Fig. 2, to ensure the compliance of INTEGRATE with data protection regulations. The main parts are the following:

1. *A double pseudonymization procedure:* Before using the data two pseunodymizations take place. The first is performed at the hospital where the hospital holds the link for the second step of de-anonymization and the second takes place on a Trusted Third Party which holds the link for the first step of de-anonymization.
2. *Encryption and the introduction of a central data controller:* The data controller is the person or organization who/which determines the purposes and means of processing personal data. Such data controller has to ensure the compliance with the current data protection legislation. The INTEGRATE partners have decided to delegate this task to the Centre for Data Protection (CDP), a non for profit organization initially established under Belgian Law within the framework of the FP6 EU funded ACGT project [15].
3. *The introduction of a Trusted Third Party (TTP):* A Trusted Third Party (TTP) is a party which at least two other parties trust. In the context of a data protection environment, a Trusted Third Party is regarded as a trustful custodian for personal data or the codes/keys/links that identify the data subject and which shall ensure the privacy of the data subject. It is the duty of the TTP to *de facto* anonymize the patient's data by pseudonymizing the data and to enable the de-anonymization of the pseudonymized data at the same time.
4. *Contracts between the participating hospitals and research-entities with INTEGRATE CDP*
5. *Contract between the TTP and INTEGRATE CDP*
6. *An informed consent of each patient:* This is for ethical reasons and to cover the unlikely scenario where personal data is inadvertently processed.

Using this scheme participating researchers could do their research and analyses in full compliance with data protection laws.

3.2. The INTEGRATE security framework

Trust and security remains a key factor influencing adoption of innovative solutions in the domain of eHealth and biomedicine [16]. Several efforts to respond to this challenge have been reported, ranging from situation-based access control [17] to adopting privacy-by-design principles [18]. The INTEGRATE security framework provides a technological solution that covers the identified INTEGRATE security requirements (for example a contextual attributes [19] requirement) and guarantees compliance of the complete INTEGRATE platform to the legal framework governing the project. It is not shown in the high level architecture since it spans across all layers ensuring the secure access to relevant information.

It consists of modular components, respectively dealing with authentication, authorization, audit and privacy enhancing techniques. It focuses on creating generic, re-useable components that are developed according to the design principles defined in the reference architecture. The developed central security framework (see Fig. 3) contains following main components:

- *Identity Manager:* provides a front-end for user and service identity management and incorporates the modules for identity provision on a technical level, i.e. Identity Providers (IdPs) supporting Web Single Sign-On (SSO) and WS Security Token Service (STS) capabilities.
- *Access Manager:* is a management front-end that enables Access Control (AC) policy authoring (maintained in the Policy Access Point (PAP)) and facilitates the definition of extended attributes (managed by the Policy Information Point (PIP)). Both PAP and PIP can be queried by the central authorization service for evaluating access requests.

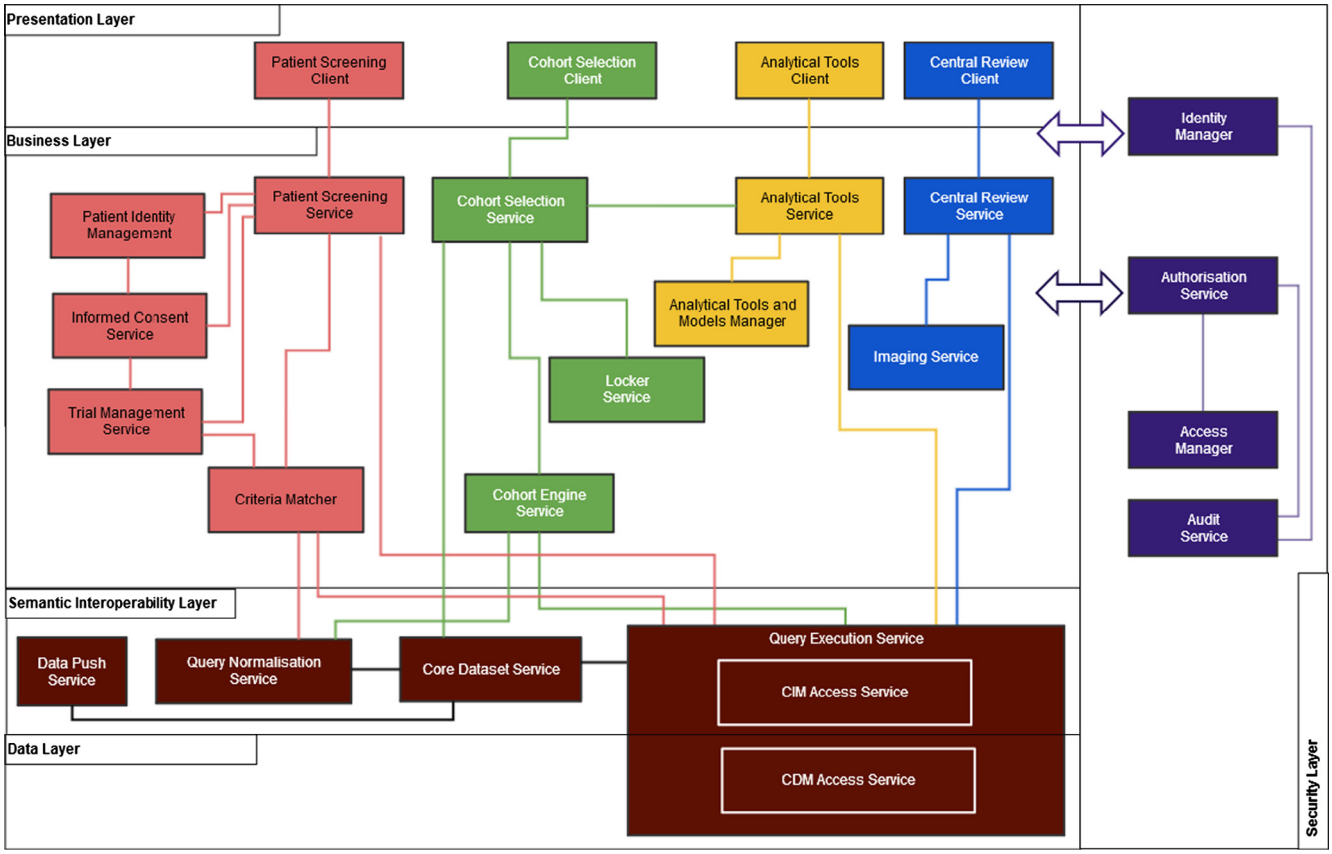

Fig. 1. A high level view of the INTEGRATE technical Architecture.

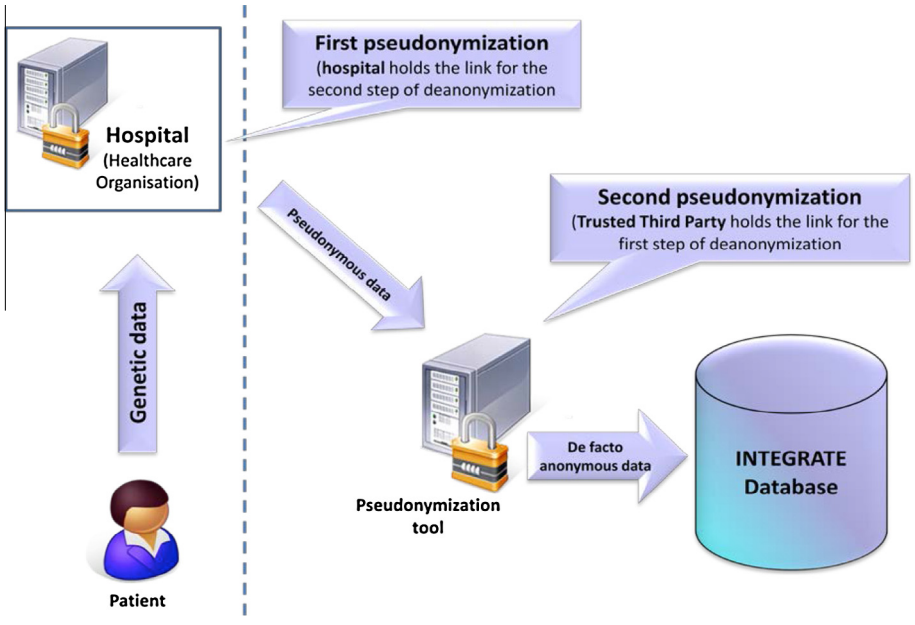

Fig. 2. INTEGRATE data protection approach.

- *Authorization Service:* evaluates access request from all over the INTEGRATE infrastructure based on the security policies defined in the PAP and PIPs. The Policy Decision Point (PDP) is the central engine who finally generates the access decision for each incoming access request and is compliant with the eXtensible Access Control Markup Language (XACML). The architecture and standards used allow this service to be easily implemented as a distributed service (scalability).
- *Audit Service:* provides auditing mechanisms and visualization and is integrated with the authorization and authentication services.

3.3. The INTEGRATE Semantic Interoperability Layer

The purpose of the INTEGRATE Semantic Interoperability Layer (SIL) is to provide a homogeneous access to diverse clinical trial

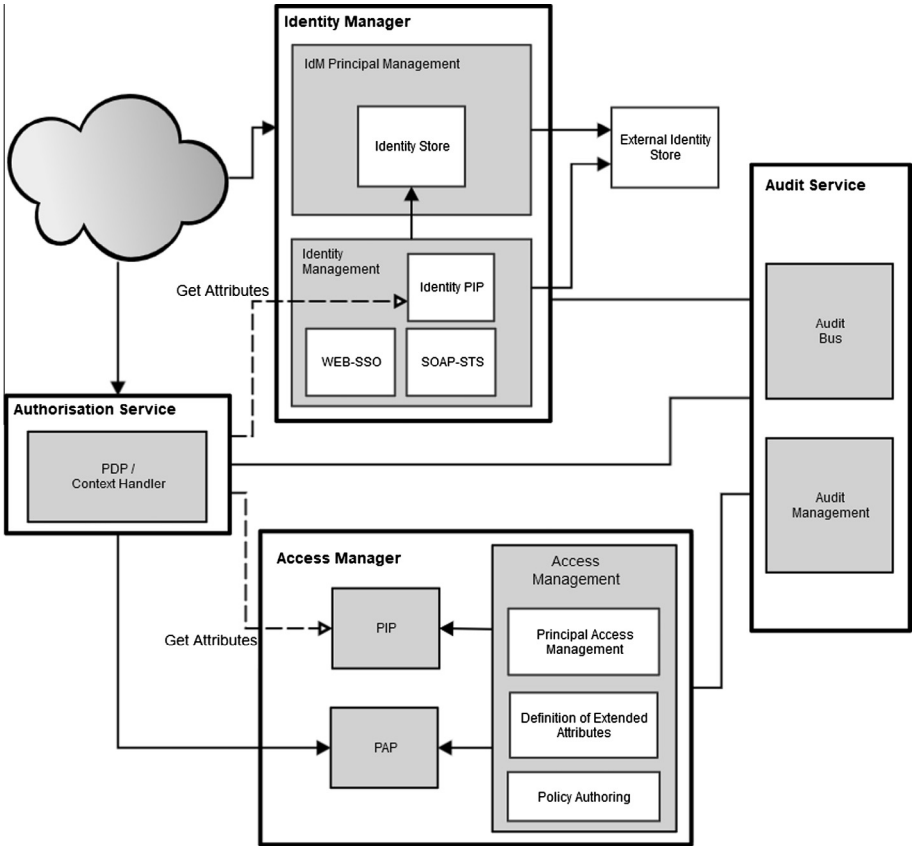

Fig. 3. INTEGRATE Security Framework Overview.

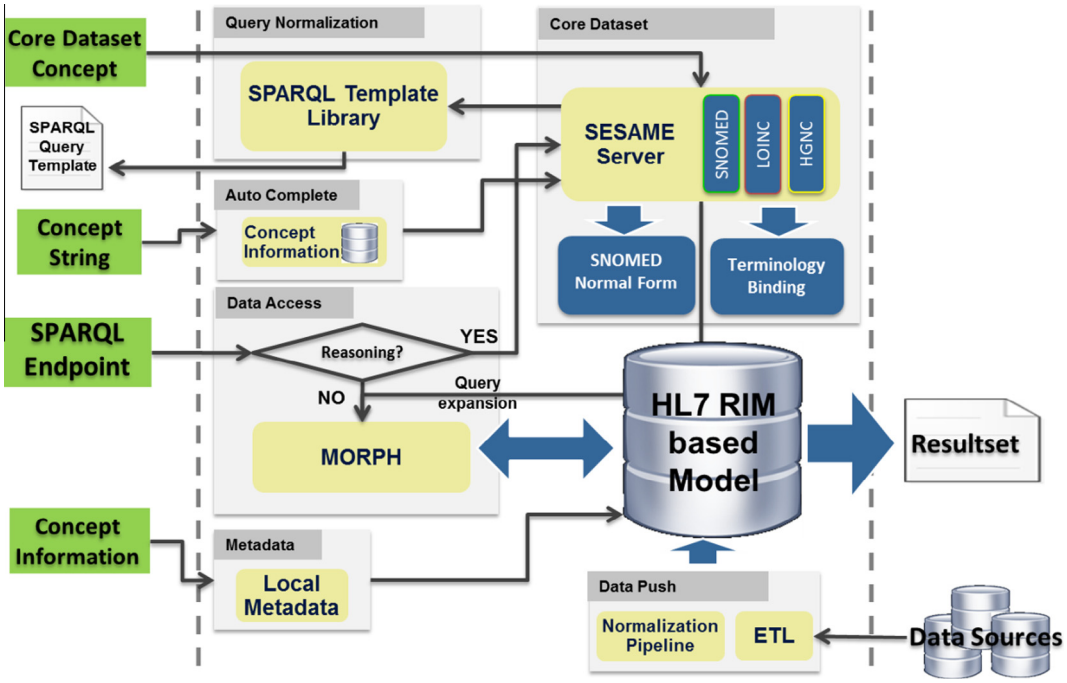

Fig. 4. INTEGRATE Semantic Interoperability Layer architecture.

data [1]. The main components of the SIL, shown in Fig. 4 are the following:

(a) Query oriented services such as the Query Normalization Service.
(b) Access oriented services such as the Query Execution Service.
(c) Binding and semantic reasoning services such as the Core Dataset or the Autocomplete Service.
(d) Data homogenization services such as the Normalization Pipeline.
(e) The relational model component to store data, the Common Data Model (CDM).

The Query Normalization Service includes a Query Template Library, defined by a set of SPARQL query templates based on the CDM. There are five types of possible queries identified: Observation, Procedure, Substance Administration, Entities (devices, products, genes) and Demographic information. The Query Execution Service provides an endpoint to retrieve data from the CDM and through an SPARQL wrapper [20]. Queries can be expanded to include the semantics from the Core Dataset. Once the query is executed, the Query Execution Service returns the results to a SPARQL result set format. The Core Dataset includes a set of standards, vocabularies and relationships among concepts used to populate the INTEGRATE CDM. It is mainly based on SNOMED CT, LOINC and the database of Human Gene Names (HGNC) that covers general clinical terms, laboratory observations and unique names for every known human gene. While the Autocomplete and Metadata Services provide Core Dataset concepts matching strings in GUIs from end-user tools. The Normalization Pipeline is a process to transform clinical data into a common standard to avoid different representations of the same information. Such normalization exploits the SNOMED Normal Form [21] normalization mechanism. Finally, the CDM is the relational scheme of the SIL based on the HL7 Reference Information Model (RIM). HL7 RIM was selected due to the extended representation capabilities for clinical data and to the large number of legacy systems allowing HL7-based exports [22]. Such a large flexibility to represent clinical data is one of the main advantages of HL7 RIM, but it requires normalization mechanisms to reduce representation ambiguities. This issue was solved in INTEGRATE with the Query Normalization Service [23].

3.4. End-user applications

The application layer includes four individual components: *DECIMA*, *CPR*, *NONA* and the *Analysis Platform Tool*. Bellow we present more details on each one of the aforementioned components.

3.4.1. DECIMA – patient screening for clinical trials

Enhancing the participation of patients in clinical trials is an important prerequisite for streamlining the execution of clinical research [24]. The successful completion of a trial is dependent on achieving a significant sample size of patients enrolled into the trial within a limited time period. However, the process of eligibility determination is extremely challenging and time-consuming, often mandating manual chart review [25,26]. Such reviews can involve repeated readings of the patients' Electronic Health Record (EHR) for multiple trials, across every visit. Although institutions participating in clinical trials spend significant resources in conducting eligibility screening, the cost of screening is generally not fully compensated through the contracts supporting the trials [27]. To facilitate recruitment, we developed an application that supports efficient identification of eligible patients for clinical trials through the automatic evaluation of eligibility based on matching the characteristics of the patient population required by the trial to the patient data available for instance in the hospital EHR.

The characteristics of the patients that are considered eligible to be enrolled in a particular trial are described by a set of free-text eligibility criteria. These criteria are both syntactically and semantically complex, which makes their automatic evaluation with the patient data in order to assess the eligibility of the patient for a trial a challenging task. Although there exists a considerable amount of work in the area trying to formalize and represent criteria such as ERGO [28], OCRe [29] or ElixR [30], focusing on focusing on temporal constraints such as [31], clustering trials based on detected in criteria UMLS concepts [32], detecting disease relatedness [33] and analyzing trends [34] still there is no complete solution [36] the automatic formalization of free text of criteria. To automate the evaluation of eligibility of patients for trials it is necessary to (1) extract and represent the semantics of the eligibility criteria in a machine-processable way, (2) define a semantically-enhanced patient data model, and (3) develop an application that automatically matches each criterion with the relevant data available for each patient [35]. We address all these aspects and propose a scalable, efficient and pragmatic application enabling automatic evaluation of eligibility of patients for a relevant set of clinical trials.

The extraction and representation of the semantics of the trial criteria can be made more efficient by detecting structure and patterns in criteria, and by formalization of the criteria as described in [36]. Our solution covers the flexible formalization of criteria and of other relevant trial metadata and the efficient management of these representations [37]. For instance, "No prior chemotherapy", No prior cardiovascular disease", and "No history of cancer" consist of the same pattern – "No prior/No history of"- combined with different concepts. Other examples of frequently occurring patterns are: "Diagnosis of ()", "At least () since ()", "() confirmed by ()", "Planned/scheduled for ()", etc. Each criterion is linked to the relevant patterns (used as templates) and concepts, and to one or more formal representations that will enable the evaluation of the criterion with the patient data. In the current implementation each template corresponds to a script that is instantiated with the relevant concepts (leveraging widely-adopted ontologies such as SNOMED-CT and LOINC) and carries out the evaluation of the criterion on the patient data.

The approach to structuring eligibility criteria by identifying and extracting contextual patterns and ontology concepts is described and evaluated in [38]. We started from the criteria of several hundreds of breast cancer trials and classified them based on subject and structure, manually extracting relevant contextual patterns. The set of patterns was iteratively extended following evaluation of coverage (based on regular expressions) in larger sets of cancer trials, yielding 130 distinct core patterns (that can be combined into more complex patterns). We detected a large overlap in criteria and patterns across trials, which was leveraged by our solution. The evaluations have shown a good coverage for the selected set of patterns and that adding new trials to the Trial Metadata Repository requires little effort, enabling the application to efficiently handle updates.

The approach is generic, integrating loosely coupled components with well-defined standard interfaces and making use of prominent standards in the healthcare domain, such as BRIDG for the Trial Metadata Repository [39] in which we preserve trial descriptions (e.g. eligibility criteria, associated formalisms and other relevant metadata). We combine the use of formal

representations of eligibility criteria with a pragmatic implementation in which templates are defined, linked to execution logic, and extensively reused. The solution can integrate different formalisms, data representations and models to fit the needs of each new healthcare environment.

The matching of criteria to patient data requires shared semantics between criteria and patient data, so underlying mappings need to be developed. We rely on the INTEGRATE standards-based semantic interoperability solution to provide shared semantics between formalized trial information and patient data, in order to support automatic linkage and matching of trial criteria to patient data. Other existing representations of eligibility criteria like ERGO [28], OCRe [29] or ElixR [30], are using the same ontologies used in our solution in their annotation phase.

Fig. 5 depicts the main components of the DECIMA application. The Patient Screening application front-end supports the clinical users in the process of evaluating the eligibility of their patients for the available clinical trials. This component relies on the Trial Metadata Repository to provide the trial descriptions including the formalized criteria, and on the Patient Screening Service to retrieve the template/execution logic of each criterion from the Trial Metadata Repository and the relevant patient data through the Semantic Interoperability Layer (consisting of the Common Information Model and the Common Data Model), and to run the Criteria Matcher. The Criteria Matcher evaluates whether a trial criterion is satisfied by the relevant available patient data. The Trial Metadata Repository manages all the metadata regarding clinical trials and is an essential component for this application. It maintains among others the trial acronym, the trial description, the target accrual range, the number of enrolled subjects, and the eligibility criteria including all the defined formalisms. In the implementation described in [37] the Trial Metadata Repository stores for each criterion the applicable contextual pattern and the matching Groovy script including the queries corresponding to the core dataset concepts present in the criterion.

Fig. 6 shows the application screen that depicts for a patient and a selected trial the value of the match for each criterion. In this screen the application also displays a brief description of the trial and a link to the trial description available on clinicaltrials.gov. Criteria that are not matched by the patient data are depicted with red, green indicates a criterion that is matched by the data, and grey is used for criteria that are unresolved (i.e. that could not be automatically evaluated for instance because no data was available in the patient file addressing this criterion, or because the criterion is "subjective" and requires the assessment of the clinician).

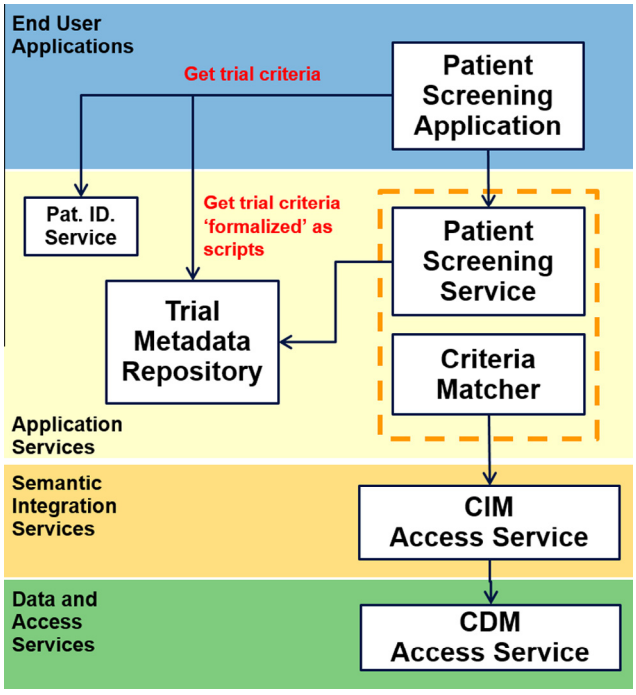

Fig. 5. Main components of the DECIMA application.

For each criterion the user can inspect in this screen the available patient data based on which the matching was carried out (i.e. evidence), and validate or override the assessment of the tool. The user can also decide on and validate the unresolved criteria, which is a feature needed for instance for criteria that rely on the assessment of the clinician. For more information on DECIMA the interested reader is forwarded to recently published papers for DECIMA [37,40].

3.4.2. Central review for pathology (CRP)

Effective and efficient collaboration of clinical trial participants remains an important challenge in the modern multi-centric setting of such, post-genomic trials [15]. The primary use of this Central Review for Pathology tool [41] is to enhance the collaboration among groups of expert pathologists and to enable efficient, high-quality decision making for patients participating in a clinical trial.

Traditionally, microscopy slides were exchanged and pathologists were traveling to perform a central review. Grading criteria for pathology images are complex and have changed over time. Uniform, reproducible grading criteria are critical for clinical trials basing randomization and treatment strategy on these findings [42]. While significant efforts have been made to provide tools to pathologists to ensure accurate uniform grading, these appear to be underused and thus a direct benefit for this intervention could not be measured. Strategies for increasing awareness and use of similar tools in future studies are needed [43].

The advent of digital scanners of microscopy slides, and the ability to exchange the scanned images through digital networks is allowing new modalities for the central review of pathology data. INTEGRATE is leveraging these abilities by the CRP. Using the tool, a panel of expert pathologists is nominated. Then they can annotate the same set of samples together. The CRP provides to the stakeholders all the necessary tools and functionality to support distributed, collaborative image review by multiple experts, to efficient manage the whole procedure and to provide for the required auditing functionality. It enables collaboration and knowledge sharing by providing modules for messaging, scheduling, role and group management, as well as mechanisms for the definition of custom review protocols and expandable business workflows.

By using CRP, the moderator – which is the nominated expert for defining the parameters and conducting the process of a central review – imports digital pathology data into the platform and defines custom review protocol templates [44]; each such a template corresponds to an image or a study type (e.g. definition of parameters for PgR analysis, HER2 analysis etc.). Finally he can register new review protocols by selecting: (a) the analysis types to be conducted, (b) the patients (digital images) to participate into the process and (c) the experts (reviewers) who are going to review these digital data. All the experts who participate into the review process should score and submit their answers to the platform and the moderator is responsible for monitoring the whole process. In the case where there are conflicts between experts' answers he is initiating a process of resolution.

The most important points that the tool addresses are the following: (a) Multiple specialists participate in the review process, having access remotely to the same set of digital pathology data in real time, (b) The qualitative results are greatly improved

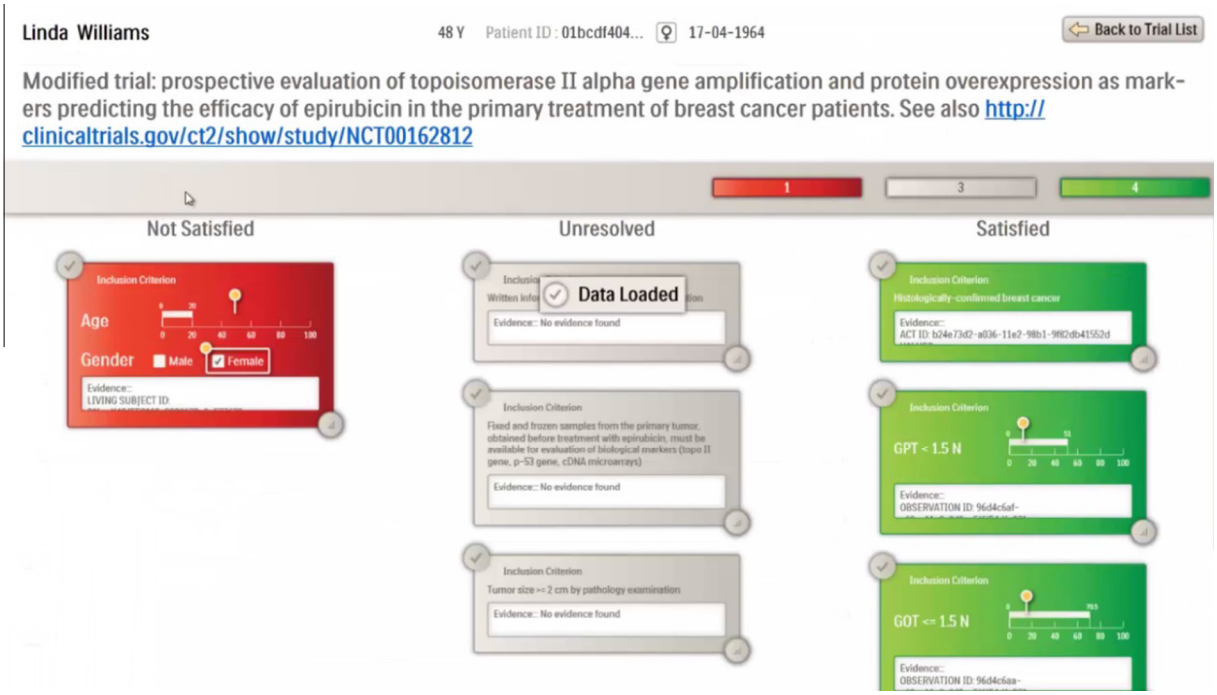

Fig. 6. Patient screening tool: Detailed view of the criteria match for a selected patient.

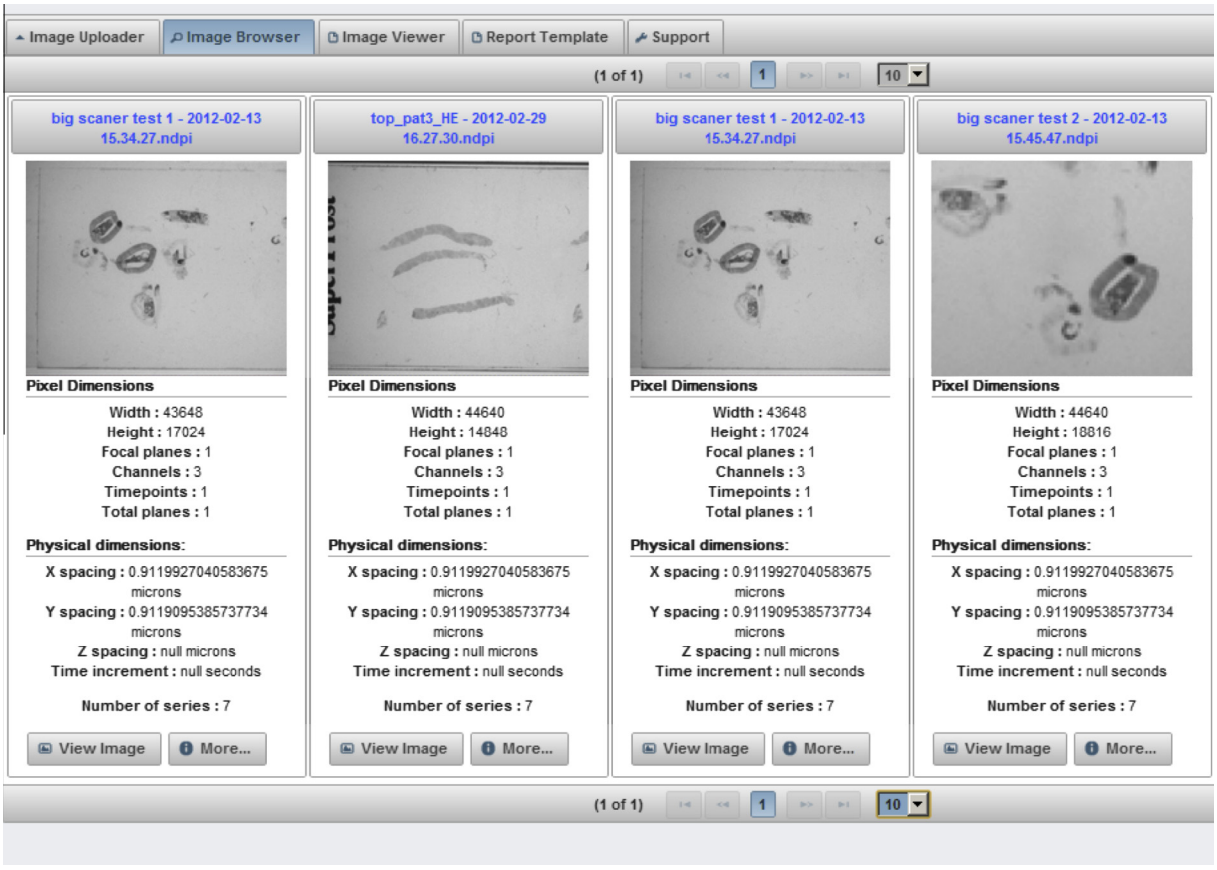

Fig. 7. A typical interface of the CRP tool including the imaging shared workspace.

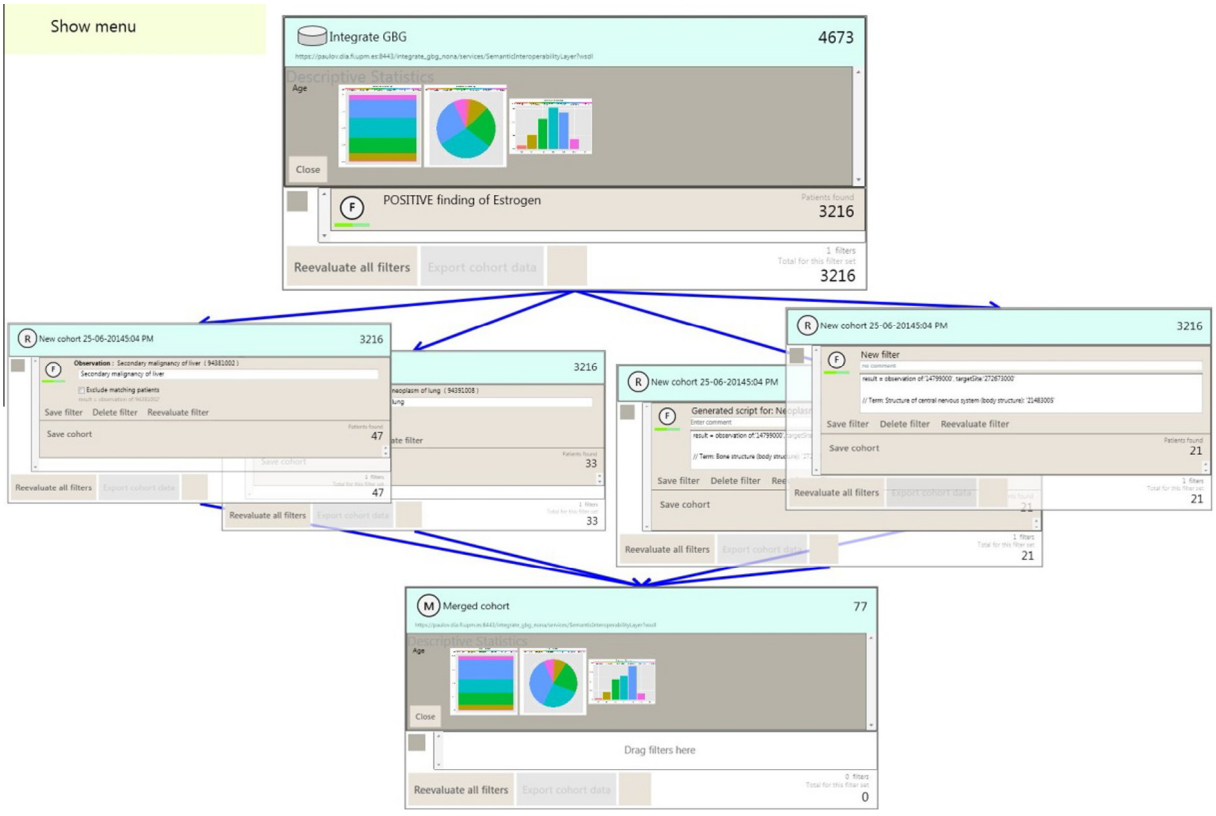

Fig. 8. Distribution of Estrogen Receptor (ER) + Breast Cancer (BC) patients with metastatic disease based on the target site.

through the utilization of knowledge sharing tools, (c) Reduce of the time needed to conduct a review process, (d) The review results are stored in a central repository and then can be easily reused at any time. A typical interface of the CRP tool including the imaging shared workspace is shown in Fig. 7.

3.4.3. NONA – cohort selection

Retrospective use of clinical trial data by clinical researchers covers a wide variety of research investigations including data exploration, quality assessment, generation and retrospective validation of research hypotheses, etc.

While clinicians involved in research hold the clinical knowledge and expertise that would enable them to formulate and test hypotheses based on the exploration and analysis of the collected data, they are often limited by the unavailability of user-friendly and intuitive visual tools giving them easy access to the data. The filtering of datasets to select relevant cohorts of patients (sets of patient data with shared desired characteristics) to be used for further analysis is often time-consuming and requires a lot of manual steps and/or advanced informatics skills (e.g. scripting, querying, etc.).

To streamline the process of exploring and filtering retrospective datasets, we developed a flexible solution that enables clinical users to efficiently and intuitively build patient cohorts with desired characteristics, perform basic analyses and counts, and export these cohorts to be further used in advanced analyses. To facilitate collaboration we also provide functionality for sharing both the filter stacks and the corresponding cohorts among teams of end-users. Advanced features allowing for the splitting and merging of defined filter stacks are also implemented.

The application provides two mechanisms of interaction between the clinical researcher and the underlying data repositories. Predefined (visual) templates can be used for building complex cohorts by clinical users without any need for informatics knowledge. For the power-users (e.g. bioinformaticians, biostatisticians) we also provide the possibility to build cohorts by advanced scripting and by directly sending queries to the data repository. By relying on existing widely-used ontologies (e.g. SNOMED-CT, LOINC) the solution incorporates reasoning in the process of building filters and allows the clinical users to explore the concepts available in each dataset and their hierarchies, make use of synonyms, etc.

Fig. 8 depicts an example of use of the application. First a dataset is selected, a cohort of ER + breast cancer patients is built, and the age distribution is inspected. Then filters are applied to select metastatic patients per target site (liver, lung, bone and CNS). Finally the four cohorts are merged and the age distribution is again visualized.

3.4.4. AP – Analysis Platform

The Analysis Platform [45] is a web-based, collaborative environment that enables the qualitative analysis of large multi-level clinical datasets and the derivation of predictive models. Its development was based on prior related results [46].

The tool supports the following types of analysis defined by multiple scenarios within the project:

- Descriptive statistics for rapidly assessing the variability, dependency and the distribution of certain clinical characteristics across patient population.
- Comparison tests and evaluation of the response rate of different examined regimens when applied to a certain patient population.
- Identification of specific clinical parameters that are surrogate markers for survival, involving the modeling of time to event data in survival analysis.

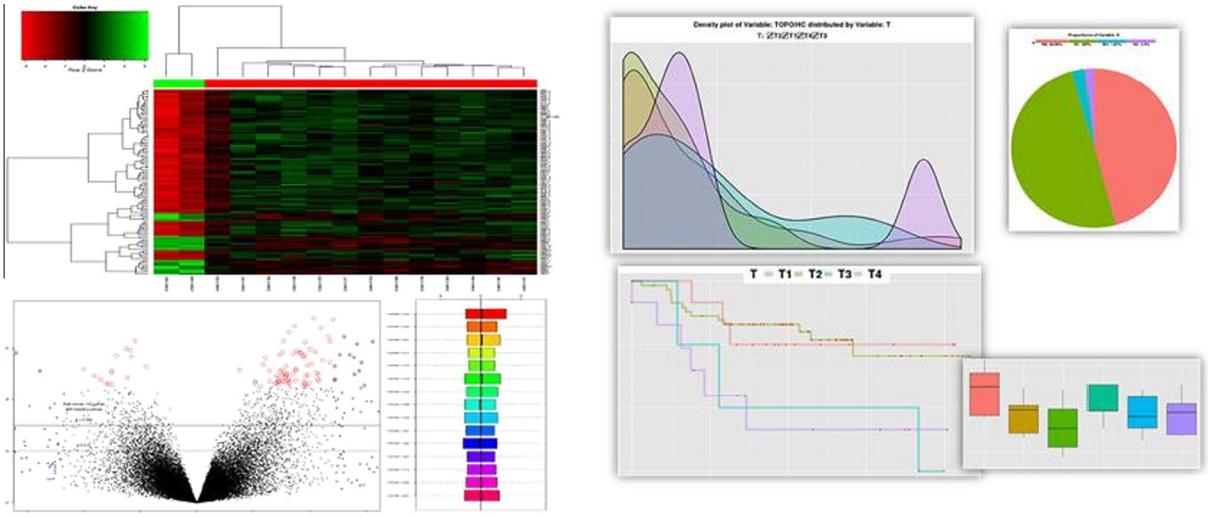

Fig. 9. Indicative (a) genomic analysis and (b) descriptive statistics results.

- Quality control tests to the genomic data, identifying statistically significant genomic information that discriminates subpopulations (i.e. patients achieved pathological complete response versus patients who didn't), and apply unsupervised learning techniques to the entire genomic information.
- Assisted predictive analysis model when homogeneous data (i.e. gene expression) is used for building, running and evaluating the predictive efficacy of the model.
- Heterogeneous integration modeling framework where multimodal data are fused for the development of multi-scale models for predicting drug response, and assessing candidate biomarkers.

The statistical and predictive modeling analysis software scripts are implemented in R language and use publicly available libraries. The platform also integrates an engine for creating dynamically statistical and predictive analysis reports by enabling integration of R code and Latex documentation while it is equipped with an internal database that stores all the metadata information for every executed analysis (e.g. type of the analysis, selected variables used for the analysis, execution time, status of the analysis etc.).

The front-end of the Analysis platform assists users in interacting with the platform's functionality in a user friendly manner. A general pipeline workflow using the platform is as follows:

- The platform authenticates the user with the provided credentials and interacts with the Common Information Model (CIM) to retrieve the data.
- Data are then displayed in tabular format and a filtering functionality allows the user to constrain a request by obtaining subpopulations and build cohorts based on specific ranges of values.
- For the selected cohort multiple tools or models can be scheduled for execution in a single step.
- The layout of the platform communicates with the back-end functionality and the required software, and the overall analysis workflow is presented in a functional diagram format.
- A table with metadata information for each completed or pending analysis is displayed to the user. Additionally the user can view, edit or compare the reports of completed analyses.

The platform supports descriptive statistics for rapidly assessing the variability, dependency and the distribution of certain clinical characteristics across patient population, comparison tests and the evaluation of the response rate of different examined regimens when applied to a certain patient population. In addition, it allows the identification of specific clinical parameters that are surrogate markers for survival, involving the modeling of time to event data in survival analysis. Quality control tests to the genomic data are also supported, identifying statistically significant genomic information that discriminates subpopulations, and unsupervised learning techniques as well for the entire genomic information. Furthermore assisted predictive analysis models are available when homogeneous data (i.e. gene expression) are used for building, running and evaluating the predictive efficacy of the model. Finally it allows the fusion of multi-modal data for the development of multi-scale models for predicting drug response, and assessing candidate biomarkers. Fig. 9 shows (a) some indicative genomic analysis results showing a heat-map on top, a volcano map and a quality assessment plot and (b) descriptive statistics result outputs showing density plots (top), a pie chart (top right), survival analysis plots (bottom left) and boxplots.

The major advantage of the Analysis framework is that brings all the functionality needed for biomarker selection and model testing within a single easy to use framework that doesn't require knowledge of any specific software environment (e.g. Matlab, R) while offering all the functionality needed within a user-friendly interface allowing even non-experts to perform statistical analysis and modeling tasks on their data. To achieve this goal, the programming aspects of the different environments and languages adopted for implementing the framework's facilities, and the connectivity process which allows the interaction between these components are kept at the back-end of the framework, hiding the complexities of the computational infrastructure.

4. Evaluation

Evaluation and validation of software product is vital to both the acquisition and development of it. The importance of the various characteristics of software quality depends on the intended usage or objectives of the system. INTEGRATE followed standardized software development practices and addressed the quality issues during the whole life cycle of the software in order to reduce the likelihood of defects and the cost incurred by them both to users and to developers.

For quality assurance, norms defined from the International Organization for Standardization (ISO) such as the Software Product Quality Requirements and evaluation [12] (SQUARE) has been used as a reference model. For each tool, the applicable functional and non-functional requirements according to ISO/IEC 25023 [47] had been defined in the early stages of the project and monitored by the developers during the whole software life-cycle.

For the evaluation of the tools, the quality characteristics from the product quality model of the ISO/IEC 25000 series along with the System Usability Scale (SUS) for global assessment of systems usability were used. At the evaluation phase different type of users, such as physicians, system developers and statisticians participated. Having such a diverse target group of evaluators, the evaluation questions had to be simple, accurate, easy to understand (especially for non IT experts), non-time consuming and without loss of functionality/quality. For that reason the crucial subcharacteristics of software quality measures from ISO/IEC 25000 series have been translated into simple questions in natural language. The evaluation form of each tool was a list of such questions where the evaluator had to answer with a degree of satisfaction with Likert scale [48]. The generic evaluation form which has been adopted and extended to the needs of each tool developed by the INTEGRATE project can be found in the Supplementary Material Table 1. We also used the System Usability Scale [49] (SUS) for global assessment of systems usability.

The four tools implemented by the INTEGRATE consortium, aim to support tasks of users in patient care and in clinical research. In the patient care setting, DECIMA supports the screening of patients on eligibility for clinical trials and the Central Review for Pathology (CRP) images tool assists pathologists on structured review and analysis of pathology images, and resolution of conflicting results. In the clinical research setting, NONA supports cohort selection and the analysis platform tool (AP) performs statistical and bioinformatics analysis of clinic-genomic data. The tools underwent two evaluations at different time points and geographic locations. All the evaluations of the four tools were quantitative and qualitative with an exception of NONA which underwent only qualitative evaluation since the current conceptual prototype at this time point was not in the technology readiness level for a first version that would allow live and quantitative evaluation.

The first evaluation workshop held at Heraklion - Greece, in June 2014. A panel of experts coming from around the world and representing potential end-users of these tools were invited to participate to the event. In order to gather opinions coming from different environments and working realities, the proposed end-users were identified in several different European and non-European cancer centres. The panel comprised of 2 oncologists, 1 research nurse, 3 bioinformaticians and 2 pathologists. Specific evaluation scenarios and training sessions conducted before the evaluation process, while all the end users signed consent forms.

The second evaluation took place at workshops at the pilot sites, i.e. the Frankfurt University Hospital (FUH), the Maastro Clinic (Maastro), the German Breast Group premises (GBG) and the Institute Jules Bordet (IJB) during October 2014. GBG and IJB were partners of INTEGRATE while FUH and Maastro were external evaluators of the INTEGRATE tools. These evaluation workshops tested ad-hoc solutions that were implemented from prior users' remarks and opinions. It was also a new possibility to confront the tools to different users, and explore their future deployment and possible exploitation.

The DECIMA tool was evaluated at the Heraklion workshop with external users (not participating at the INTEGRATE consortium) and at Maastro, IJB, and FUH. The setup for the evaluation at the hospitals (Maastro, IJB, and FUH) was similar. DECIMA was installed on a laptop with internet access, while the patient and trial data were stored anonymized in the INTEGRATE distributed testbed. The users had access to the DECIMA tool and to data extracted from the local EHR system in the hospital and exposed through the SIL. In each session, the evaluator was introduced to the INTEGRATE project and given a small tutorial on the DECIMA tool. Then, the evaluator was requested to execute specific tasks in DECIMA and to provide feedback about the workflow, the performance and the usability. At FHU and Maastro the time needed for the decision was also recorded. The start time was the moment when data inspection started, either in DECIMA or the local system. The end time was the moment the decision was taken (i.e. pressing the enroll button, or verbally announcing that the patient is excluded). Timing was measured manually, with the accuracy of 5 s. Switchover to new data requires quite some time; this is excluded from the measured per-patient time.

At Maastro the evaluation was conducted with a trial nurse and consisted of screening a large number of patients and recording the timings and trial assignments. The evaluation was conducted in two 4-h sessions, on two days. Screening performance has been investigated at Maastro before [50]. The historical performance for patients not eligible is 120 s, while eligible patients require on average 480 s.

These numbers are the benchmarks that we will refer to. For the evaluation, 270 patient records were checked against 13 trials with in total 210 criteria. The screening time in our evaluation using DECIMA for the patients not eligible is on average 120 s as well. Out of a total of 64 eligible patients six were assigned to a trial. The screening time for the eligible patients is on average 373 s improving the time required here by 23%. We have to note that (a) the users of the system were inexperienced and we expect that as they will gain experience the time will greatly improve and (b) the time required to retrieve the data for one patient is about 20–60 s which we expect will soon be reduced by future refined implementations. Even though this test does not give evidence for a strong numerical improvement, the potential for improving the screening performance is recognized by the user.

At FHU one trial coordinator and one trial nurse familiar with the details and criteria of trials participated with a cohort of 258 patient records against 5 trials and 60 criteria. The service performance was in line with our expectations from the current system (about 10–30 s to retrieve the data for one patient). The amount of time spent considering each patient for trials is fairly consistent and around approximately 300 s. The evaluation at IJB contained 29 patients and 4 trials for screening with a total of 17 criteria. Evaluators were two research nurses part of the multidisciplinary recruitment/screening team and one administrative assistant having less knowledge of the patient screening domain and the evaluation period was two days. This evaluation did not include timing results due to the lack of trial knowledge of the participants. At IJB evaluation sessions were generally focusing on usability of the tool and work flow discussions.

The patient screening workflows on the three pilot sites show considerable differences. They range from a single user at a single decision point, without the chance to collect extra information on the patient, to a more comprehensive process where multiple users review the patient in multiple steps and information is continuously updated. Despite these differences, DECIMA was seen to have potential to improve the screening process; while some additional work is needed to close the gap between the prototype and a fully-deployable system. Overall the end-users evaluated positively the software quality of DECIMA. One of the evaluators pointed out that DECIMA was not specifically tested to run concurrently with many applications, and suggested that the underlying systems should be tested for this, if that is a desired feature. It was also suggested that more testing might be necessary with respect to fault tolerance. For the evaluation the users answered the SUS score questionnaire and a questionnaire focusing to the main software characteristics

Table 1
Qualitative evaluation results of DECIMA tool.

	Maastro1	FUH1	FUH2	IJB1	IJB2	WorkShop1	WorkShop2	Workshop3
Functionality	3.00	5.00	4.00	3.00	5.00	5.00	5.00	5.00
Efficiency	2.33	4.33	4.00	3.67	5.00	4.67	5.00	3.33
Usability	5.00	4.50	5.00	4.00	5.00	5.00	5.00	5.00
Reliability	2.00	4.00	2.00	3.00	5.00	4.00	4.00	5.00
Security	4.00	3.50	4.00	3.50	5.00	5.00	5.00	5.00
Quality in Use	2.67	5.00	4.33	4.67	4.50	5.00	5.00	4.67
SUS score	72.5	82.5	92.5	75	95	100	92	95

from the ISO2500 series. Specifically the ends users had to answer one question for functionality, three questions for efficiency, two for usability, two for reliability, two for security and three for quality in use. Table 1 shows the evaluation results of the DECIMA tool. Each row represents one software quality category while the last row shows the SUS scores. Each column represents the evaluation of an end user and the label indicates the place of the evaluation, e.g. FUH1 is the first evaluator from Frankfurt University Hospital. The values are the average of the answers from one evaluator for the specific category. The evaluator had to answer using a satisfaction degree on the Likert scale with values ranging from 1 (strongly disagree) to 5 (strongly agree). Values over three represent high level of the specific software characteristic and are highlighted in the table with green, values between 2.5 and 3 are in low risk and are highlighted in the table with orange and values below 2.5 are considered high risk and are highlighted in the table with red.

The second tool is the Central Review for Pathology platform and aims to evaluate the process of reviewing digital pathology images from one or more groups of specialists. The first evaluation (workshop at Heraklion) was performed with two pathologists. A Surgical Pathologist and a Cancer researcher from IJB. In general, the participants had no difficulty in understanding on how to use the platform for accomplishing their tasks. The users confirmed that the tool can be useful in processes where multiple users are located in different parts of the globe and have to work together on reviewing the same set of digital pathology images. However the evaluators expressed some concerns in respect of platform's usage in routine practice where it is particularly important that patient data (and their images) are stored and managed in a central, secured and isolated infrastructure (of a research center or a hospital), and integrated with the Lab Information System, especially regarding the sample traceability. The qualitative evaluation results for the Central Review for Pathology platform are shown in Table 2. Similarly to Table 1 in Table 2 we have the average scores for specific software characteristic categories per evaluator and we use the same color coding. The last row shows the SUS scores of each evaluator.

The third tool, the Analysis Platform is designed to be used by oncologists with at least some experience in clinical trial data analysis and by other researchers such as bioinformaticians and biostatisticians. The first evaluation was performed with one oncologist and one bioinformatician. The second evaluation was performed with three clinicians at the IJB. In general the comments from the users and the overall scoring indicate that the platform provides a fast and easy to-use solution for performing specific statistical and predictive analysis, as addressed from the user needs of INTEGRATE, under a consistent and secure framework. The experts made also some very useful remarks regarding the functionality of the platform such as the scale of the survival curves should be between 0 and 100 irrespective of the data, to allow comparison between curves, and easier publication. Also, the distinction between direct and scheduled data retrieval did not seem to be clear enough. In general, as derived by the evaluation forms results, the Analysis tool is of good quality. However usability was negatively affected by some incomplete functions. The end users agreed that optimization work on the prototype is required in order to be moved to actual clinical deployment. The qualitative evaluation results for the Analysis platform are shown in Table 3. Similarly to Table 1 in Table 3 we have the average scores for specific software characteristic categories per evaluator and we use the same color coding. The last row shows the SUS scores of each evaluator.

The NONA tool was evaluated at GBG and IJB. The target users are researchers who deal with cohort selection as part of their daily work. In clinical practice, this turns out to be two individual work contexts: one where a cohort is selected and the detailed data are taken to a detailed study, and trial feasibility where a cohort size is estimated as preparation for a clinical trial. The first evaluation at Crete focused on the first type of cohort selection while the evaluation at pilot sites focused on the trial feasibility. NONA was evaluated by a senior bioinformatician at GBG, a senior Project Manager, a Translational Research Coordinator, a Clinical Fellow in the clinical trials data management Unit, a senior statistician and an epidemiologist from the Statistics and Epidemiology data center at IJB. All the end users agreed that there are established tools and workflows for trial feasibility, and a new tool will find itself in direct competition with these. There are still situations, mainly centered on data sharing and multisite cooperation that are not covered by existing tooling. This is where NONA can make a contribution. User-friendliness and accessibility are other opportunities for NONA. These can enable a wider audience to perform cohort selection, and allow more streamlined workflows for non-

Table 2
Qualitative evaluation of the Central Review for Pathology.

	Reviewer mode			Moderator mode		
	User1	Workshop1	Workshop2	User1	Workshop1	Workshop2
Functionality	4.40	4.40	4.40	5.00	4.57	5.00
Efficiency	5.00	5.00	3.00	4.00	5.00	3.00
Usability	4.50	5.00	5.00	5.00	4.50	5.00
Reliability	3.50	4.00	4.00	4.00	5.00	4.50
Security	3.00	5.00	5.00	3.00	5.00	
Quality in use	4.50	4.25	4.00	5.00	4.00	4.50
SUS score	92.5	87.5	67.5	87.5	82.5	75

Table 3
Qualitative evaluation of the INTEGRATE Analysis platform.

	IJB1	IJB2	IJB3	Workshop oncologist	Workshop bioinf.
Functionality	4.50	4.00	3.50	5.00	5.00
Efficiency	5.00	3.67	4.33	5.00	5.00
Usability	5.00	4.50	4.50	4.75	4.00
Reliability	4.00	4.00	5.00	5.00	5.00
Security	5.00	3.00	3.00	5.00	3.00
Quality in use	5.00	3.13	1.88	4.38	3.38
SUS score	97.5	75	65	65	75

specialists. The feature requests from the end users provided valuable input for closing the gap between the current conceptual prototype, and a first version that will allow live and quantitative evaluations.

For the INTEGRATE tools, the evaluation workshop has resulted in valuable feedback on the current state and in clear directions for the future development and deployment. The experts confirmed the relevance of the use cases for which the four tools were developed and overall gave a positive feedback regarding the tools' utility and functionality. The SUS usability scores ranged from 65 to 90 on a scale of 0–100. The average SUS score from published studies has been measured by Sauro et al. [51] as the 62.1 but as a gold standard is often used the 68. A SUS score above 68 would be considered above average and anything below 68 is below average. This indicating that the INTEGRATE tools reach an acceptable level, but that there is still plenty of space for improvement. This evaluation workshop also resulted in valuable suggestions for their improvement and their forthcoming development.

Overall the evaluations show the relevance of the tools in daily practice, and the added value of the INTEGRATE approach to create tools for clinical end users. As witnessed by the SUS scores, the tools have been of high usability, while requiring modest development effort. The evaluations also show that engineering and optimization work remains in order to allow a clinical deployment of the INTEGRATE framework and its tools. This includes issues on user interface, additional features requested by the evaluators, customizations to allow proper integration in the everyday workflow of users, service performance and stability.

5. Related work

There have been several approaches so far trying to provide solutions for the efficient collection, integration and management of multi-centric clinical trials offering analytical tools and integrated environments.

From the US the caBIG project [52], launched by the National Cancer Institute, was aiming to create a virtual network of interconnected data, individuals and organizations that collaborate in order to redefine the way that cancer research is conducted. The project created several tools to this direction but eventually was discontinued due to many problems. For example the solutions were technology based and not designed around the needs of clinical and researcher's needs, legislative barriers were not tackled and tried to address all problems in clinical and basic research (see [53] for an overview). INTEGRATE on the other hand involved clinicians and researchers in all phases from the design to evaluation, produced clear legal results and tried to address specific questions for multi-centric clinical trials.

At the same time within Europe the ACGT project [54] tried to provide a novel infrastructure for advancing clinico-genomic trials on cancer whereas the recently finished p-Medicine project [55] tried to integrate trial data focusing however on personalized services [56] and patient empowerment [57,58]. The INTEGRATE partners capitalizing experiences from the ACGT project created novel solutions and technologies building and exploiting previous results whereas they shared the semantic interoperability solution with the p-Medicine project proving the wide applicability of the INTEGRATE's developed solution.

On the commercial front, solutions like the IBM Cognos [59] and Microsoft's Caradigm Intelligence [60] platforms provide business intelligence services to pharmaceutical and life sciences companies conducting clinical trials. However, although they manage and integrate large amounts of data they do not deal with multi-centric trials across institutions and national borders.

Besides these projects other initiatives tried to develop clinical research data warehouses such as tranSMART and I2B2. I2B2 [61] has the goal to provide clinical investigators with the software tools necessary to collect and manage project-related clinical research data in the genomics age as a cohesive entity. tranSMART [62] on the other hand is a knowledge management platform that enables scientists to develop and refine research hypotheses by investigating correlations between genetic and phenotypic data, and assessing their analytical results in the context of published literature and other work. However, both these platforms are too generic not focusing to the individual characteristics of the cancer disease. In addition the INTEGRATE platform offers solutions to all levels from the ethical, security, interoperability and integration to patient screening, cohort selection, collaborative pathology review and analysis.

6. Discussion and conclusions

Lack of interoperable, integrated solutions [63] cost fortunes to the translation industry. The INTEGRATE project focused on building informatics solutions that help improve the efficiency of clinical research in oncology and the data and knowledge flow between clinical research and clinical care. This integrated effort was designed to provide real-life solution to multi-centric clinical trial integration with the goal to facilitate all the necessary processes involved. Significant needs to streamline the execution of clinical research and to speed up the transfer of results to clinical practice were identified. Such needs actually represent large opportunities for software solutions focused on providing support for the efficient identification of suitable patients for clinical trials, for collaboratively carrying out research in the multi-centric setting, and for retrospective or prospective data analysis and hypotheses generation. The solutions proposed by INTEGRATE successfully address all these scenarios. Fig. 10 maps the key applications developed by INTEGRATE on the relevant clinical research workflows and depicts where they can contribute to closing the loop between clinical research and practice in oncology.

The tools and services developed in INTEGRATE focus on providing support in the multi-site distributed setting in which consortia of organizations jointly carry out translational research.

To deal with the inherent sources of heterogeneity in a large multi-centric context (e.g. different clinical information systems and infrastructures, different languages and countries, different terminologies/ontologies and information models) first an appropriate ethical and security framework was needed for allowing data processing and usage across borders and institutes. A key lesson here is that compliance to trial related legislation, especially to the data protection laws, is a critical success factor for any research-network. INTEGRATE created a Data Protection Scheme relying on double pseudonymization, encryption and introduction of a central data controller and a trusted third party,

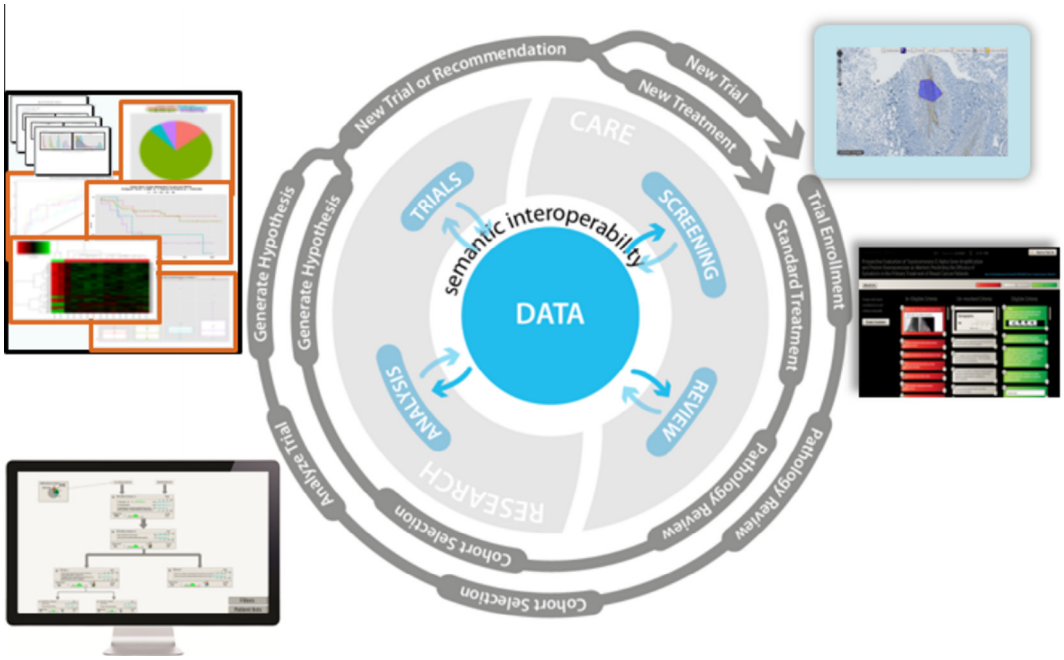

Fig. 10. Mapping the key applications developed by INTEGRATE on the relevant clinical research workflows.

contracts between the participating entities and patient informed consents. The INTEGRATE security framework covers all identified requirements through the developed mechanism for identity and access management, authorization and auditing.

One of the key contributions of the INTEGRATE project was the SIL leveraging widely adopted healthcare standards and ontologies/terminologies. The tools on the application layer rely on the adoption and deployment of the SIL. The effort needed to deploy it at new clinical sites and to extend to new scenarios and clinical domains impacts the success of the INTEGRATE approach. To assess the feasibility and scalability of the INTEGRATE approach the coverage and suitability of existing ontologies was evaluated while at the same time the effort of deploying the INTEGRATE environment at a new site was estimated. The sizes of the sets of relevant concepts and their relative frequency to estimate the cost of data transformation, of building the necessary semantic mappings, and of extending the solution to new domains, were evaluated and compared. The analysis showed that the proposed approach is both feasible and scalable [64]. During the analysis of several widely-used medical ontologies in the clinical domain such as the SNOMED-CT, the LOINC and the MedDRA, on their suitability to capture the semantics of the clinical trial eligibility criteria, of the clinical trial data (e.g., Clinical Report Forms), and of the corresponding patient record data that would enable the automatic identification of eligible patients, gaps in all studied ontologies were identified. These gaps are to be expected taking into account the purpose and the focus of each of the available ontologies and the very high rate of growth and change of the medical knowledge in oncology. The gaps identified are limited and adequate extensions to address them in the INTEGRATE solution in an efficient way have been proposed. This pragmatic approach allows the leverage of the huge effort that has been invested over time in the building and validation of these ontologies. A valuable lesson here is that, by using existing ontologies, rather than developing new domain specific ontologies [65], significant savings related to knowledge maintenance are gained while the INTEGRATE solution becomes future-proof. As clinical knowledge growth and changes, the new versions of these widely adopted ontologies will capture the new state of knowledge making use of the contribution of a large community of experts. In addition, with the use of general biomedical standards such as the SMOMED-CT, MedDRA and LOINC vocabularies and the BRIDG data model, this approach tested the application of those standards in one disease, rather than developing something in one disease and expecting it to apply to other diseases.

Furthermore, the SIL is used by the application layer to offer a set of diverse services for patient care and clinical research: DECIMA supports patient screening, CRP assists pathologists on structured review and analysis of pathology images, NONA supports cohort selection and the AP performs statistical and bioinformatics analysis of clinico-genomic data. These tools have been iteratively evaluated with clinical users and have shown good results and significant potential in several relevant scenarios as described in this paper. For example the SUS score in all tools ranges from 65 to 90 whereas the average SUS score from published studies is 62. However, despite the high technical readiness level of these tools, in order to be fully integrated in the daily clinical workflow, additional optimization work should be performed on GUI, performance and stability. Another valuable lesson here is that technological tools should always involve end-users in all phases of the development in an iterative process and that usability is equally important to the effectiveness of the tools in order to gain user acceptance.

Conflict of interest

We confirm that there are no conflict of interests for our publication

Acknowledgments

The authors gratefully acknowledge the financial support of the European Commission for the Project INTEGRATE.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbi.2016.05.006.

References

[1] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, Health Inform. Sci. Syst. 2 (3) (2014).

[2] A. Ryan, Towards Semantic Interoperability in Healthcare: Ontology Mapping from SNOMED-CT to HL7 version 3, Australasian Ontology Workshop (AOW2006), Hobart, Australia, in: M.A. Orgun, T. Meyer (Eds.), Conferences in Research and Practice in Information Technology (CRPIT), vol. 72, 2006.

[3] M. Tsiknakis, V.J. Promponas, N. Graf, M.D. Wang, S.T.C. Wong, N. Bourbakis, C. S. Pattichis, Guest editorial, computational solutions to large-scale data management and analysis in translational and personalized medicine, IEEE J. Biomed. Health Inform. 18 (3) (2014) 720–722.

[4] L. Martin, A. Anguita, N. Graf, M. Tsiknakis, M. Brochhausen, S. Rüping, A. Bucur, S. Sfakianakis, T. Sengstag, F. Buffa, H. Stenzhorn, ACGT: Advancing Clinico-genomic trials on cancer – four years of experience, MIE 2011, Oslo, Norway, August 2011, Stud. Health Technol. Inform. 169 (2011) 734–738.

[5] F. Köpcke, H.U. Prokosch, Employing computers for the recruitment into clinical trials: a comprehensive systematic review, J. Med. Internet Res. 16 (2014) e161.

[6] C. Safran et al., Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper, J. Am. Med. Inform. Assoc. 14 (1) (2007). Jan/Feb.

[7] H. Kondylakis, Giorgos Flouris, Irini Fundulaki, Vassilis Papakonstantinou, Manolis Tsiknakis, Flexible Access to Patient Data through e-Consent, MobiHealth, 2015.

[8] J. Iavindrasana, A. Depeursinge, P. Ruch, S. Spahni, A. Geissbuhler, H. Müller, Design of a decentralized reusable research database architecture to support data acquisition in large research projects, Stud. Health Technol. Inform. 129 (2007) 325–329.

[9] Data Protection Directive, 2015, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML> (last visited December 2015).

[10] Institute of Medicine (US) Forum on Drug Discovery, Development, and Translation, Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary. Washington (DC): National Academies Press (US), 3, Challenges in Clinical Research, 2010, Available from: <http://www.ncbi.nlm.nih.gov/books/NBK50888/>.

[11] E. Roelofs et al., Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies, Radiother. Oncol. (2010), http://dx.doi.org/10.1016/j.radonc.2010.08.009.

[12] ISO/IEC 42010:2007, Systems and Software Engineering – Recommended Practice for Architectural Description of Software-intensive Systems, 2007.

[13] OASIS SOA Reference Model TC, 2016, <https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm> (last visited February 2016).

[14] T. Souza, R. Kush, J.P. Evans, Global clinical data interchange standards are here!, Drug Discov Today 12 (2007) 174–181.

[15] M. Tsiknakis, M. Brochhausen, J. Nabrzyski, J. Pucaski, S. Sfakianakis, G. Potamias, C. Desmedt, D. Kafetzopoulos, A semantic grid infrastructure enabling integrated access and analysis of multilevel biomedical data in support of post-genomic clinical trials on Cancer, IEEE Transact. Inform. Technol. Biomed. 12 (2) (2008) 191–204. Special issue on Bio-Grids.

[16] W. Wilkowska, M. Ziefle, Privacy and data security in E-health: Requirements from the user's perspective, Health Inform. J. 18 (3) (2012) 191–201.

[17] M. Peleg, D. Beimel, D. Dori, Y. Denekamp, Situation-based access control: privacy management via modeling of patient data access scenarios, J. Biomed. Inform. 41 (2008) 1028–1040.

[18] OASIS Privacy by Design Documentation for Software Engineers, 2016, <http://docs.oasis-open.org/pbd-se/pbd-se/v1.0/csd01/pbd-se-v1.0-csd01.doc> (last visited February 2016).

[19] B. Claerhout, K. de Schepper, D. Pérez-Rey, A.I.D. Bucur, Contextualisation of ABAC Attributes through a Generic XACML Functionality Extension Mechanism, HEALTHINF, 2013, pp. 52–57.

[20] F. Priyatna, O. Corcho, J. Sequeda, Formalisation and Experiences of R2RML-based SPARQL to SQL query translation using Morph. World Wide Web Conference (WWW), 2014, pp. 479–490.

[21] K.A. Spackman, Normal forms for description logic expressions of clinical concepts in SNOMED RT, in: Proceedings of the AMIA Symposium, American Medical Informatics Association, 2001, p. 627.

[22] R. Alonso-Calvo, D. Perez-Rey, S. Paraiso-Medina, B. Claerhout, P. Hennebert, A. Bucur, Enabling semantic interoperability in multi-centric clinical trials on breast cancer, Comput. Methods Programs Biomed. 118 (3) (2015) 322–329.

[23] S. Paraiso, D. Perez Del Rey, A. Bucur, B. Claerhout, R. Alonso-Calvo, Semantic normalization and query abstraction based on SNOMED-CT and HL7: supporting multi-centric clinical trials, IEEE J. Biomed. Health Inform. (2014), http://dx.doi.org/10.1109/JBHI.2014.2357025.

[24] C. Shivade, C. Hebert, M. Lopetegui, M.C. de Marneffe, E. Fosler-Lussier, A.M. Lai, Textual inference for eligibility criteria resolution in clinical trials, J. Biomed. Inform. 58 (2015). S211–S218.

[25] L. Penberthy, R. Brown, F. Puma, B. Dahman, Automated matching software for clinical trials eligibility: measuring efficiency and flexibility, Contemp. Clin. Trials 31 (2010) 207–217.

[26] G. Joseph, D. Dohan, Recruiting minorities where they receive care: institutional barriers to cancer clinical trials recruitment in a safety-net hospital, Contemp. Clin. Trials 30 (2009) 552–559.

[27] L. Penberthy, B. Dahman, V.I. Petkov, J.P. DeShazo, Effort required in eligibility screening for clinical trials, J. Oncol. Pract. 8 (2012) 365–370.

[28] S.W. Tu, M. Peleg, S. Carini, M. Bobak, J. Ross, D. Rubin, I. Sim, A practical method for transforming free-text eligibility criteria into computable criteria, J. Biomed. Inform. 44 (2) (2011) 239–250.

[29] I. Sim, S.W. Tu, S. Carini, H.P. Lehmann, B.H. Pollock, M. Peleg, K.M. Wittkowski, The Ontology of Clinical Research (OCRe): an informatics foundation for the science of clinical research, J. Biomed. Inform. 52 (2014) 78–91.

[30] C. Weng, X. Wu, Z. Luo, M.R. Boland, D. Theodoratos, S.B. Johnson, EliXR: an approach to eligibility criteria extraction and representation, J. Am. Med. Inform. Assoc.: JAMIA 18 (Suppl 1) (2011) 116–124.

[31] Z. Luo, S.B. Johnson, A.M. Lai, C. Weng, Extracting temporal constraints from clinical research eligibility criteria using conditional random fields, AMIA Annu. Symp. Proceed. 2011 (2011) 843.

[32] T. Hao, A. Rusanov, M.R. Boland, C. Weng, Clustering clinical trials with similar eligibility criteria features, J. Biomed. Inform. 52 (2014) 112–120.

[33] M.R. Boland, R. Miotto, C. Weng, A method for probing disease relatedness using common clinical eligibility criteria, MEDINFO, 2013, pp. 481–548.

[34] C. Weng, A. Yaman, K. Lin, Z. He, Trend and Network Analysis of Common Eligibility Features for Cancer Trials in clinicaltrials.gov, ICSH, pp. 130–141, 2014

[35] C. Weng, S.W. Tu, I. Sim, R. Richesson, Formal representation of eligibility criteria: a literature review, J. Biomed. Inform. 43 (2010) 451–467.

[36] K. Milian, R. Hoekstra, A. Bucur, A. ten Teije, F. van Harmelen, J. Paulissen, Enhancing reuse of structured eligibility criteria and supporting their relaxation, J. Biomed. Inform. 56 (2015) 205–219.

[37] A. Bucur, J. van Leeuwen, N. Chen, B. Claerhout, K. de Schepper, D. Pérez-Rey, R. Alonso-Calvo, L. Pugliano, K. Saini, Supporting Patient Screening to Identify Suitable Clinical Trials, MIE, 2014, pp. 823–827.

[38] K. Milian, A. Bucur, A. ten Teije, Formalization of clinical trial eligibility criteria: Evaluation of a pattern-based approach, in: 2012 IEEE International Conference on Bioinformatics and Biomedicine, 2012, pp. 1–4.

[39] J. van Leeuwen, A. Bucur, B. Claerhout, K. de Schepper, D. Pérez-Rey, R. Alonso-Calvo, BRIDG-based Trial Metadata Repository – Need for Standardized Machine Interpretable Trial Descriptions, HEALTHINF, 2014, pp. 453–458.

[40] A. Bucur, DECIMA, J. Clin. Bioinform. 5 (S-1) (2015) S11.

[41] A. Iliopoulos, I. Karatzanis, M. Tsiknakis, V. Sakkalis, K. Marias, A collaborative central reviewing platform for cancer detection in digital microscopy images, in: 6th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation – The CHIC Project Workshop (IARWISOCI). November, 3–4, 2014, Athens.

[42] W.A. Woodwar, N. Sneige, K. Winter, et al., Web based pathology assessment in RTOG 98-04, J. Clin. Pathol. 67 (9) (2014) 777–780.

[43] M.G. Wallis, K. Clements, O. Kearins, et al., The effect of DCIS grade on rate, type and time to recurrence after 15 years of follow-up of screen-detected DCIS, Br. J. Cancer 106 (2012) 1611–1617.

[44] S.C. Lester, S. Bose, Y.Y. Chen, et al., Protocol for the examination of specimens from patients with invasive carcinoma of the breast, Arch. Pathol. Lab. Med. 133 (2009) 1515–1538.

[45] G. Manikis, E. Maniadi, M. Tsiknakis, K. Marias, Multi-Modal Medical Data Analysis Platform (3MDAP) for analysis and predictive modelling of cancer trial data, in: 6th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation – The CHIC Project Workshop (IARWISOCI). November, 3–4, 2014, Athens.

[46] K. Marias, D. Dionysiou, V. Sakkalis, N. Graf, R.M. Bohle, P.V. Coveney, S. Wan, A. Folarin, P. Büchler, M. Reyes, G. Clapworthy, E. Liu, J. Sabczynski, T. Bily, A. Roniotis, M. Tsiknakis, et al., Clinically driven design of multi-scale cancer models: the ContraCancrum project paradigm, Interface Focus, 2011.

[47] ISO/IEC DIS 25023, Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Measurement of System and Software Product Quality.

[48] R. Likert, A technique for the measurement of attitudes, Arch. Psychol. 140 (1932) 1–55.

[49] J. Brooke, SUS-A quick and dirty usability scale, Usability Evaluation in Industry 189, 194, 1996, pp. 4–7.

[50] E. Roelofs, L. Persoon, S. Nijsten, W. Wiessler, A. Dekker, P. Lambin, Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial, Radiother. Oncol. 108 (1) (2013) 174–179.

[51] J. Sauro, J.R. Lewis, Correlations among prototypical usability metrics: evidence for the construct of usability, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2009, pp. 1609–1618.

[52] K. Kosman, A. Talaei-Khoei, A.H. Ghapanchi, P. Ray, A framework for assessing clinical trial informatics systems: case of caBIG and its clinical trial management system, IJCIH 1 (3) (2012) 269–280.

[53] A. Califano, A.M. Chinnaiyan, G.M. Duyk, S.S. Gambhir, T. Hubbard, D.J. Lipman, L.D. Stein, J.Y. Wang, O.T. Bartlett, C.L. Harris, An Assessment of the Impact of the NCI Cancer Biomedical Informatics GRID (caBIG) Available online <http://deainfo.nci.nih.gov/advisory/bsa/archive/bsa0311/caBIGfinalReport.pdf>2011.

[54] A. Bucur, S. Rüping, T. Sengstag, S. Sfakianakis, M. Tsiknakis, D. Wegener, The ACGT project in retrospect: lessons learned and future outlook, ICCS (2011) 1119–1128.

[55] J. Marés, L. Shamardin, G. Weiler, A. Anguita, S. Sfakianakis, E. Neri, S.J. Zasada, N. Graf, P.V. Coveney, p-medicine: A medical informatics platform for integrated large scale heterogeneous patient data, in: AMIA Annu. Symp. Proc., 2014, pp. 872–881.

[56] H. Kondylakis, L. Koumakis, M. Tsiknakis, M. Marias, E. Genitsaridi, G. Pravettoni, A. Gorini, K. Mazzocco, Smart recommendation services in support of patient empowerment and personalized medicine, Multimed. Serv. Intell. Environ. (2013) 39–61.
[57] H. Kondylakis, E. Kazantzaki, L. Koumakis, I. Genitsaridi, K. Marias, A. Gorini, K. Mazzocco, G. Pravettoni, D. Burke, G. McVie, M. Tsiknakis, Development of Interactive Empowerment services in support of personalized medicine, eCancer Med. Sci. J. 8 (2014) 400.
[58] I. Genitsaridi, H. Kondylakis, L. Koumakis, K. Marias, M. Tsiknakis, Evaluation of personal health record systems through the lenses of EC research projects, Comput. Biol. Med. 59 (2015) 175–185.
[59] IBM Cognos Platform, 2015, <http://www-01.ibm.com/software/in/analytics/cognos/platform/> (last visited December 2015).
[60] Microsoft Caradigm Intelligence Platform, 2016, <https://www.caradigm.com/en-us/solutions-for-population-health/intelligence-platform/> (last visited February 2016).
[61] S.N. Murphy, S.E. Churchill, L. Bry, H. Chueh, T. Cai, S. Weiss, et al., Instrumenting the health care enterprise for discovery research in the genomic era, Genome Res. 360 (13) (2009) 1278–1281.
[62] B.D. Athey, M. Braxenthaler, M. Haas, Y. Guo, tranSMART: An open source and community-driven informatics and data sharing platform for clinical and translational research, AMIA Jt. Summits Transl. Sci. Proc. 13 (2013) 6–8.
[63] H. Kondylakis, D. Plexousakis, V. Hrgovcic, R. Woitsch, M. Premm, M. Schüle, Agents, Models and Semantic Integration in Support of Personal eHealth Knowledge Spaces, WISE, 2014, pp. 496–511.
[64] A. Ibrahim, A.I.D. Bucur, A. Dekker, M.S. Marshall, D. Pérez-Rey, R. Alonso-Calvo, H. Stenzhorn, S. Yu, C. Krykwinski, A. Laarif, K. Mehta, Analysis of the Suitability of Existing Medical Ontologies for Building a Scalable Semantic Interoperability Solution Supporting Multi-site Collaboration in Oncology, BIBE, 2014, pp. 204–211.
[65] M. Brochhausen, G. Grigonyte, L. Martín, N. Graf, J. Haller, B. Smith, M. Tsiknakis, The ACGT Master Ontology and its applications – towards an ontology-driven cancer research and management system, J. Biomed. Inform. 44 (1) (2011) 8–25.