ICT-2010-270253

INTEGRATE

Driving excellence in Integrative Cancer Research through Innovative Biomedical Infrastructures

STREP
Contract Nr: 270253

Deliverable:
D8.1 Public summary of the project

Due date of deliverable: (28-02-2011)
Actual submission date: (02-03-2011)

Start date of Project: 01 February 2011 Duration: 36 months

Responsible WP8: Philips

Revision: <outline, draft, **proposed**, accepted>

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Service	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (excluding the Commission Services)	

0 DOCUMENT INFO

0.1 Author

Author	Company	E-mail
Anca Bucur	Philips	anca.bucur@philips.com

0.2 Documents history

Document version #	Date	Change
V0.1		Starting version, template
V0.2		Definition of ToC
V0.3		First complete draft
V0.4		Integrated version (send to WP members)
V0.5		Updated version (send PCP)
V0.6		Updated version (send to project internal reviewers)
Sign off		Signed off version (for approval to PMT members)
V1.0		Approved Version to be submitted to EU

0.3 Document data

Keywords	
Editor Address data	Name: Anca Bucur Partner: Philips Address: HTC34 6-033 Phone: +31 40 2749699 Fax: +31 40 27 44797 E-mail: anca.bucur@philips.com
Delivery date	02.03.2011

0.4 Distribution list

Date	Issue	E-mailer
February 2011		AI_integrate_all@natlab.research.philips.com
		Loukianos.gatzoulis@ec.europa.eu

Table of Contents

1 PROJECT INFO

General Info	
Acronym	INTEGRATE
Name	Driving excellence in Integrative Cancer Research through Innovative Biomedical Infrastructures
Web page	www.fp7-integrate.eu
Reference	FP7-ICT-2009-6-270253
Contract type	STREP-CP-FP-INFSO
Start date	February 01, 2011
End date	January 31, 2014
Duration	36 months
Total budget	5803 K Euro
EC contribution	3974 K Euro
Coordinator	Philips

Participants			
No.	Name	Short name	Country
1	Philips Electronics Nederland BV	PENB	NL
2	Breast International Group	BIG	B
3	Foundation for research and technology Hellas	FORTH	GR
4	Custodix	CUSTODIX	B
5	Institut Jules Bordet	IJB	B
6	Universidad Politecnica de Madrid	UPM	SP

2 VISION

Significant effort and financial investments in biomedical research and in the healthcare industry have resulted into a wealth of data, information and knowledge with the potential to bring along large qualitative improvements in patient outcome. Paradoxically, despite the huge volumes of data generated in clinical practice, very little of it is used for pushing the frontiers of medical science forward. Lack of data still hampers biomedical research, while the secondary use of the large amounts of patient data collected in clinical practice is very low. On the clinical side, the transfer of clinical research outcomes to clinical practice is often slow, creating a widening knowledge gap between research and care. Additionally, to improve patient outcomes and reduce healthcare costs, we need to replace expensive trial-and-error approaches with effective individualized treatments.

Large efforts dedicated to biomedical research generated many exciting discoveries world-wide, but have not brought so far the desired breakthrough benefits in the clinic, in dramatically improving the patient outcomes[1].

Synergy between clinical research and the VPH community will speed up the development and validation of multiscale predictive models in breast cancer - an area in which modelling significantly lags behind due to high variability in molecular/genetic and tissue level characteristics –, improve patient outcomes and reduce costs. INTEGRATE will enable this through a dynamic environment, collaboration tools and secure access to comprehensive datasets.

Fragmented efforts, lack of common methodologies and research frameworks, and lack of sufficient and high quality data result in improper validation and make reproducibility of research results a difficult task. As a consequence, many new drugs and biomarkers fail in the market, at the end of very expensive development processes, and breakthrough discoveries may be abandoned without having the chance to prove their real value. The huge potential of the current biomedical research cannot be fully exploited and confirmed in the clinical practice in the absence of a coordinated and systematic approach.

The same fragmentation is apparent at the level of infrastructures, systems and tools used in basic and clinical research and in clinical care. While the healthcare industry continues to improve its capabilities for electronic data capture, a gap remains in the ability of IT systems to deliver knowledge and insight back to the very researchers and clinicians they are intended to support. The consequences are inefficiency and duplication of data and effort, and hampered wide-scale multidisciplinary collaboration. The low integration, having as main barriers the lack of interoperability and the low adoption of common standards and terminologies, also turns access to sufficient and high quality data for carrying out basic and clinical research into a real challenge.

Furthermore, we lack accurate predictive models able to integrate all the data and knowledge available at the various dimensional scales into a comprehensive view that would enable the selection of the most suitable therapy yielding the best possible patient outcome. Often research is carried out in isolation, duplicating effort, re-exploring dead-ends, and without being able to share data, results and knowledge. Channelling all that effort, expertise and knowledge towards a common goal of building the large picture of an individual in the context of a complex disease could facilitate discovery at unprecedented speed and scale.

The VPH Initiative (VPH-I) identified the need for an all-comprising integration of data, information, knowledge and wisdom as a huge challenge with a tremendous impact[2].

[1] Integrating Genomic Medicine Into Clinical Practice For Common Chronic Diseases Still In The Early Stages, JAMA News Release, 2008.

[2] A Vision and Strategy for the VPH in 2010 and beyond, www.biomedtown.org/biomed_town/VPH/vphonoe./vphnoe_vision/VPH%20Vision%20submitted%20141209 pdf

Achieving that objective requires the development of specialized infrastructures able to support sharing of high volumes of heterogeneous biomedical data at all spatial scales, and wide-scale research collaboration across disciplines, institutions and industries. Next to sharing data, the biomedical community needs to be able to jointly develop and share models, methodologies and tools. Only this close, multi-faceted collaboration can efficiently exploit the exponential growth of the available data and knowledge in the biomedical field and turn it into benefits for the patients.

The above needs have been recognized world-wide with many high profile initiatives, such as caBIG[3] and Sage Bionetworks[4], having as a mission to bring together researchers and their data and knowledge, build tools and infrastructures enabling sharing and collaboration, support reuse of data, models and tools, and promote common standards and interoperability.

INTEGRATE shares this vision and aims to build solutions that support a large and multidisciplinary biomedical community ranging from basic, translational and clinical researchers to the pharmaceutical industry to collaborate, share data and knowledge, and build and share predictive models for response to therapies, with the end goal of improving patient outcome. Moving away from empirical medicine, towards evidence-based personalized care has the potential to both dramatically improve patient outcome and to reduce costs.

Our infrastructure will bring together heterogeneous multi-scale biomedical data generated through standard and novel technologies within post-genomic clinical trials and seamlessly link to existing research and clinical infrastructures, such as clinical trial systems, eCRFs, and hospital EHRs, and to relevant external biomedical infrastructures. A unique quality of the INTEGRATE approach is the full commitment of the Breast International Group[5], as a partner in the project, to contribute patient data and the extensive basic, translational and clinical research expertise of their network to build solutions based on challenging but realistic use cases.

The INTEGRATE Vision

Within the VPH context the vision of INTEGRATE is to empower the clinical researcher with the unique opportunity to access breast-cancer multi-scale data from 29 European cancer centers and develop breast cancer predictive model methodologies related to specific clinical questions for optimizing therapy, identifying high-risk patients, etc.

Within INTEGRATE the clinical researcher will be able to retrieve temporal (e.g. before and after therapy), multi-level (e.g. from microarray to MRI/PET), data from specific population groups (e.g. postmenopausal women), with specific characteristics (e.g. that have received specific therapy regime) and then extract/develop predictive biomarkers/models (e.g. based combination of imaging/genetic biomarkers) that could answer a question such as “can these models/biomarkers help predict the specific therapy outcome for a patient in order to avoid unnecessary/costly treatment?”

[3] http://cabig.cancer.gov/

[4] http://www.sagebase.org/

[5] http://breastinternationalgroup.org/

3 INTEGRATE APPROACH

The INTEGRATE environment will facilitate collaboration between cancer research centres and other relevant stakeholders such as basic research organizations and pharmaceutical companies, and will be evaluated and validated in the context of the NeoBIG research program of BIG, in challenging but realistic scenarios in the breast cancer domain. This also requires compliance with a wide set of regulations and laws in the context of security, safety and privacy protection.

The NeoBIG clinical program was developed specifically to enhance and accelerate biomarker discovery and validation within the arena of early drug development in breast cancer. In this context, BIG aims to build on its already well established research collaborations to introduce a new model for the conduct of early breast cancer trials in the era of molecular oncology. This approach – to be launched in a series of international neo-adjuvant trials – will be based on the evaluation of new promising *targeted therapies* in *different molecular breast cancer subgroups*. NeoBIG brings together cancer research centres, hospitals and specialised laboratories to carry out an innovative program of prospective neo-adjuvant breast cancer trials with novel targeted therapies. The advantage of the neo-adjuvant setting is the ability to have an early surrogate of treatment efficacy as well as the opportunity to take serial biopsies to molecularly characterize the tumour and its response to treatment at multiple data levels. The ultimate aim is to perform biomarker research and validation in the pre-surgical (neo-adjuvant) setting to accelerate their implementation in the post-surgical (adjuvant) setting. This clinical program will also provide valuable input for the development of the INTEGRATE modelling framework.

The applicability of the INTEGRATE concepts and technologies will go far beyond the breast cancer domain. We will ensure the extensibility of our solutions based on requirements and use cases provided by communities of experts in other disease domains, ensuring the applicability and promoting the further adoption of our solutions beyond the breast cancer research and treatment community.

As we believe that the efforts towards integration and collaboration will only be successful if applied at a large scale, we aim to provide solutions to a wide community, much beyond the already very large network of BIG. Additionally, we have set collaboration and aim to join forces with an important initiative in the US, Sage Bionetworks[4].

We are aware that to achieve adoption of such scale we need to make use whenever possible of existing standards and terminologies. Additionally, building solutions based on standards will enable us to make use of previous efforts in data sharing, modelling and knowledge generation, and to access important external sources of data and knowledge.

Therefore, the INTEGRATE project aims to support a new integrative approach in medicine by building flexible research infrastructures enabling large-scale collaboration and sharing of data, knowledge, models and tools, and improving semantic interoperability with existing relevant systems. Essential steps for achieving interoperability improvement include the definition of sound information models describing the clinical research systems, building on existing research results when possible. Electronic health records too need to be properly modelled; to that end we will adopt the appropriate state-of-the-art representation formalisms such as HL7 CDA, the OpenEHR Reference Model, ISO/EN 13606, etc.

The INTEGRATE research infrastructure will store and manage a wide range of datasets, such as clinical, bio-molecular, imaging, models, annotations and other metadata, and put a strong focus on data privacy and security.

The semantics of the clinical terms should be captured by standard terminology systems such as SNOMED CT, ICD, LOINC. The scalability of the solution needs to be achieved by modularization, e.g. instead of aiming at inclusion of the complete SNOMED terminology (more than 300 000 concepts) we will identify a core subset that covers the chosen clinical domain and the datasets in our repositories. Such core data set shall be validated both by clinical and knowledge engineering experts to assure proper coverage and soundness. In the process of identifying the core data set and the corresponding mapping tools, care will be taken to allow to easily extend the core data set, should the inclusion of new concepts become necessary. We will rely when possible on existing initiatives and previous efforts in terminology development and standardization.

Utilizing the core dataset we will devise a mapping system between the information models of the INTEGRATE research infrastructure and the information models of the clinical research systems and those of EHRs. As this semantic dataset will be mapped to concepts from existing standardized vocabularies (e.g. well established and widely used clinical terminologies such as SNOMED CT, ICD-10), we will foster scalable semantic interoperability not only among the clinical infrastructures within the project, but also towards other healthcare and research organizations adhering to the adopted standards.

The INTEGRATE project aims to support sustainability beyond the scope of the research project, building a long lasting clinical and translational research infrastructure that will promote scientific collaboration among cancer research centres in Europe (and the rest of the world), pharmaceutical companies, and biomedical research communities well beyond the FP7 funding period. We will actively promote our approach and solutions in wide user communities and in other disease domains.

The INTEGRATE project will carry out the research and development work in a user-driven manner, promoting close collaboration between the ICT researchers and the clinical researchers participating to the project. We will define realistic use cases based on requirements input provided by our clinical partners and by our external pool of expertise. By collaboration and use of shared standards we aim at convergence with similar initiatives focused on large scale integration in the biomedical area.

4 OBJECTIVES

BUILD INFRASTRUCTURE COMPONENTS AND TOOLS FOR THE STORAGE, SHARING AND MANAGEMENT OF DATA, INFORMATION, KNOWLEDGE AND MODELS

INTEGRATE will build reusable components based on which we will set up a dynamic infrastructure supporting our user community to store, manage and share biomedical data, models, tools, methodologies, and knowledge.

We provide support for building and linking comprehensive datasets, for format conversion and annotation. We will provide standard-based interfaces and services/tools enabling users to store, query and manage heterogeneous multi-scale data, predictive models, annotations and other types of metadata preserved in our repositories.

The current heterogeneity in healthcare-related research, manifest at the level of methodologies, workflows, data processing, and ICT infrastructures, tools and services, has significant negative impact on medical knowledge discovery, on the validation of clinical research results and on the adoption of the new results in clinical care for more predictive, individualized, effective and safer healthcare. On the ICT side, a main barrier is the lack of interoperability among relevant infrastructures, services and tools, due to the low adoption of common standards and terminologies. INTEGRATE will build solutions based on established standards for storing, annotating and exchanging biomedical data, metadata, models and knowledge. The use of established standards and terminologies also supports the integration with existing infrastructures and the access to external relevant repositories adhering to those standards and terminologies.

BUILD TOOLS TO ENABLE COLLABORATION

We will support the description and execution of shared, multi-disciplinary and multi-site workflows. An important requirement for emergent collaborations is a shared workspace that is accessible to all collaborators. Ideally, this workspace should include all the important transactions that have taken place among scientific workers. In addition to helping a group of collaborators learn from past transactions and take the best step forward, the workspace will facilitates stigmergy, i.e., it will enable a worker's contribution to stimulate others to build on that contribution without any direct communication between the workers.

In addition to a shared and open workspace, emergent collaboration requires meta-level information that highlights the significance of the transactions that are occurring or have occurred within a group of collaborators. Since many workers may be making their contributions at any point in time, which of these many contributions should other members be focusing on and building on at that time? How should they prioritize their effort? Meta-level information that indicates the significance of transactions within a group can help a group determine which issues it should tackle first and which contributions should be given more weight than others.

Specifically, tools enabling the linking and tagging features which make them suitable for generating the meta-level information that indicates the significance of transactions among a group of workers will be developed.

BUILD PREDICTIVE MODELS AND A MODELLING METHODOLOGY AND FRAMEWORK

The INTEGRATE project will propose an approach and a methodology and build a framework enabling the development of multi-scale predictive models of response to therapy in breast cancer, making use of multi-level heterogeneous data provided by clinical trials in the neo-adjuvant setting. The models developed will be based on

realistic clinical research scenarios and on comprehensive datasets from rigorously conducted clinical trials. The models will also be used to validate the INTEGRATE approach and the appropriateness of the INTEGRATE infrastructure.

By proposing a methodology and building a framework for predictive models development within clinical trials we support more efficient development and validation of such models and faster adoption into clinical practice through the process of clinical trial validation.

The aim of our biomedical modelling and simulation research will be focused on predicting the responsiveness of patients to specific drugs. This could lead to a more targeted and personalized treatment of the patient, avoiding at the same time a great deal of suffering due to unnecessary or ineffective treatment.

Next to building models at individual scales, we aim to integrate those models in a complex multi-scale model with the potential to provide a robust and accurate prediction of the patient response to specific therapies and drugs.

> The challenge will be to evaluate the relative contributions of multiple levels of data, both molecular and clinical, in predicting breast cancer outcome and response to anticancer agents. Developing integrative models that combine clinical and complex multilevel molecular factors, such as gene expression patterns, functional proteomics, traditional clinico-pathological risk factors and treatment information, will also increase our understanding of the complex genotype-phenotype inter-relationship involved in breast cancer.
>
> It is conceivable that one or more gene-expression classifiers could be combined into one model together with traditional clinico-pathological parameters that still retain important prognostic information…
>
> *C. Sotiriou and M. Piccart,*
> *Taking gene expression profiling to the clinic: when will molecular signatures become relevant to patient care, Nature Reviews. 2007.*

Combining these models and the enhanced integration of the multi-dimensional genome data will result in increased understanding of breast cancer biology, large scale collaboration, the validation of biomarkers in rigorous clinical settings, and will facilitate the ultimate goal of individualized therapy in cancer patients.

LINK TO EXTERNAL SOURCES OF INFORMATION

Biomedical research often relies on access to the many external repositories of data, information and knowledge. In the INTEGRATE project we will provide uniform standardized interfaces to external resources relevant to our user community. The external data will be used in the development of predictive models and to provide input to various analysis and communication tools.

It is not realistic to assume that all available relevant data of the collaborating organizations will be retrieved and stored in the INTEGRATE repositories. Instead, we will provide a dynamic infrastructure able to assist the users to carry out complex analyses combining data in the INTEGRATE environment with data in the existing systems and repositories of participating organizations.

Next to those, other external sources will be accessed, such as databases of clinical studies, scientific literature, drug-drug interactions and adverse events, biomedical repositories (such as public datasets), models repositories, clinical guidelines and treatment protocols.

In the bioinformatics field, an extremely large body of relevant external data resources is available. During the project, an approach to making those data resources available to the platform will be developed. Based on the use cases and scenarios, a selection of resources will be made available within the platform. The available data spans a wide range of molecular types and interactions, both on a content level (the “raw” data) as on a description level (various ontologies, vocabularies and literature).

With respect to accessing external sources of data and knowledge we will also focus when possible on existing standards. Standards for representation of web-enabled data and knowledge, such as RDF[6], are already widely adopted. Several relevant public biomedical knowledge repositories, such as GO[7] and PubMed[8] adhere to the RDF serialization format.

ENABLE SEMANTIC INTEROPERABILITY TO EXISTING RESEARCH AND CLINICAL INFRASTRUCTURES

The ability to interface to existing medical research infrastructures is an important objective of INTEGRATE, as it is the basis for reaching a large community of users. The adoption of our solutions depends on an efficient link of the data in the INTEGRATE repositories with data in existing research infrastructures to facilitate reuse of data and avoid the need for multiple data entry. Furthermore, to promote the fast adoption of the clinical research results into clinical care, we need to also target standards-based interoperability to existing clinical infrastructures. Interoperability will enable the INTEGRATE dedicated infrastructure and tools to efficiently exchange and share data when needed with the existing research and clinical infrastructures. The interoperability with existing research and clinical systems will be tackled based on relevant data and communication standards in clinical research and care (e.g. CDISC, MedDRA, HL7, OpenEHR, etc.) and terminologies (e.g. SNOMED, ICD, LOINC, etc.).

From the technology viewpoint, in order to provide an efficient, robust and semantically interoperable solution, one needs to move from plain keyword matching to a combined approach where keywords are mapped to higher level concepts with clearly defined semantics. Such concepts are usually organized in concept hierarchies and include domain specific attributes and relations. Reasoning at this level, rather than at keyword level, is expected to enable us to move from error-prone lexical matching to more robust semantic-aware solutions.

VALIDATION OF THE INTEGRATE ENVIRONMENT AND TOOLS

The capability of the INTEGRATE environment to achieve the above mentioned objectives will be demonstrated through prototypes implementing realistic clinical scenarios. These scenarios will also allow us to demonstrate the interoperability of our solutions with existing infrastructures.

The prototypes will be first deployed at the sites of the INTEGRATE clinical partners, members of the BIG network, and validated in concrete scenarios. In a second stage, we will also promote our solutions towards external user groups that expressed interest in INTEGRATE and with which we have set up collaborations.

[6] http://www.w3.org/2001/sw/wiki/RDF

[7] http://www.geneontology.org/

[8] http://www.ncbi.nlm.nih.gov/pubmed

5 TECHNICAL APPROACH

INTEGRATE will cover technological research, service development, system integration, testing, and uptake activities in a complex interaction. Therefore a clear and well-structured methodology for the project is essential.

Phase one: Definition

The definition phase marks the beginning of the project. There are three fundamental project objectives, of the utmost importance to the overall success of the project, to be addressed in the definition phase: 1) User acceptance, 2) availability of technology and 3) exploitability of the *INTEGRATE* solution.

- **Creating scenarios** The scenarios will be deduced from the domain settings defined by users and provide the framework for all subsequent user requirements specifications.
- Defining **functional and non-functional user requirements specifications** based on the user scenarios involves addressing seamless interoperability, openness, and specific user requirements for the selected user cases must also be integrated in the requirements specification.
- Defining **trust and security user requirements specifications** involves identifying not only trust (i.e. privacy, auditability and assurance) and security issues, but also legal and ethical issues, which in turn translate to new trust and security requirements.
- Defining **societal user requirements specifications** will be done by correlating ethical, regulatory and policy issues with the deployment and wide spread use of the *INTEGRATE services*. Aspects of e.g. social acceptance, regulatory frameworks for surveillance and control, privacy of data, governmental provisions, etc. will be addressed and integrated.
- **Volatility and Evolution of Requirements.** During the time it takes to develop a system the users' needs may mature because of increased knowledge brought on by the development activities, or they may shift to a new set of needs because of unforeseen organizational or environmental pressures.

Phase two: Research and development

The complexity of the domain which is addressed by the *INTEGRATE* project necessitates that a spiral process of requirements analysis, elicitation, technological implementations, integration and validation is adopted. Specific techniques have also been selected for the elicitation, negotiation and agreement of requirements as well as their validation. These techniques are ***scenarios*** – during the requirements specification phase - and ***prototyping*** – during the R&D phase.

Phase three: System integration and testing

With the successful completion of all tasks in the research and development phase, the project has reached the stage, where realization of the environment and its validation by users in user environments are possible.

Phase four: Evaluation, validation, and uptake

The final phase of the project will be concerned with validation of the prototype *INTEGRATE* platform in the various user cases as well as overall project evaluation and preparation for take-up activities and exploitation.

In this phase the INTEGRATE services and environment will be exposed to the use cases described in the user scenarios.

Validation of the prototypes will involve user testing (letting the users execute the user scenarios on location). Also the socio-economic and security aspects will be tested, by involving focus groups and questionnaires. Validation will be performed jointly by the users and the developers and documented in a report for each user scenario.

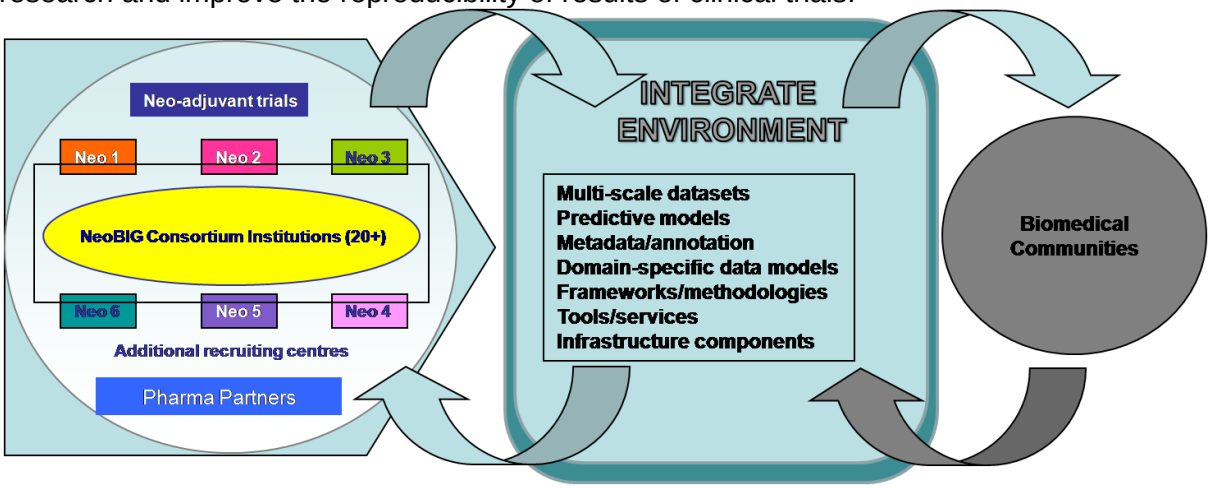
6 EXPECTED RESULTS

INTEGRATE targets specific research objectives in a sharply focused approach to address data integration, collaboration as well as modelling and simulation. Special emphasis is given in making sure that the new tools, services and applications to be developed in *INTEGRATE* will also be evaluated on their effectiveness and their ability to interface with existing medical research infrastructures.

The partnership within INTEGRATE with the Breast International Group enables us to bring the VPH community closer to the world of clinical research, and provide a path towards clinical validation of VPH models. Clinical trials are a formal, well regulated and statistically rigorous process and represent the way in which research results are validated and can be introduced into clinical practice.

INTEGRATE will build infrastructure components and tools providing uniform, secure and standardized interfaces and enabling the sharing into a large biomedical community of comprehensive datasets and knowledge generated within clinical trials. Access to this data and knowledge has the potential to fuel new research in the biomedical community, which is currently hampered by the lack of sufficient and high quality datasets.

The infrastructure and tools developed by the INTEGRATE project will also support BIG to promote in the clinical community new methodologies and define standards concerning the collection, processing, annotation and sharing of data in clinical research and improve the reproducibility of results of clinical trials.



We exploit the unique opportunity of the NeoBIG-empowered collaborative environment to combine multi-scale biomarkers (from genetic level to tissue level including imaging biomarkers) collected in neo-adjuvant studies. The aim is to identify predictive and prognostic biomarkers that can assess novel targeted therapies in a quick and robust manner. This will take the form of a 'use-case' VPH scenario emanating from and being deployed within the INTEGRATE environment. The goal is to demonstrate that the predictive power of responsiveness can be enhanced by using multi-scale biomarker signatures. The development of predictive models in an environment build together with the clinical trials community enables us to benefit of an accelerated adoption process towards clinical practice.

Important exploitation goals of INTEGRATE are the sustainability of our solutions beyond the duration of the project and the use of the project outcomes by stakeholders outside of the consortium. We believe that a successful dissemination of our results is key to the wide-scale adoption of our solutions.

7 IMPACT

Our vision is to drive research excellence in oncology through a unique accessible biomedical infrastructure integrating diverse mega-datasets, building predictive bionetworks and offering advanced tools to guide the development of effective human therapeutics and diagnostics. These comprehensive datasets will also become available to the biomedical research community through the INTEGRATE infrastructure.

7.1 More predictive, individualized, effective and safer healthcare

The development in the INTEGRATE project of a modelling framework and of predictive models of response in the context of post-genomic clinical trials has the potential to contribute towards a 'modelling aided' optimal treatment design for cancer patients that will positively influence the treatment outcome. We also aim to initiate a paradigm shift in breast cancer treatment selection supported by cancer treatment planning, treatment monitoring and outcome prediction in silico (i.e on the computer).
To improve patient outcome in oncology, the research and clinical communities need to move away from predominantly empirical approaches, towards enabling "the rational and tailored use of cancer therapies for individual patients"[9].
More effective healthcare is achieved by selecting the optimal therapy for an individual patient. Using the INTEGRATE platform it will be possible to efficiently run trials that could shed light in the optimal prediction for the candidate therapeutic schemes/schedules based on the patient's specific data. Safer healthcare could be achieved by reducing adverse effects of therapy within dedicated studies making use of the INTEGRATE environment.
Therefore, the envisaged outcome is expected to function as an attractor back to Europe to the research activities of the "complementary" pharmaceutical industry.

7.2 Improved interoperability of biomedical information and knowledge

Taking into account the need for reuse, efficiency and wide-scale integration, the INTEGRATE project will have a strong focus on standards-based interoperability. We will build a flexible infrastructure consisting of interoperable components interconnected by standard interfaces, we will develop uniform access to relevant external resources and services, and we will insure interoperability with relevant existing infrastructures in clinical research and care. The INTEGRATE project will assess relevant existing infrastructures from an adopt, adapt and interoperate perspective.
Interoperability with other VPH infrastructures and the development of shared standards are of prime focus for our project. As collaboration and integration at a global level is essential for the success of the VPH efforts and also a desired way of working for our clinical partners, we will aim at interoperability and harmonization with important similar initiatives in the biomedical area, adherence to relevant existing standards and contribution to emerging standards and shared methodologies.
INTEGRATE aims to build an environment providing to its users full support for collaboration and sharing of complex multi-level datasets and models, but also access to relevant external data, knowledge and services. At the same time, we aim to enable the biomedical research community to benefit of the comprehensive datasets

[9] Genomic markers for decision making: what is preventing us from using markers?, Coyle, V.M., Johnston, P.G., Nat. Rev. Clin. Oncol, 7, 90-97, 2010

preserved by the INTEGRATE environment, and of our predictive models and tools. This bi-directional exchange can only be efficiently achieved through standards-based semantic interoperability.

7.3 Social Impact

Allowing for discoveries in the laboratory to be quickly transferred to the clinical management and treatment of patients can bring important societal benefits by significantly improving patient outcomes. Additionally, there is a strong need to enable the rational and personalized use of treatments that suit individual patients, and to move away from the current predominantly empirical approach.

Providing the necessary infrastructure, tools and services to the clinical research community will enable them to reduce costs by more efficiently setting up and carrying out clinical trials; better reuse of data, knowledge and tools; reduced duplication of efforts; easier access to all relevant information out of external sources; and more insightful generation of new research hypotheses. In the end this means quicker validation of new discoveries in clinical trials and transfer of new results into clinical care to become part of new treatments and improve patient outcomes. Providing standards-based infrastructure and services enabling biomedical researchers to build comprehensive molecular and clinical datasets will also support the definition of validated disease models that can improve the speed and efficiency of therapeutic drug development.