

IMI1 Final Project Report Public Summary

Project Acronym: EHR4CR

Project Title: Electronic Health
Record systems for Clinical Research

Grant Agreement: 115189

Project Duration: 01/03/2011 - 29/02/2016

1. Executive summary

1.1. Project rationale and overall objectives of the project

Unmet medical needs, chronic diseases, ageing populations, and the emergence of personalised medicine are amongst the factors contributing to a growing healthcare demand and continuing research for effective and safe innovative medicines. The development of new medicines is recognised as critical to advance improvements in healthcare. As most new medicines are developed by the pharmaceutical industry in collaboration with academic and healthcare organisations, clinical research programs are overseen by national and international regulatory bodies. However, the discovery and development of new medicines that are effective and safe for routine use in patients have become increasingly challenging.

Pharmaceutical innovation faces numerous R&D challenges causing significant study delays and increased costs. Importantly, over the last 12 years, the average cost of conducting clinical trials has increased three-fold. In 2005, the cost of researching, developing and achieving regulatory approval for a new chemical or biological entity was estimated at € 1.1 billion. The number of drug development programs has grown by an average of 6% per year from 2002 to 2011, with growth continuing through the recent economic downturn. In parallel, clinical research is evolving and growing in complexity and labour intensity. This is, in part, due to the need to conduct large clinical trials that provide definitive evidence of clinical benefits and safety, and to the ever increasing demand from regulators and payers to also generate value-based evidence which requires conducting further studies in order to assess the “real-world” comparative effectiveness, safety and cost-effectiveness of innovative medicines compared to existing therapies.

The main bottlenecks in current clinical research include sub-optimal protocol designs, slow and lengthy patient recruitment, and labour-intensive and time-consuming clinical study conduct. Specific issues relevant to conducting clinical trials include the difficulty in evaluating patient populations and in optimizing protocol design, the effort involved in identifying suitable patients for clinical trials, the manual and redundant re-entry of data, the reliability of data sources for clinical trials, and the difficulty in detecting and reporting infrequent adverse events.

The widespread adoption of EHR systems in Europe and worldwide represent vast, rich, and highly relevant health data sources which have the potential be reused for research, to address these bottle necks. There is a growing realisation that the ability to effectively integrate and inter-operate advanced EHR systems within health care networks for clinical research purposes represents a breakthrough opportunity to enhance academic research, to speed up and streamline existing processes and to build greater efficiency. Potential applications of interest include clinical trial feasibility, patient recruitment, clinical trial execution and drug surveillance reporting.

However, such developments require acceptance by patients, the public and the health service community. The lack of interoperability between EHR systems currently limits the ability to efficiently combine the data across large populations for research analysis. Key challenges need to be overcome, at a European scale, to provide a platform to support clinical research that functions across many EHR systems, complies with ethical, legal and privacy requirements that differ from country to country, and is sustainable through a scalable business model.

1.2. Overall deliverables of the project

The EHR4CR project (2011-2016) with a budget of +16 million Euros, has involved 35 academic and private partners (10 pharmaceutical companies) and is one of the largest of the IMI PPPs in this area. The consortium also included 11 hospital sites in France, Germany, Poland, Switzerland and the United Kingdom. It was part-sponsored by the European Commission through the Innovative Medicines Initiative (IMI).

The EHR4CR project has developed a robust and scalable platform that can utilise de-identified data from hospital EHR systems, in full compliance with the ethical, regulatory and data protection policies and requirements of each participating country. The EHR4CR platform supports distributed querying to assist in clinical trials feasibility assessment and patient recruitment. The platform can connect securely to the data within multiple hospital EHR systems and clinical data warehouses across Europe, to enable a trial sponsor to predict the number of eligible patients for a candidate clinical trial protocol, to assess its feasibility and to locate the most relevant hospital sites. Applications for internal use are offered to connected hospitals to assist them to efficiently identify and contact the patients who may be eligible for particular clinical trials. Contrary to other initiatives, EHR4CR designed a solution which is compliant to EU legislation and respects the position of hospital and patients. One of the key aspects is that patient level data never leaves the connected hospitals.

This development has required securing acceptance from the patients, the public and the research and health service communities. Therefore, in parallel to the technical developments, senior level decision makers, ethics boards and industry executives and scientists, have been involved in consultations to provide strategic insights into the most robust and acceptable technical and procedural approaches that should be taken to ensure privacy protection and compliance with European and national/regional regulations on data protection.

EHR4CR has shown that such a platform can significantly improve the efficiency of designing and conducting clinical trials, reducing time and costs, reducing administrative burdens, optimising protocol feasibility assessments, accelerating patient recruitment, making study conduct more efficient, enabling the participation of European hospitals in the more clinical trials and thereby potentially increasing research income.

The **European Institute for Innovation through Health Data** (i~HD) <http://www.i-hd.eu> is a not-for-profit organisation that has been established in 2015, arising in part out of the EHR4CR project, to develop and promote best practices in the governance, quality, semantic interoperability and uses of health data, including its reuse for research. An important role of i~HD is to provide independent governance oversight of clinical research platforms and their expanding networks of hospitals.

The first EHR4CR service provider, Custodix <https://www.custodix.com>, is now launching its operational platform, **InSite** (www.insiteplatform.com), for Europe-wide deployment, to be governed by i~HD. An early adopter *Champion Programme* has been launched as a first step in building a pan-European network connected to the InSite Platform. The objectives are to start building a network and community of hospitals open to data re-use for research, to further validate and improve the technology and to refine the business model, creating a win for all stakeholders. The *Champion Programme* serves at proving the value of Real World Data for clinical research and the InSite technology on a wide scale.

1.3. Summary of progress versus plan since last period

The final year of the project was undertaken through the granting of a one-year no cost extension to the original project term, specifically in order to prepare the ground for the launch of an early adopter programme of use of the sustainable and commercially viable platform, and to launch a complementary not for profit Institute.

Custodix, a Belgian SME and project partner, the first-appointed Service Provider by the consortium, was able to take a substantial portion of the tools and services developed in the project, combining these with the results of other recently completed European projects, in order to develop a commercial grade and robust platform, tools and services. This product, the InSite platform, has been made ready for initial contracted use, beyond the project, through an early adopter programme known as the *Champion Programme*.

By the end of the reporting period six of the original EFPIA partners from EHR4CR, plus one non-project Pharma company and one global CRO, had agreed to join the champion programme, licensing the use of the InSite platform and sponsoring the connection of additional hospitals across Europe. The research sponsors have agreed to “pool” the hospitals they each bring on board, thereby sharing this initial network of hospitals. They have furthermore agreed to collaborate in a joint evaluation of the success of the platform.

The second main objective of the no cost extension year was to enable the formation and launch of a new not-for-profit Institute, the European Institute for Innovation through Health Data (i~HD). Its articles of association were developed, and the International not-for-profit association was legally registered in Belgium during 2015. i~HD has participated in meetings and planning of the champion programme, since it will take responsibility for developing the codes of practice and standard operating rules governing this ecosystem, preparing the ground for a wider role in governing the use of clinical research platforms in general. It should be noted that compliance to the governing role is voluntary, although it is being specified formally within the contractual agreements signed as part of the champion programme. i~HD is starting to play other roles in promoting best practices in the uses of health data for clinical research and for clinical care, that are described elsewhere in this report and in final year project deliverables.

Another final year success was the completion and publication (or forthcoming publication) of the EHR4CR business model, a Cost Benefit Assessment and a Business Impact Analysis.

However, due to an unexpected reversal of the commitment of some EFPIA partners to convert their in-kind shortfall into in-cash, the project was suddenly placed in the position of being around €1.2 million short at the beginning of the fifth year. This impacted especially on the pilot sites, which had to stop work before they were able to undertake all of their intended evaluations, but it also had a knock-on effect on the engineering work plan. This has in particular impacted on the extent of the implementation and the extent of deployment in the pilot sites of the combined third and fourth scenarios, for clinical trial and execution and serious adverse event reporting. It is hoped to attract further research funding in future to enable the finalisation of these services, which were fully specified and designed, since there are clear value propositions for many stakeholders in delivering these solutions.

1.4. Significant achievements since last report

The significant achievements of the last year have largely been summarised in Section 1.3, and are detailed more fully within Section 1.5.

These achievements are:

1. Development of a commercial version of the clinical research platform (InSite), combining EHR4CR results with some from other European projects.
2. Launch of the Champion (early adopter) Programme of the commercial platform, during 2016-2017, and a funded plan to expand the network of connected hospitals to around 24.
3. Launch of the European Institute for Innovation through Health Data, a not-for-profit Institute to develop codes of practice, provide governance oversight and promote better quality and interoperability of health data, across Europe.
4. Completion and publication of work on the EHR4CR business model, Costs Benefit Assessment and Business Impact Analysis.
5. Significant effort in promoting the project, its results, and its sustainability strategy across Europe, and internationally, through publications, conferences and meetings with key decision makers.

1.5. Scientific and technical results/foregrounds of the project

Through a combination of a consortium that brought collectively many years of experience from previous relevant EU projects and the global conduct of clinical trials, an approach to ethics that has engaged many important stakeholders across Europe to ensure acceptability. This engagement has resulted in a robust iterative design methodology for the platform services that was anchored throughout on requirements and an underlying Service Oriented Architecture that has been designed to be scalable and adaptable, EHR4CR has successfully delivered a sound, useful and societally-acceptable pan-European solution for the reuse of hospital EHR information to support clinical research studies.

This has been achieved through a sophisticated work-plan covering many socio-technical areas, including engagement with key stakeholders, information governance, robust business modelling, scenario and requirements analysis, software engineering of many components, tools and services. All of this has been closely connected to eleven very engaged hospital pilot sites from across Europe, who contributed to all of these areas, provided deployment environments and conducted evaluations. The activities and results of these various areas of the work plan are summarised below.

Stakeholder engagement

The objective of this activity has been to determine and document the concerns, needs, opportunities and perceived challenges of the complex network of stakeholders impacted by EHR4CR. Incentives and disincentives for participation by each stakeholder in both the EHR4CR development stage and the long-term sustainability platform were also explored. The initial phase of

the work was to conduct an extensive pilot exercise in Scotland to develop a best practice approach for local national stakeholder identification, analysis engagement and management during the development phase of the EHR4CR project and to make recommendations to those who will be involved in rolling out these activities across the full European scope of the project. The Scottish Stakeholder Management pilot activity was completed and documented in D1.1. Stakeholders engaged with included senior NHS staff involved in patient care and research support, academic and NHS researchers, NHS IT staff, patients and carers, ethics committee members and medico-legal experts. Interactions ranged from individual face-to-face meetings through to and group meetings with the Scottish Health Informatics Programme. The meetings focused on the key objectives of EHR4CR: Protocol feasibility; Facilitation of patient identification and recruitment; Clinical trial execution, evaluation of the safety of medicines and business model development.

The Scottish experience was used by the University of Glasgow to develop a set of generic materials for use throughout Europe. These materials included: a project summary; a structured questionnaire in two versions targeting a) ethics committee members, b) all other stakeholders; recommended approaches for identifying and approaching stakeholders and for interview conduct. The Draft Generic Stakeholders Engagement Materials were presented at the EHR4CR meeting in Frankfurt, made available for comment on the EHR4CR SharePoint site and were presented and analysed at a special workshop on EHR4CR at a training meeting of 90 members of Scottish Ethics Committees.

An in depth interview survey was then conducted by senior members of the project consortium in their countries. Thirty-seven interviews were conducted, by telephone or in person, the interviews lasting between 60 and 90 minutes each. The interviewees included chairpersons of ethics committees, patient association leads, national policy makers and opinion leaders, and senior executives in health care provider organisations and academia. The results showed that there was strong support for the objectives of EHR4CR, with the strongest motivating factors for participation being perceived to be greater income generation from industry sponsored clinical trials and the ability to improve the efficiency of conducting clinical trials. There was recognition that there might be some concerns amongst patients and healthcare professionals, primarily about privacy protection, which would need to be addressed. Another concern which surfaced, to a moderate extent, was whether the key data fields required for successful patient recruitment would be available in most hospital EHRs. These results were presented in D1.2. A follow on survey with regulatory and EFPIA stakeholders was originally conceived and prepared, but in the end was not considered likely to yield significant new insights, and so was not pursued.

Scenario and requirement development for patient recruitment

The detailed description for the first scenario, protocol feasibility, was developed during 2011, initially by interviewing several EFPIA study protocol managers who were undertaking feasibility assessments by traditional means. These interviews helped to derive a generic (cross-company) workflow and to identify the key decision points and the way these are currently informed. Issues and bottlenecks were also identified. From these inputs a proposed new EHR4CR-enabled workflow was developed (see D1.1) and used as an input to workshops and iterative document reviews that eventually led to the production of a formal Software Requirements Specification. This SRS was presented to and formally reviewed by the public and EFPIA partners involved in WPG2 (dealing with

the design and implementation of the platform) and a “freeze candidate” version was eventually agreed in early 2012. This version was included in D1.1, and served as the requirements basis for the implementation of that service (Figure 1 shows an example Use Case diagram from that SRS.)

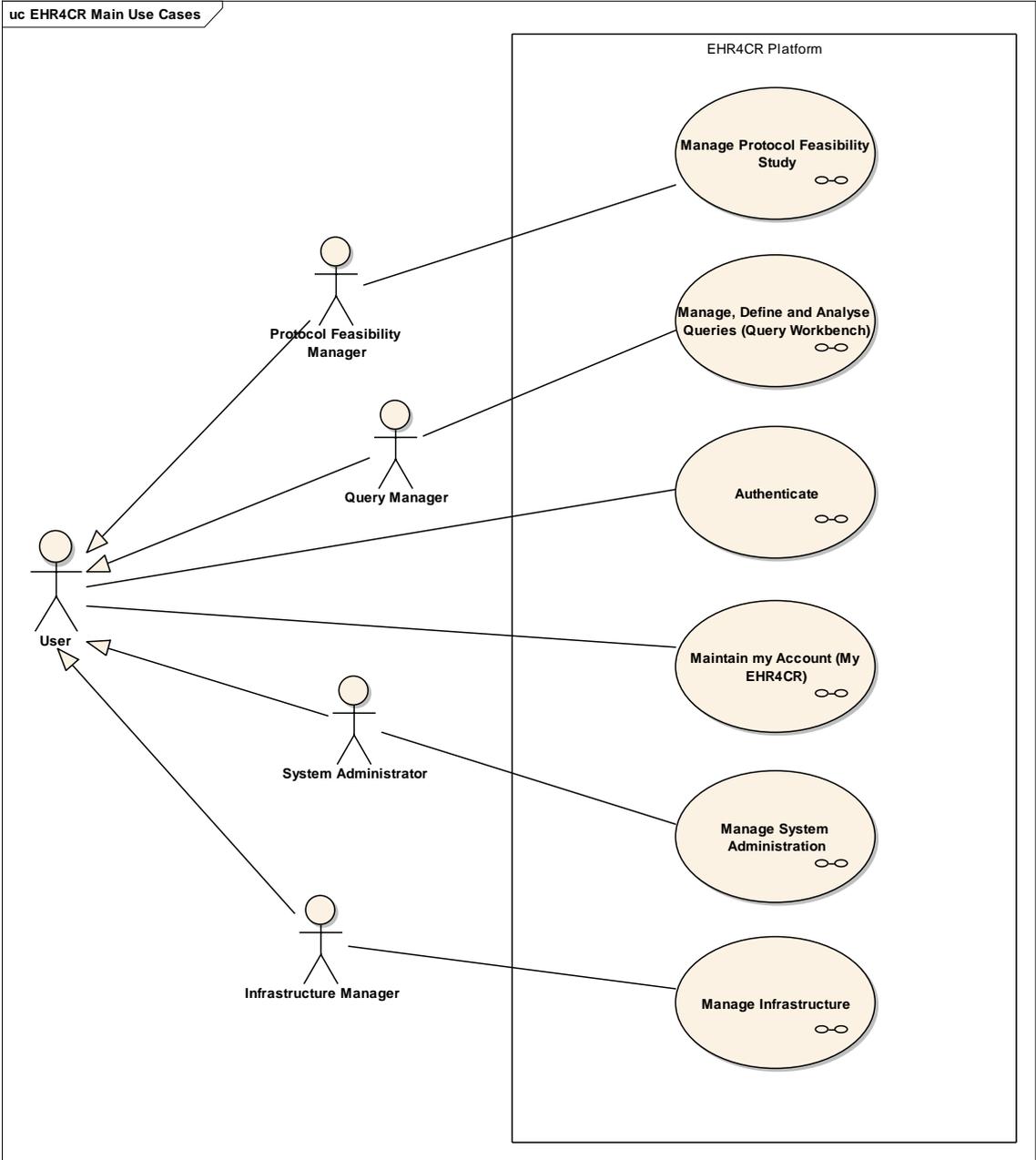


Figure 1: EHR4CR main use case for protocol feasibility

In March 2012, the formalisation of the second EHR4CR scenario was launched (Patient Identification and Recruitment). Several user stories were initially collected. A workshop held in Dusseldorf in April developed the in-depth requirements, resulting in a formal capability description that fed into the second Software Requirements Specification (SRS). The SRS was further developed at a second workshop in Paris involving the pilot sites, and has formalised requirements statements for 45 use

cases. The SRS successfully passed its final Stage Gate, which required minor updating and some restructuring. It was handed over to WPG2 for implementation, and included in D1.2.

All of the requirements for both Scenarios 1 and 2 were fed into an issue tracking system, JIRA, for easier maintenance and cross-checking with the implementation work plan. Relevant experience from other European projects was effectively sought (e.g. the Debug IT and the TRANSFoRM project).

In November 2013, the formalisation of the combined third & fourth EHR4CR scenario was launched (clinical trial execution & severe adverse event reporting). Several user stories were initially collected. A workshop held in Leverkusen in May 2013 developed the in-depth requirements, resulting in a formal capability description that fed into a Software Requirements Specification (SRS). The SRS was further developed at a second workshop in Leverkusen involving the pilot sites, and has formalised requirements statements for 28 use cases. The SRS successfully passed its final Stage Gate, which required minor updating and some restructuring. It was handed over to WPG2 for implementation in November 2013, and included within EHR4CR D1.3.

During the final years of the project the focus was on developing the appropriate framework for the certification of service providers with products that conform to these specifications. An analysis was undertaken of the certification frameworks of other comparative bodies within Europe, including EuroRec, ECRIN and UKCHIP, and an eventual model was developed. This clearly defines the role and workflow of a proposed Conformity Assessment body, the role it would play for different granularities of organisation, product or service being assessed, and an initial proposal for the fee structure. D1.4 presents an introduction to the approach being taken on conformity assessment, primarily reporting the rationale and approach, since the actual assessment criteria are still in development. The criteria are being derived from the three Software Requirements Specifications described above, drawing on prior work and expertise in this field from EuroRec and the eClinical Forum. This work is being taken forward during 2016 through the joint collaboration between The European Institute for Innovation through Health Data (see deliverable 9.19) and EuroRec.

Ethics and governance

The initial activities, undertaken mainly in year 1, aimed towards: (a) outlining the ethical and privacy issues and (b) identifying relevant regulations and legislative issues at EHR4CR pilot sites, (c) contributing to WP5 work regarding potential ethical implications. Work was therefore undertaken to compile the most relevant documents and regulations regarding ethical issues in clinical research. Secondly effort was focused on identifying local requirements in all EHR4CR pilot sites to guarantee smooth and uneventful implementation of EHR4CR. A major survey of the legislation in each of the pilot site countries was undertaken during year 2.

A template was used as the basic resource to collect information regarding ethics, legal, regulatory frame and privacy issues at the microenvironment of the pilot site hospitals. The areas in which the template focused were approvals, permissions, legal exemptions, data subjects, classes of information (identifiers, sensitive personal data, clinical data categories), de-identification policies and processes, pseudonymisation, longitudinal linkage management, re-identification, audit and information security policies. Templates were electronically sent to all pilot sites, and followed up through email and telephone interviews to enable their completion.

The main focus of the work on ethics and information governance in years 2-3 was to develop a pilot site information package. This was co-developed with the pilot sites. The highlight event was a June 2012 workshop in Berlin, comprising pilot site representatives and site Data Protection officers. They walked through the first two scenarios in detail and identified the information flows and protection measures that would satisfy their concerns. Version 1 of the pack was produced during summer 2012, and was used at the pilot sites when seeking approvals to connect local data sources to the first demonstrator version of the EHR4CR Protocol Feasibility Service. However, there were some critical issues that delayed agreement among the sites to release data, which prompted adding draft Standard Operating Rules to a new version of the pilot site information package. A code of conduct was also developed and drafted, with EFPIA partner involvement on the content to ensure acceptance. The project tracked the evolving landscape of the new Data Protection Regulation, helping prepare for platform compliance to meet the expected stipulations.

Throughout the IMI programme a need for a coordinated approach to address data protection concerns became more and more evident. During 2013 EHR4CR has lead a cross-project collaboration, through Sanofi and TMF, on ethics and data protection with a number of other projects from DG Connect (FP7) and IMI to tackle some common ethics and data protection challenges together (the 'Convergence Initiative'). Through this, early EHR4CR work on a data re-use code, initially in response to information governance needs within this project, was generalised to be applicable across all similar IMI & FP7 projects. This was shared with data protection officers at DG Connect and IMI as well as with several Pharma companies, and the ethics boards of other projects (e.g. eTRIKS, EMIF). The document was finalised through consensus building among ethics experts and then published by IMI as the Code of practice on secondary use of medical data in European scientific research projects.

The EHR4CR ethical and information governance needs were further formalised through the development of Standard Operating Rules to detail more explicitly than the above Code how each of the scenarios should be implemented in order to protect patient privacy.

The main focus of work in the final two years has been to finalise these Standard Operating Rules for the EHR4CR platform and services, for the first two scenarios which are intended to be deployed commercially. Through a combination of face-to-face meetings, teleconferences and document exchange the ethical and governance task force has specified the constraints and good practices that should be followed by the service provider of the platform and by hospital (data provider) and research users. A risk analysis of user roles was performed, and a detailed risk assessment was undertaken and used as a tool to verify if all of the necessary mitigation instruments are in place, and to identify gaps that will be filled during 2016. The ENSO funded additional task (Task 9.6) on consent and trust models is mentioned here as the work has interfaced with the other aspect of ethics and governance in WP1. Early on it became clear that a consent model would not be scalable for EHR4CR, which was confirmed by consultation with various patient representative groups within Europe. Most of the work therefore focused on the trust model, defining its key characteristics and the implications this has for the governance of the EHR4CR platform (published in D9.17). Members of the information governance task force also contributed to submissions made by partner institutions to the Article 29 Working Party.

The current working version of the EHR4CR governance principles and Standard Operating Rules are presented in D1.4. These are being finalised over the coming months and will be used to govern the Champion Programme (described in deliverable 8.7). They will be complemented by more detailed standard operating procedures, covering topics such as access controls, incident management, audit and monitoring.

Evaluation of the scenario implementations

The design and implementation of each of the scenarios was progressed as independent cycles: this proved helpful in enabling the project to experience the full cycle for Protocol Feasibility and using the lessons learned for downstream scenarios. During year 3 an evaluation was undertaken of the Protocol Feasibility deployment using replicated test data to validate that equivalent patient counts were returned from each hospital site in response to the same distributed query.

A usability evaluation of the query builder for the Protocol Feasibility scenario was undertaken during that year. The usability testing was conducted co-jointly by WP1 (the “Task 1.5 team”), WP7 & Custodix during Q4 2014 (from September 2014 to December 2014). A total of 38 testers participated to the evaluation conducted in two iterative rounds (22 for the first round vs. 16 for the second round). 7 EFPIA partners were represented (Amgen, AZ, Bayer, GSK, Lilly, Novartis & Sanofi), 2 academic institutions (AP-HP & HUG) and 5 countries (France, Germany, UK, Spain & Switzerland). Participants received self-training prior to the testing and then performed a script composed by 3 tasks to execute for an estimated period of one hour and a half. This study design allowed the team to assess if the training provided was adequate enough for the user to make proper use of the platform. A questionnaire was also submitted by the team to the testers, so at the end their input allowed Custodix to enhance the system accordingly. A publication was submitted to the BioMed Research International journal. Due to a delay in its implementation and deployment, the Patient Identification & Recruitment scenario evaluation was achieved in a total of 6 sites over the 11 that participated. Some partners were not able to participate in this evaluation because they were not able to find any appropriate active clinical study protocol at their site to evaluate. The methodology and initial results were presented in D1.3 and in D7.3.

D1.4 provides an overview of the evaluation methodology that was used for the three implemented scenarios, and presents the qualitative and quantitative results obtained (the learning outcomes). This complements the pilot site deliverable 7.4, which reports more on the status of deployment, but also include some aspects of the valuation is undertaken. D1.4 presents an evaluation of the protocol feasibility query builder, the evaluations that were possible to undertake using the patient recruitment tools and services that were available in early 2015, and the investigations that were undertaken to prepare the ground for the clinical trial execution (clinical trial data exchange) scenario.

At the time of writing this report it was hoped that further pilot site evaluation work could be undertaken, but in practice this was not possible due to an unexpected (substantial) funding shortfall. Some hoped-for final evaluations were therefore not undertaken during year 5.

Business model innovation

The initial activity was to undertake an environmental scan survey and to align/optimize the strategic planning activities relevant to designing a sustainable EHR4CR business model and value proposition. Following the EHR4CR kick-off meeting in Goteborg in March 2011, an EU electronic survey was developed and conducted in 2 waves (from June-September 2011) using a EHR4CR e-questionnaire i) with participating stakeholders from the public and private sectors involved in the EHR4CR consortium and ii) with non-participating informed stakeholders. The results were highly consistent between the two groups and confirmed a high interest towards the EHR4CR objectives and scenarios, as well as the relevance of developing a customized value proposition to address the respective needs of key stakeholders. In order to build further awareness and momentum in Europe, a scientific manuscript presenting the EHR4CR e-survey objectives, methods, results and conclusions was published in late 2011.

The strategic framework for the development of the EHR4CR Business Model and Value Proposition was developed in Q2.2011 and submitted to the consortium as a draft didactic document for internal dissemination and consolidation, with a particular focus on the PEST and SWOT analyses. The business model strategic approach and core building blocks were presented at the EHR4CR annual meeting in Frankfurt in October 2011. The strategic needs and broad framework for the development of a EHR4CR cost-benefit and budget impact assessment were addressed in Q3. 2011 for top up (ENSO) funding.

A EHR4CR business model innovation strategic forum (BMI-SF) was constituted in late 2011 to build awareness and advocacy amongst designated business leaders from participating EFPIA partners, and to gather their strategic expert guidance and business intelligence relevant to the design of a business model that will be sustainable and relevant to the pharmaceutical industry at project completion. The EHR4CR BMI-SF met for the first time in Frankfurt in October 2011. Following the meeting, participants received the BMI-SF meeting highlights and a business model questionnaire to provide further comments on the proposed vision, mission and values of the EHR4CR platform, PEST & SWOT considerations, key success factors, awareness-building strategies and representation at future meetings.

Considerable progress was made during year 2 through a more focused Business Model Innovation Task Force, comprising selected senior academic and industry members of the project. Consensus was reached on the sustainability organisations to be set up: (1) an *Institute*, to manage the standards and specifications for each of the four scenarios, how the platform components interact, the certification of systems, and the accreditation of service providers and data providers; (2) *Service Provider(s)* to operate platform services and possibly develop and/or licence independent implementations of the platform, initially arising from the project results, but with the model that would enable third parties to also become Service Provider(s).

During year 3 business model assumptions were developed, and presented as draft projections based on those assumptions to the consortium for feedback. After such feedback, work was undertaken to refine the underlying assumptions and to obtain relevant data. Business model projections, especially for Pharma, were then produced and evaluated through the BMI-SF, by senior executives from the EFPIA partners. A successful BMI-SF meeting was held in Basel, in November 2012; the EHR4CR Vision, Mission and Values, and the main features of the Business Model Canvas were presented, discussed and endorsed.

Using business modelling best practices and the perspective of a service provider, a sustainable business model was designed to establish how to best create, deliver and capture value from exploiting the EHR4CR platform and services. An EHR4CR business model simulation forecast the expenses, revenues, balance sheets, and profitability ratios that could be derived by service providers for exploiting the EHR4CR platform and services, confirming a significant and sustainable business potential in Europe, and beyond. This advanced simulation, and the results of the probabilistic sensitivity analyses performed, suggested that the EHR4CR business model appears profitable and sustainable over a 5-year time horizon, contingent upon the swift adoption of EHR4CR services at project completion, and upon achieving scalable market penetration. In parallel, a state-of-the-art EHR4CR cost-benefit assessment using the perspective of the pharmaceutical industry was undertaken to establish the added value of the EHR4CR services compared to current practices for this customer segment. A new not-for-profit Institute was proposed to establish the governance, promote the re-use of EHR data for clinical research, maintain specifications and standards, provide certification and accreditation services, meet the standards for ethics, governance, interoperability and quality of EHR4CR services, maintain an open source reference implementation, and provide oversight, training, and education, in compliance with legal and audit requirements.

In 2014, 4 Business Model Innovation Task Force (BMI-TF) workshops and Cost-Benefit Assessment (CBA) Expert panel meetings were organised. These undertook both the original intended year 4 work on sustainability and the ENSO funded Cost Benefit Assessment (Task 9.16). The work on sustainability built on the concepts developed during the third year of the project, of (i) an eventual commercial service provider to deploy and operate the platform and services, connect to hospitals and contract with research organisations to provide critical feasibility and patient recruitment support; (ii) a not for profit Institute to oversee the proper information flows and to actively promote the benefits and governance of the research uses of health data. Working also in collaboration with other EC projects with an interest in the Institute, the vision, mission, strategic objectives, governance model and operating model of the European Institute for Innovation through Health Data were worked out during the year, and have led to its formation as a legal entity.

A state-of-the-art cost-benefit assessment (CBA) using the perspective of pharmaceutical industry was conducted to assess the value of EHR4CR solutions compared to current practices. The CBA included the development of a core CBA model and analysis based on the actual person time, as well as the development of a confirmatory analysis using a complementary methodological approach based on R&D cycle time. The CBA methods and results were accepted for poster presentation at the 17th European ISPOR Conference (November 2014, Amsterdam). A scientific abstract and poster were developed and presented, and the abstract was published (see Figure 2). A full scientific manuscript has been developed for submission to a peer reviewed international journal.

The team also developed a comprehensive budget impact analysis (BIA) using the perspective of pharmaceutical industry to assess the budget impact of adopting EHR4CR solutions compared to current practices. The analysis was presented during EHR4CR webinars and the scientific manuscript has been submitted to a peer reviewed international journal.

This activity contributed to establishing the European Institute for Innovation through Health Data, including the definition of its vision & mission, scope, strategic objectives, governance model, and

organizational structure, and provided strategic guidance for the development of a sustainable business model and business plan.

Cost-Benefit Assessment of the Electronic Health Records for Clinical Research (EHR4CR) European Project

Beresniak Ariel¹, Schmidt Andreas², Provee Johann³, Bolanos Elena⁴, Patel Neelam⁵, Ammour Nadir⁶, Sundgren Mats⁷, Ericson Mats⁸, De Moor Georges⁹, Kalra Dipak¹⁰, Dupont Danielle¹

¹Data Mining International, Geneva, Switzerland, ²Hoffmann-La Roche Ltd, Basel, Switzerland, ³Bayer Healthcare, Leverkusen, Germany, ⁴Eli Lilly, Alcobendas, Spain, ⁵Eli Lilly, Surrey United Kingdom, ⁶Sanofi Aventis Chilly-Mazarin, France, ⁷AstraZeneca, Mølndal, Sweden, ⁸Amgen, Neuilly-sur-Seine, France, ⁹University of Ghent, Ghent, Belgium, ¹⁰The European Institute for Health Records (EuroRec), Ghent, Belgium

ABSTRACT

Objective

The widespread adoption of electronic health records (EHR) provides a unique and novel opportunity to improve the efficiency of clinical research processes. The European EHR4CR (Electronic Health Records for Clinical Research) 4-year project has developed a technological platform to enable the reuse of EHR data for clinical research. The objective of this cost-benefit assessment (CBA) is to assess the value of EHR4CR solutions compared to current practices.

Methods

Three clinical research scenarios were selected: Protocol feasibility assessment (S1), patient identification for recruitment (S2), and clinical study conduct (S3). The approach considered that the estimated reduction in actual person-time and costs of performing EHR4CR S1, S2, S3 applied to a Phase II or Phase III oncology clinical trial as a reference case would accelerate time to market (TTM). Probabilistic sensitivity analyses were conducted.

Results

When converting the efficiency gains realized with the EHR4CR platform into potential financial value for achieving faster TTM, the absolute mean cost-benefit for the global pharmaceutical oncology sector was estimated at €161m (S1), €46m (S2), €1904m (S3), €204m (S1+S2), and up to €2119m (S1+S2+S3) when the three scenarios were used sequentially.

Conclusions

The results confirm that optimizing clinical trial design and study conduct with the EHR4CR platform would generate substantial added value for pharmaceutical industry.

Disclosure

The EHR4CR project is mandated by the Innovative Medicines Initiative (IMI) co-funded by the European Commission and the European Federation of Pharmaceutical Industries and Associations (EFPIA).

BACKGROUND AND OBJECTIVE

Pharmaceutical innovation faces important research and development (R&D) challenges, including significant delays and escalating R&D costs.

- The average cost of developing a new drug is estimated at €2 billion¹.
- The average cost of clinical trials has increased three-fold over the last 12 years².
- Clinical development process is lengthy and can last 8 to 10 years^{3,4}.

Major R&D bottle necks include sub-optimal protocol designs, slow patient recruitment, and labor-intensive and time-consuming clinical study conduct.

- On average, there are 2-3 protocol amendments implemented per clinical trial. This number exceeds 3.5 for Phase III clinical trials⁵.
- Each protocol amendment takes an average additional 61 days to implement and costs over USD450,000⁶.
- Almost half of all trial delays are caused by participant recruitment problems⁶.
- 48% of sites miss their enrollment targets in Phase II and III multi-center study⁷.
- The percentages of studies that complete enrolment on time are 18 % in Europe, 17 % in Asia-Pacific, 15% in Latin America, and 7 % in the US⁸.
- Western Europe (69%) and Eastern Europe (75%) have the lowest achievement rates of targeted levels compared to North America (98%), Latin America and Asia Pacific⁷.
- Over 70% of data are duplicated between EHR and clinical trial systems⁹.

Pharmaceutical industry must transform its R&D processes to deliver innovative medicines more efficiently by:

- Improving the feasibility assessment of clinical trial protocol;
- Enabling the identification of suitable patients and speeding up their recruitment;
- Accelerating clinical study conduct, including serious adverse events (SAEs) reporting.

The expected benefits of enhancing these clinical research scenarios are summarized in Table I.

Table I. Expected benefits of improving clinical research scenarios

Clinical Research Scenarios	Expected Benefits
Protocol feasibility assessment	<ul style="list-style-type: none"> More patient-centric protocol Improved site and patient experience of study Avoidance or reduction of protocol amendments Increased probability of achieving the clinical trial objectives Reduced administrative burden of clinical research Significant reduction in actual person-time and costs Improved efficiency and reduced clinical trial cycle time
Patient identification and recruitment	<ul style="list-style-type: none"> Better and faster patient and clinical site targeting Accelerated patient recruitment Increased probability of achieving the clinical trial endpoints Reduced administrative burden and clerical tasks Significant reduction in actual person-time and costs Improved efficiency and reduced clinical trial cycle time Increased overall efficiency
Clinical study conduct and reporting of serious adverse events (SAE)	<ul style="list-style-type: none"> Clinical trial seamless execution (reduced paper work) High quality data and no data re-entry (reduced risk of error) Improved patient safety Significant reduction in actual person-time and costs Improved efficiency and reduced clinical trial cycle time

The EHR4CR European project, one of the largest public-private research partnership between the European Union and the EFPIA, has developed an advanced technological platform to enable the trustworthy reuse of hospital EHRs data for clinical research in Europe, and beyond. The objective of this CBA is to assess the potential value for pharmaceutical industry of adopting EHR4CR clinical research solutions versus current practices, as applied to oncology trials as a reference case.

CONCLUSIONS

By enabling the trustworthy reuse of hospital-based EHR patient-level data for clinical research, the EHR4CR breakthrough platform promises to transform clinical research environments, to enhance current practices, and to improve the overall efficiency of clinical research current frameworks. This CBA is the first study to assess the value of the EHR4CR platform compared to current practices. The results confirm that the EHR4CR-enabled clinical research scenarios, whether used individually, in combination, or in sequence within a clinical trial workflow, appear highly efficient, reducing the actual person-time and operational costs for conducting Phase II-III clinical trials in oncology as the reference case. Should the efficiency gains realized with the EHR4CR platform translate into achieving marketing authorization faster, and delivering innovative medicines to healthcare sooner, this economic evaluation establishes that the overall benefits and added value to global pharmaceutical industry would be substantial. While these results suggest an early adoption of EHR4CR solutions, further research is warranted to assess the EHR4CR value in real life context, once the platform has been fully deployed for enhancing clinical research, across clinical trial phases and therapeutic areas, in Europe, and beyond.

References

- EFPIA. The Pharmaceutical Industry in Figures. Key Data 2013. EFPIA Report 2013. Update 1-27.
- Wolster P. The cost of clinical trials. Drug Discovery & Development 2005; May 18.
- Tufts University. Rising protocol complexity, execution burden varies widely by phase and TA. C Tufts Center for the Study of Drug Development Impact Report 2010, 12.
- Mestre-Ferandez J, Sussex J, Traves A. The R&D cost of a new medicine. Office of Health Economics 2012.
- Alcumalde M. Aggregated EMR: Mitigating Trial Risk through Quality by Design Protocols. Applied Clinical Trials online Aug. 2014.
- Kalra D, Schmidt A, Potts HMM et al. Case Report from the EHR4CR Project. A European Survey on Electronic Health Records Systems for Clinical Research. #Health Connections 2011; 1: 108-13.
- Tufts University. Analysis and Insight into Critical Drug Development Issues, Impact report. Tufts Center for the Study of Drug Development 2013; 15: 1-4.
- Evaluate Pharma September 2013

Disclaimer

This project has received funding from the European Union Seventh Framework Program (FP7/2007-2013) under the Joint Undertaking Grant Agreement n° 115189 of the Innovative Medicines Initiative (IMI), a public-private partnership between the European Union and the European Federation of Pharmaceutical Industries and Associations (EFPIA). The sole responsibility for the content of this scientific communication lies with the authors and does not necessarily reflect the opinion of the European Union or EFPIA. The European Commission and EFPIA are not responsible for any use that may be made of the information contained therein.

METHODS

Using the perspective of pharmaceutical industry, for the first time, this CBA assesses the value of EHR4CR solutions compared to current practices for enhancing three clinical research scenarios (S), whether used separately, in combination, or in sequence, namely:

- Protocol feasibility assessment (S1),
- Patient identification and recruitment (S2), and
- Clinical study conduct and SAEs reporting (S3).

A resource utilization assessment was conducted by EFPIA partners to estimate the actual person-time and related costs for performing S1, S2, S3 under current practices and under EHR4CR conditions for a Phase II or Phase III clinical trial in oncology (reference case). For illustrative purposes, Table II summarizes the sum of the minimum and the sum of the maximum values of the estimated number of days and costs under current practices and under EHR4CR conditions.

Table II. Resource Utilization Assessment

Scenarios	Estimated # Days in Actual Person-Time		Estimated Costs (€)	
	Current Practices	EHR4CR	Current Practices	EHR4CR*
S1: Protocol Feasibility Assessment	148.5-389	73.3-194.5	177.9-346.9	177.9-346.9
S2: Patient Identification and Recruitment	39.9-100.3	20.2-46.2	100.0-200.0	100.0-200.0
S3: Study conduct and SAE	797.03-1000.0	408.03-600.0	1000.0-1000.0	1000.0-1000.0

Source: EFPIA Partners Resource Utilization Assessment (2013) * Excluding EHR4CR fees for service

The CBA assumptions (Table III) were validated by a multidisciplinary expert panel composed of academic experts, clinical research senior scientists (EFPIA partners), and expert health economists.

Table III. CBA Assumptions

- Global multicenter Phase II or Phase III oncology study
- Study conducted in 10-15 countries
- Involving 100-150 clinical centres
- Clinical study size: 450-1500 patients
- Full time equivalent (FTE) = 0.687/day (estimated weighted average daily wage)
- 1 FTE = 220 days
- 50% reduction in actual person-time and costs for S1,S2,S3 tasks impacted by the EHR4CR platform
- S1 assumes a reduction of 50% in the number of protocol amendments under EHR4CR conditions

Assuming that the estimated reduction in actual person-time and cost achieved with the EHR4CR platform would directly translate into accelerated TTM for oncology compounds, the potential mean benefits for the global pharmaceutical industry were derived by applying the full distribution of global market values (2012) of oncology products⁸. Absolute and relative CBA were conducted using probabilistic sensitivity analyses (10,000 Monte-Carlo simulations).

RESULTS

Compared to current practices, should the efficiency gains achieved with EHR4CR Scenarios 1, 2 and 3 used individually or in combination ultimately translate into achieving faster TTM, the corresponding estimated absolute cost-benefit for the global pharmaceutical oncology sector were derived, as described in Figure 1 and in Table IV.



Figure 1. Estimated cost-benefit of EHR4CR clinical research scenarios

Table IV. Cost-Benefit Assessment Results

Scenarios	Mean Costs Current Practices (SD)	Mean Costs EHR4CR* (SD)	Mean Benefit (€)	Absolute Cost-Benefit† (€)	Relative Cost-Benefit‡
(S1) Protocol Feasibility	183'959 (24'035)	216'491 (44'999)	161'522'390	-134'967'610	0.01
(S2) Patient identification and Recruitment	185'393 (51'078)	185'393 (51'078)	457'12'633	-171'733'367	0.02
(S3) Study conduct and SAE reporting	2'448'030 (690'712)	1'999'281 (368'549)	1'906'506'416	-541'523'584	0.01
S1+S2	-	402'111 (67'555)	204'573'175	-204'171'064	0.01
S1+S2+S3	-	1'999'281 (368'549)	2'121'810'208	-1'787'810'208	0.04

SD: Standard Deviation * Based on a 50% reduction in actual person-time. Includes estimated EHR4CR fees for service † Absolute CBA (estimated costs minus benefits) ‡ Relative CBA (estimated costs divided by benefits)

Figure 2: ISPOR CBA POSTER

Architecture and Integration

The EHR4CR architecture is described by an Architecture Document (AD) structured according to the methodology described by ISO/IEC 42010. The AD document has been iteratively extended throughout the course of the project and yearly snapshots have been used to produce the WP3 deliverables. The key underlying principles of the EHR4CR Service Oriented Architecture (SOA) are loose-coupling, the use of formal contracts and abstractions, reusability and autonomy of services and more specifically the ability to run services next to the clinical data inside the hospital. Evidence that the EHR4CR architecture succeeded in achieving a high level of reusability can be found in the fact that all four scenarios are relying on a large set of common modules and services, including most notably the EHR4CR query engine, authentication and authorisation modules and terminology mapping services.

Key to the SOA paradigm is the ability to deploy and publish Web Service endpoints hereby advertising relevant functional and non-functional metadata, to discover suitable Web Service endpoints and finally to bind to them, hereby fulfilling any technical constraints – such as security requirements - that might apply as described by the service's metadata. The EHR4CR platform relies on a central technical registry allowing platform service providers, be it application providers or data providers, to advertise Web Service Endpoints supporting various capabilities of the four EHR4CR scenarios. The EHR4CR registry is based on industry standards (OASIS UDDI v3) and profiles. The metadata information model used by EHR4CR for dynamic service discovery encompasses such selection criteria as: site and country identification, contact identification, Web Service security requirements and capabilities, Web Service interface specifications and bindings and clinical site-related capabilities such as available medical equipment, staff and facilities (useful for clinical site selection in the four scenarios).

Because of its distributed nature, with much of the data integration and querying capabilities taking place inside the hospitals acting as data providers, the EHR4CR architecture is tailored towards easy clinical site adoption. The EHR4CR platform provides an asynchronous message broker service allowing service providers and service consumers to interact without the need for service providers (the hospitals) to expose their Web Services to the Internet. Instead, a connected site's Web Services can securely connect to the EHR4CR message broker to receive incoming requests and to answer them with responses. The availability of this infrastructure service, together with the fact that communication is performed over the standard (and thus firewall-friendly) HTTP(S) protocol, makes a deployment in which all sensitive data is maintained over full control of the clinical site realistically achievable and ensures a buy-in from the IT departments of the hospitals. Again, the EHR4CR platform relies on industry standards to achieve this (SOAP over JMS protocol, JMS service using HTTP(S) as a transport protocol). Further, the use of an asynchronous messaging platform facilitates request buffering and thus allows clinical sites to operate the EHR4CR software with minimal system requirements.

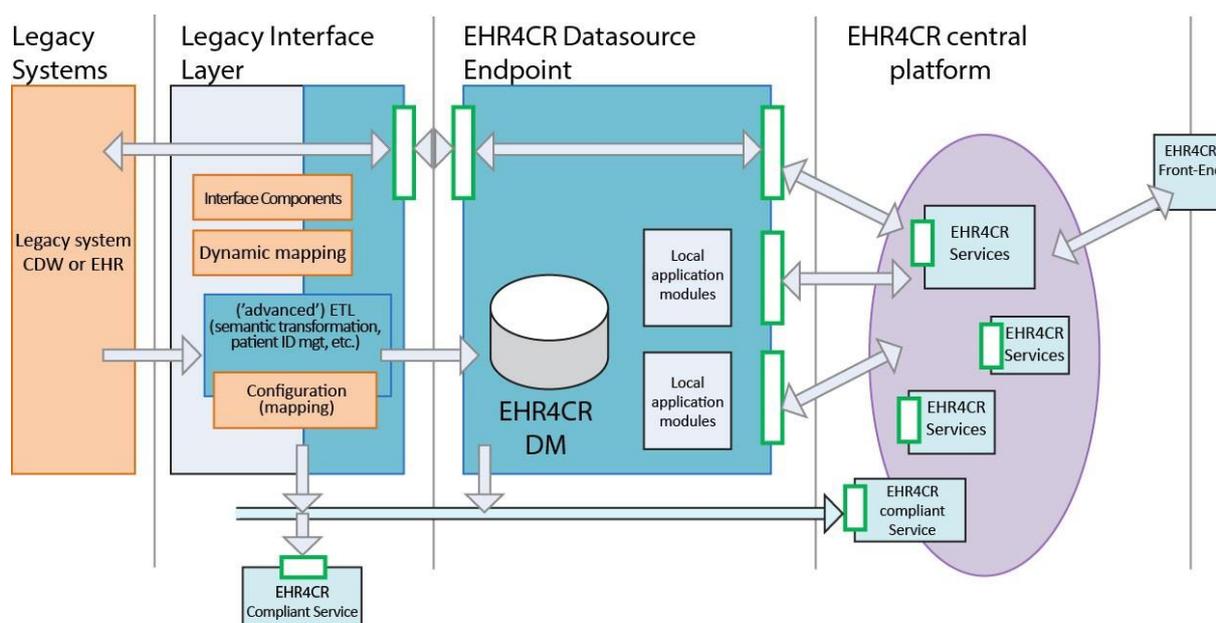


Figure 3: Simplified view of the EHR4CR Platform

The EHR4CR architecture description specifies the Web Service interfaces to be used and describes the workflows to engage in from a technical perspective in order to achieve technical compliance for each of the four scenarios. This is achieved relying fully on industry standards where possible and additionally constraining them where needed into technical profiles to be adopted in order to achieve ‘EHR4CR compliance’ with minimal effort needed from service providers. For example if the standards offers a range of options that are all to be implemented for being able to claim full compatibility or adherence, the EHR4CR architecture specifies the minimal supported functionality for achieving interoperability over the EHR4CR platform.

With respect to interfacing with existing external systems such as Electronic Health Record (EHR) systems, Clinical Data Warehouses (CDWs) and Clinical Data Management Systems (CDMSs), the EHR4CR architecture imposes the smallest possible set of technical requirements for achieving minimal functional integration with the EHR4CR software and data models. As such it is possible to tailor – by configuration - a specific EHR4CR installation for use with a wide set of legacy systems. Given the observation that the EHR4CR pilot sites are operating either under a data warehouse model developed within EHR4CR (joint effort of WP4 and WP7) or the i2b2 model, out-of-the-box integration is offered for these systems. Given the wide variety of existing EHR and Electronic Data Capture (EDC) systems, it is not feasible to provide out-of-the-box integration with the EHR4CR software. However, in order to maximize the number of EHR and EDC systems the EHR4CR software can integrate with minimal required effort, support for the third and fourth scenarios (requiring interaction with EHR and EDC systems) is fully based on IHE and CDISC specifications to which a large set of vendors are already (partially) adhering or planning to adhere.

The EHR4CR development activities have been increasingly streamlined over the course of the project, starting out at the beginning of the project with a minimal set of rules and guidelines for developers, resulting in the final year into an extensive integrated development environment addressing a wide range of industry best-practices including adoption of an agile development methodology (SCRUM), source control (subversion), continuous integration (Jenkins), automated testing (JUnit, Mockito, Selenium WebDriver), issue tracking (JIRA), build automation (Maven),

dependency management (Artifactory) and multiple deployment environments (staging environment for developer testing, production environment for the pilot activities).

A key achievement of the final project year is the delivery of a fully standardised reference software stack (in the form of a virtual machine) to support efficient deployment at future hospital sites. Our experience shows that installation of the reference software stack inside a new hospital data centre can be achieved in one day while requiring just a few hours of effort from local IT staff. Connection of hospital data sources on the ETL pipeline can be achieved with minimal effort from the hospital staff (typically less than one week of a single person's time).

Privacy and security

Patient privacy protection is one of the core principles of the EHR4CR project and by consequence the EHR4CR technical platform. This principle is consequently applied throughout the architecture of the EHR4CR platform that has been guided by the *privacy-by-design* principle. This is reflected through the following architecture views and viewpoints:

- Information view (Interface design): no EHR4CR applications outside the hospital network shall have access to patient level information.
- Deployment view: Patient-level clinical data and the software operating on it must strictly be deployed inside the boundaries of the hospital network.
- Information view: personally identifiable information residing in the clinical data warehouse is pseudonymised by default and controlled re-identification can only occur through health care professionals having a trust (care) relationship with the patient. Only after explicit approval of the patient will his/her identifying medical information be accessible by investigators or other authorised personnel.
- Security viewpoint: Services offering platform-level access to aggregated patient data shall apply additional measures to avoid (partial) re-identification through inference attacks (e.g. by using overlapping queries to single out an individual patient). All access to individual patient level information is subject to audit logging and audit logs refer to the end-user on whose behalf the patient level information is being accessed (even though the end-user may not be able to directly see it – only the aggregated result).

The above measures are the result of translating the security and privacy requirements gathered through feedback from the pilots into technical requirements and combining them with other technical requirements arising from all related design and development activities.

The EHR4CR platform relies on a number of core security services in order to facilitate interaction between end-users, the platform applications and the back-end services they are calling, including the data access services running in the hospital environments. In addition, they are required in order to support authorised pseudonymisation and controlled re-identification of patient identifying information. The EHR4CR core security services are based on widely adopted industry standards such as SAML, WS-Security, WS-Trust, XACML and XDAS. Given that these standards define a wide scope of alternatives to achieve the same underlying security goals, the EHR4CR security architecture

defines integration profiles to narrow down these alternatives and thus achieve interoperability with a minimal amount of development and configuration effort.

The underlying trust model of the EHR4CR platform relies on a centrally administered authority vouching for the trustworthiness of enrolled organisation and end-users operating on the platform. Role and privilege-based access control can be ensured by relying on the platform's central authorisation attribute and/or authorisation decision authority.

Integration of platform services and applications with the core security services is facilitated by a security integration software developer manual and the availability of several security integration applications and modules, including:

- a configurable reverse web proxy for enabling EHR4CR Single-Sign-On and credential delegation on web application (e.g. used for the EHR4CR central workbench),
- a library for securing SOAP Web Services and their clients with the central EHR4CR authentication services,
- a library for enabling policy-based authorisation on Web Services using the EHR4CR central authorisation service,
- a library for creating EHR4CR compliant audit logs and submitting them to a central audit bus allowing for user-centric and data(-subject)-centric audit trail reconstruction,
- a module for tracking data provenance information in the query engine, etc.

Next to the *privacy-by-design* principle, the EHR4CR architecture and reference implementation rely on pseudonymisation services where interaction with patient level information is needed (e.g. in the query engine querying the local clinical data warehouse). The EHR4CR pseudonymisation services are integrated with the EHR4CR core security services to ensure that pseudonymisation and re-identification requests are authorised and audited. Their capabilities include:

- Configurable pseudonymisation of Personal Identifying Information (PII)
- Replacement of identifiers (names, addresses, locations, dates) in free-text with pseudonyms or anonymised placeholders.
- Master Patient Index (MPI) capabilities for linking patient identifiers relating to the same patient and for tracking study-specific (pseudonymous) ids.

The pseudonymisation services are invoked as follows in the various EHR4CR workflows:

- The ETL pipeline contains a step to pseudonymise patient identifying information such that the clinical data warehouse will contain pseudonyms rather than patient identifiers. Next to calculating pseudonyms, the pseudonymisation service used in the EHR4CR reference implementation is capable of performing other privacy preserving tasks such as replacing names, dates and locations in free-text, encrypting sensitive attributes etc.

- In the Patient identification and Recruitment Scenario (PRS), when an authorised user is presented with possible candidate patients for recruitment (the treating physician by default), the local workbench tool will request re-identification of the patient pseudonym on his/her behalf.
- Upon recruitment of a patient, the PRS workflow will capture the study identifier under which the patient is registered in the clinical study and register that study identifier in the Master Patient Index (MPI).
- In the Clinical Trial Execution (CTE) and Adverse Event Reporting (AER) scenarios, when an eCRF form is retrieved through the EHR4CR software, the workflow engine will use the patient study id to request the corresponding local patient id from the MPI and use that information to invoke the EHR4CR query engine to retrieve relevant clinical facts from the clinical data warehouse for eCRF form auto-population.

Semantic interoperability services

The objective of the Semantic Interoperability Services is to allow:

- Clinicians in hospitals (data providers of the EHR4CR network), while using their own words, to simultaneously utilize the most appropriate reference codes for meaningful re-use of routinely collected clinical data in electronic healthcare records (EHRs) in the context of clinical research conducted at an international level.
- Investigators of the EHR4CR network to use a semantically-enabled platform to efficiently perform sophisticated web searches across European hospitals, to find clinically relevant results that can help improve clinical research.

The clinical terms normally used by clinicians are usually mapped to – often local - coding terminologies used locally for care coordination and secondary use of the clinical content. These local coding terminologies do not necessarily match with international administrative and clinical reference terminologies – such as ICD-10-CM, SNOMED CT®, LOINC®, ATC, etc. – used within the EHR4CR European network. The aim is that clinicians in hospitals can go on capturing, storing and searching their clinical content according to local terminologies while providing to the EHR4CR users a cross-border access to this important clinical information according to international reference terminologies.

In addition to maintaining a wide range of curated semantic resources (healthcare template/data elements/value sets and terminologies) the EHR4CR semantic interoperability platform also created tools and services to support the mapping between local terminologies used in the hospitals and reference terminologies used in EHR4CR queries.

EHR4CR semantic Interoperability solution has been designed and implemented to support the different actors in accomplishing their tasks within the standardization process and EHR4CR use case execution. Tools and services are used for i) authoring and maintaining the shared semantic resources of the EHR4CR mediation model and ii) supporting the definition of query specifications in the context of the EHR4CR use cases.

The EHR4CR project developed a semantic interoperability platform providing a consistent integrative semantic abstraction on top of existing application representations that enables to

mediate across heterogeneous applications - Electronic health records (EHRs) and Clinical Data Warehouses (CDWs) – storing routinely collected clinical data at hospital sites. A mediation model provides a homogeneous view of the clinical data contained within disparate databases of data providers so that data users can access these data using a library of standard queries that have been written based on the mediation model.

Electronic health records (EHR) support insurance reimbursement processes and clinical practice at the point of care. Each has different logical organizations and physical formats, and the terminologies used to describe the clinical information conditions vary from source to source. Clinical Data Warehouses (CDWs) support secondary use of clinical data and allow users to generate evidence from a wide variety of sources and support collaborative research across data sources both within and outside the hospitals. Clinical Data Warehouses (CDWs) also implement various information models and terminology models.

EHR4CR faces the challenge of improving semantic interoperability of clinical information in order to better leverage routinely collected clinical data in electronic healthcare records (EHRs) during the execution of clinical trials.

The EHR4CR Common Information Model (CIM) is a standard-based expressive and scalable mediation model, allowing dynamic mappings between data structures and semantics for consistent interpretation of clinical data accessed from varying sources. The approach is based on the realistic assumption that the co-existence between several standard semantic artefacts - namely information models (e.g. EN ISO 13606 information model and archetypes, openEHR, HL7 RIM, C-CDA and FHIR specifications, CDISC ODM, etc.) and terminologies/ontologies (e.g. LOINC, ATC, SNOMED CT, etc.) – as well as proprietary implementations for representing the content of health information in systems (EHR systems, CDWs, CTMS, EDC systems, etc.) will endure. Therefore achieving broad-based, scalable and computable semantic interoperability across multiple domains and systems requires a consistent use of multiple standards, clinical information models and terminology models. The EHR4CR project provides a mediation model – the EHR4CR Common Information Model consisting in a set of multilingual semantic resources based on multiple standards.

The EHR4CR Common Information Model (mediation model) has been developed, and can be extended, through a global consensus-based development process² in order to cover the scope of both i) eligibility criteria and data items identified from a given set of specific clinical trials (bottom up approach) and ii) standards reference clinical information models (top down approach). The EHR4CR Common Information Model is developed and evolves through repeated cycles using a "Learning by Doing" approach.

The *EHR4CR Common Information Model (CIM)* consists in a set of multilingual semantic resources based on multiple standards. The EHR4CR templates are based on FHIR resources (Patient, Encounter, Condition, Observation, Procedure and Medication Statement. FHIR-based resources were organized into categories based on HL7 CCD sections and UMLS semantic types: Demographics, Encounters, Advance directives, Problems, Family History, Social History, Alerts, Medications, Immunizations, Vital Signs, Results (lab, anatomic pathology), Procedures, Plan of Care, Lifestyle Choice, Ethical consideration. FHIR resources were enriched in order to fulfil the requirements of the

² Defined consistently with the governance principles defined by CDISC SHARE

project and represent the required semantic content. Some specific value sets were defined for some data elements of the FHIR templates.

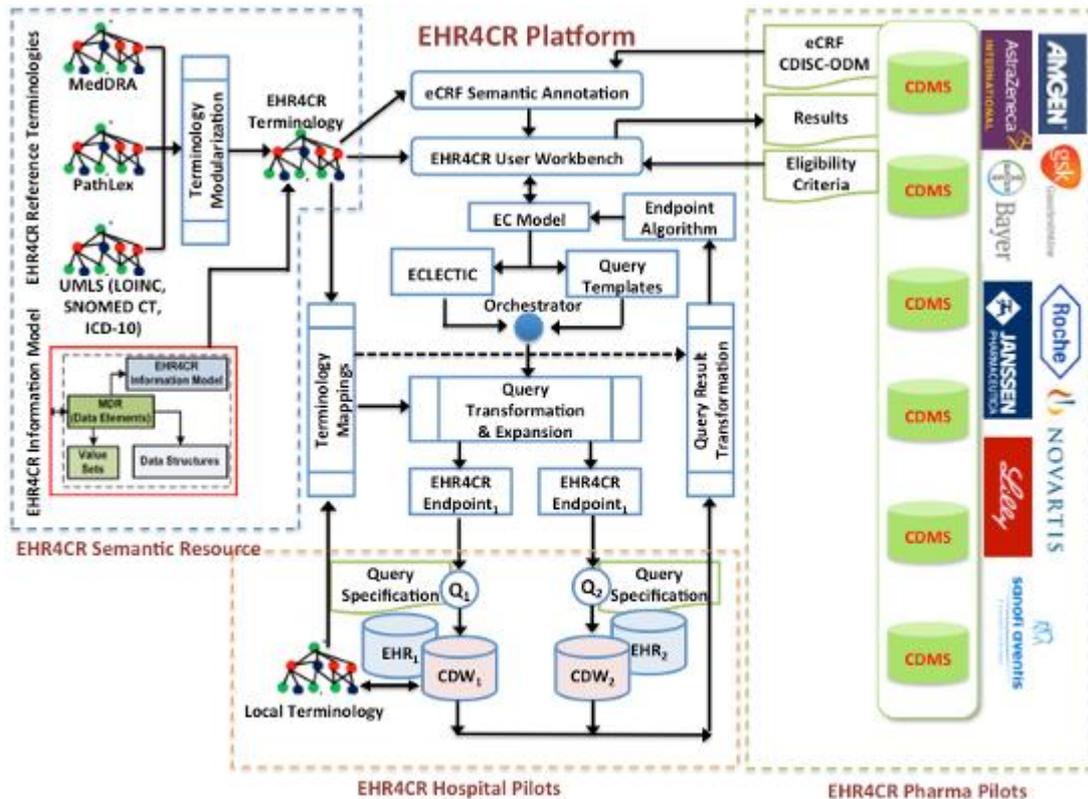


Figure 4. EHR4CR Semantic Interoperability platform: a set of EHR4CR Semantic Resources and Semantic Interoperability Services (SIS) are used during EHR4CR use case execution.

As much as possible existing resources are imported especially reference biomedical terminologies/ontologies. Some of external resources are overlapping (e.g. ICD-10 and SNOMED CT; MedDRA and SNOMED CT; NCI-T and SNOMED CT). Associations between these reference terminologies are available in UMLS. Some of the external resources need to be translated and/or extended, EHR4CR translations/extensions need to be captured and managed. At last, some specific resources need to be created. An EHR4CR terminology was created in order to create concepts that are in the scope of the project but do not exist in the selected reference terminologies. We integrated the UMLS CUI in order to allow multi-terminology binding.

Once hospital CDWs/EHRs are connected to the EHR4CR platform and source information models mapped to the EHR4CR Common Information Model, distributed queries can be specified based on the EHR4CR Common Information Model and executed over heterogeneous sources. Routinely collected clinical data can be used at different key points in trial design and execution life-cycle.

Platform services and tools

Underpinned by the architecture, privacy and security protection components and interfacing with the semantic interoperability components, a set of tools and services allow end-users to construct and execute queries, distributed to multiple connected hospitals across Europe, and review the

results of those queries. As an example, the way in which protocol feasibility services (PFS) are run is summarised below.

The PFS demonstrator includes a workbench application that allows the authoring and execution of computable Eligibility Criteria (EC) queries and allows secured sharing of feasibility studies and the associated EC queries amongst different platform users. EC queries can be built using a user-friendly graphical user interface which allows specifying Boolean and temporal constraints between individual EC (See Figure 5).

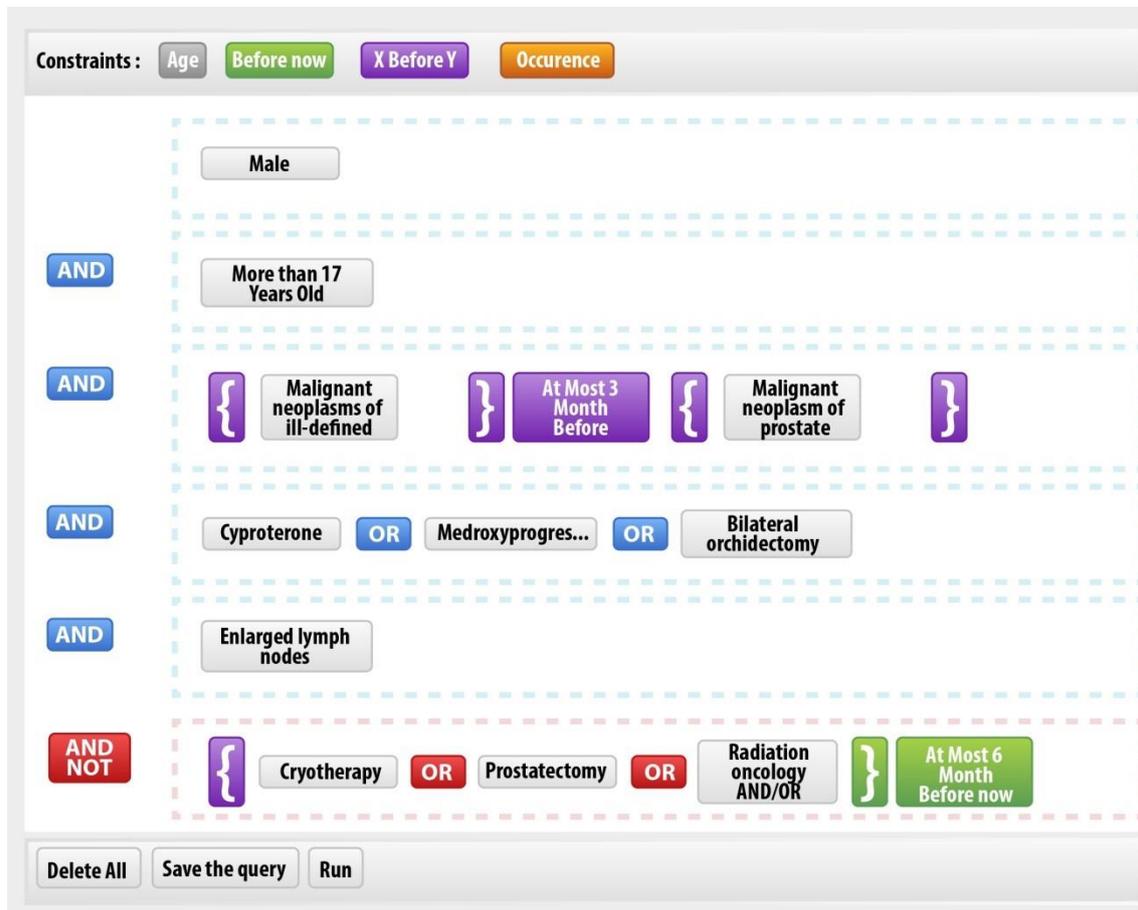


Figure 5: PFS query builder graphical user interface

After running an EC query, the results can be visualised by showing the overall results with the possibility to access break-downs on the patient demographics (age categories and gender) level, the individual eligibility criterion level as well as the results returned by the individual sites (Figure 6).

Result:

Aggregated (384)

384 patients match all criteria

Age	Male	Female	Undefined	Total
0 to 9	0	0	0	0
10 to 19	0	0	0	0
20 to 29	0	0	0	0
30 to 39	5	1	0	6
40 to 49	8	8	0	16
50 to 59	35	19	0	54
60 to 69	56	21	0	77
70 to 79	105	54	0	159
80 +	37	35	0	72

23717 patients match: born() at least 18 year before now

3253 patients match: first diagnosis((ICD-10:E11,"Non-insulin-dependent diabetes mellitus"))

827 patients match: last numericstatus((SNOMED Clinical Terms:60621009,"Body mass index")) in range (>=25.0) unit((ucum:kg/m2,"kilogram per square meter"))

1368 patients match: last numericstatus((SNOMED Clinical Terms:60621009,"Body mass index")) in range (<=40.0) unit((ucum:kg/m2,"kilogram per square meter"))

5006 patients match: last numericstatus((LOINC:4548-4,"Hemoglobin A1c/Hemoglobin.total:Mass Fraction:Point in time:Whole blood:Quantitative")) in range(>=0.0) unit((ucum:%,"percent"))

23262 patients match: not first diagnosis((ICD-10:I21,"Acute myocardial infarction"))

24031 patients match: not last medication((ATC:A10A,"INSULINS AND ANALOGUES [atc:A10A]")) at most 12 month before now

24031 patients match: not last numericstatus((LOINC:1742-6,"Alanine aminotransferase:Catalytic Concentration:Point in time:Serum/Plasma:Quantitative")) in range(>=96.0) unit((ucum:[U]/L,"international unit per liter"))

WWU (384)

Figure 6: Screenshot of a query result

For the patient recruitment service (PRS) this workbench was extended to include recruitment study coordination functions and a corresponding dashboard showing the current recruitment and accrual status at each of the clinical sites that have been invited to participate in a given study.

In order to start the recruitment process for a given study, a new study definition must be created by the study manager. The definition includes the protocol description and optionally the formal eligibility criteria to allow computer-assisted checking of patient eligibility. The study definition can be based on an existing study definition previously created through the PFS or it can be newly created if protocol feasibility checking for the given study has not been previously conducted on the platform. The study definition can also be based on an existing CDISC SDM (Study Design Model) file. The formal eligibility criteria defined for the PFS can be extended and enhanced to be used for the PRS.

Through the registry service, the study manager is able to select clinical sites of interest that expose the necessary technical interface. Following this an invitation containing the study definition will be sent to each of the selected sites. The study definition will be imported in the local study repository and the invitation will eventually be presented to the *data relationship manager* responsible for engaging the clinical site in semi-automated studies.

Once a clinical site has been invited to participate in a given study for recruitment, its participation status will be visible to the study manager. Once the clinical site accepts to participate, the number of patients in each of the various recruitment stages will be periodically made available to the study manager (Figure 7).

Entity	Status	Last Update	Potential	Candidate	Notified	Consenting	Screening	Denied	Included	Expand
Total	-	2013-10-26 08:54:38	224	390	222	363	345	244	331	[-]
FR	-	2013-10-26 08:54:38	118	64	71	37	150	29	106	[-]
Assistance Publique Hôpitaux de Paris	Running	2013-10-26 08:54:38	32	35	56	0	72	25	27	
CHU-de Rennes	Running	2013-10-26 08:54:38	66	29	15	37	78	4	79	
GB	-	2013-10-26 08:54:38	106	326	151	326	195	215	225	[+]

Figure 7: Screenshot of the clinical site recruitment status dashboard

For the clinical sites, an entirely new application was designed and implemented to support data relationship management (participations in clinical studies), local study management (user assignments and study status) and candidate patient identification and patient recruitment status tracking. After the site accepts to participate in a given study, the Principal Investigator can create a selection containing potential candidate patients to be recruited. If the study participation request includes a formal representation of the EC, these can be used at the clinical site to automatically query the data access endpoint to populate the initial list of (potential) candidate patients (computer-assisted candidate selection). The initial candidate patient list will be based on pseudonimised records and patient identifying information will not be visible until a treating physician has contacted the patient and if the patient agrees to enter the enrollment process.

Pilot sites and evaluations

The EHR4CR platform has been evaluated by demonstrating the functionality of the tools and services. These evaluations occurred at several large academic hospitals, interfacing with EHR systems, with a specific focus towards a set of medical domains mutually agreed between the pilot sites and EFPIA partners. The EHR4CR project primarily addressed the following disease areas included in the pilots: oncology, inflammatory diseases, neuroscience, diabetes, cardiovascular and respiratory diseases. These disease areas are relevant to pharmaceutical industry pipelines, and align with clinical research interests and data resources at the pilot sites.

The overall objective of the pilot site work package was to demonstrate the functionality of the tools and services provided by the platform and to evaluate the EHR4CR platform in the areas of clinical study design, execution and serious adverse event reporting with a specific focus towards a set of mutually acceptable medical domains agreed on by the demonstrator sites and EFPIA. The platform was piloted at 11 different data provider sites.

Interfaces between the EHR systems and the central EHR4CR platform were established. An inventory of data elements for pilot studies was defined. Semantic mapping between local terminologies and the central EHR4CR terminology was undertaken. Clinical data warehouses (CDWs), compliant with the EHR4CR platform and the associated extract-transform-load (ETL) processes were designed and tested. Approval of all data processing steps was gained in accordance with local ethical and legal regulations at each site.

All pilot sites installed a local endpoint with connection to a local clinical data warehouse, and eleven data provider sites in five countries are connected to the EHR4CR platform (see Figure 8 as an example for the PFS scenario).

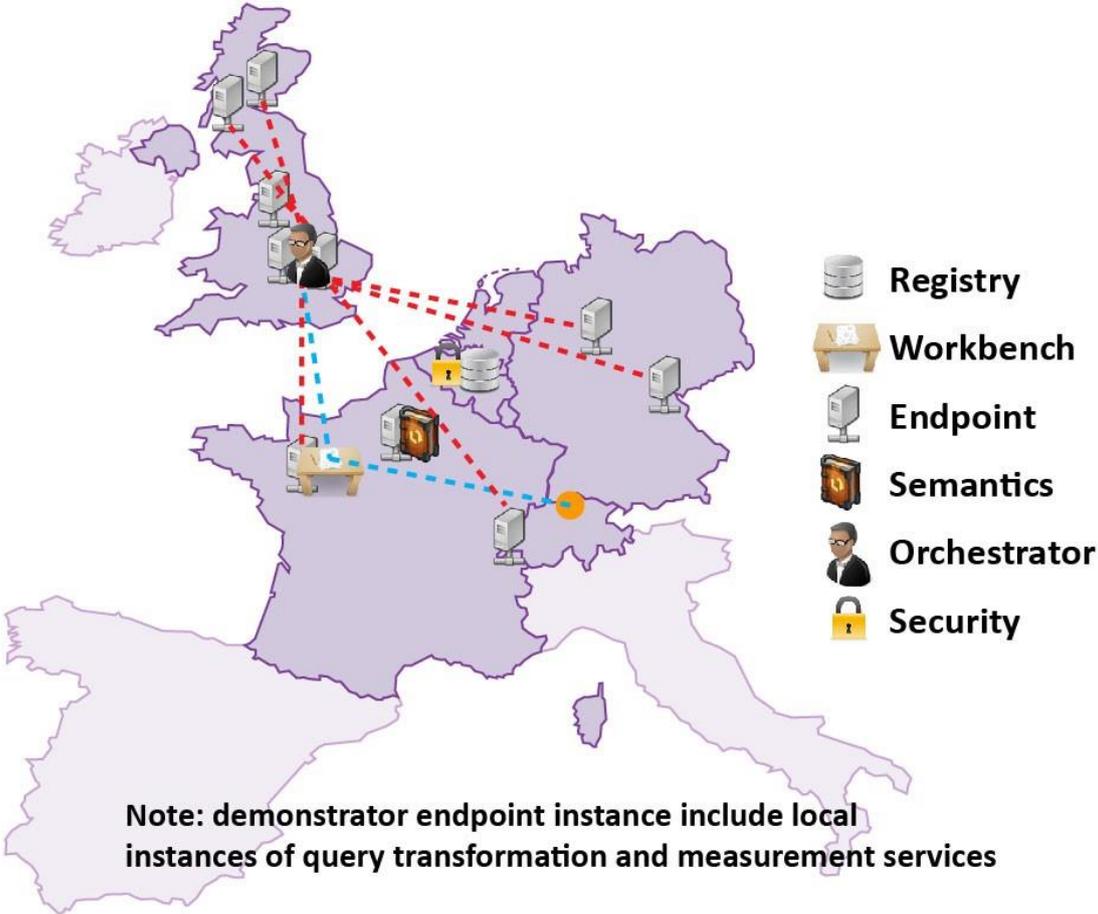


Figure 8: Overview of the main services of the Protocol Feasibility Scenario

The piloting was divided into three scenarios: protocol feasibility, patient identification and recruitment, clinical trial execution including serious adverse event reporting. Several scientific reports regarding various aspects have been published.

Data Inventories have been defined with data elements that are important for EFPIA and are available in European EHR systems for PFS, PRS and CTE/SAE. An initial top list of data elements containing 75 EHR data elements was identified by comparing common eligibility criteria used by EHR4CR

EFPIA partners at the pilot sites with available data elements in the EHR/CDW and EDC systems (see Figure 9). In addition, a wish-list was drafted of a further 21 data elements which were not available at more than 50% of the sites but deemed important.

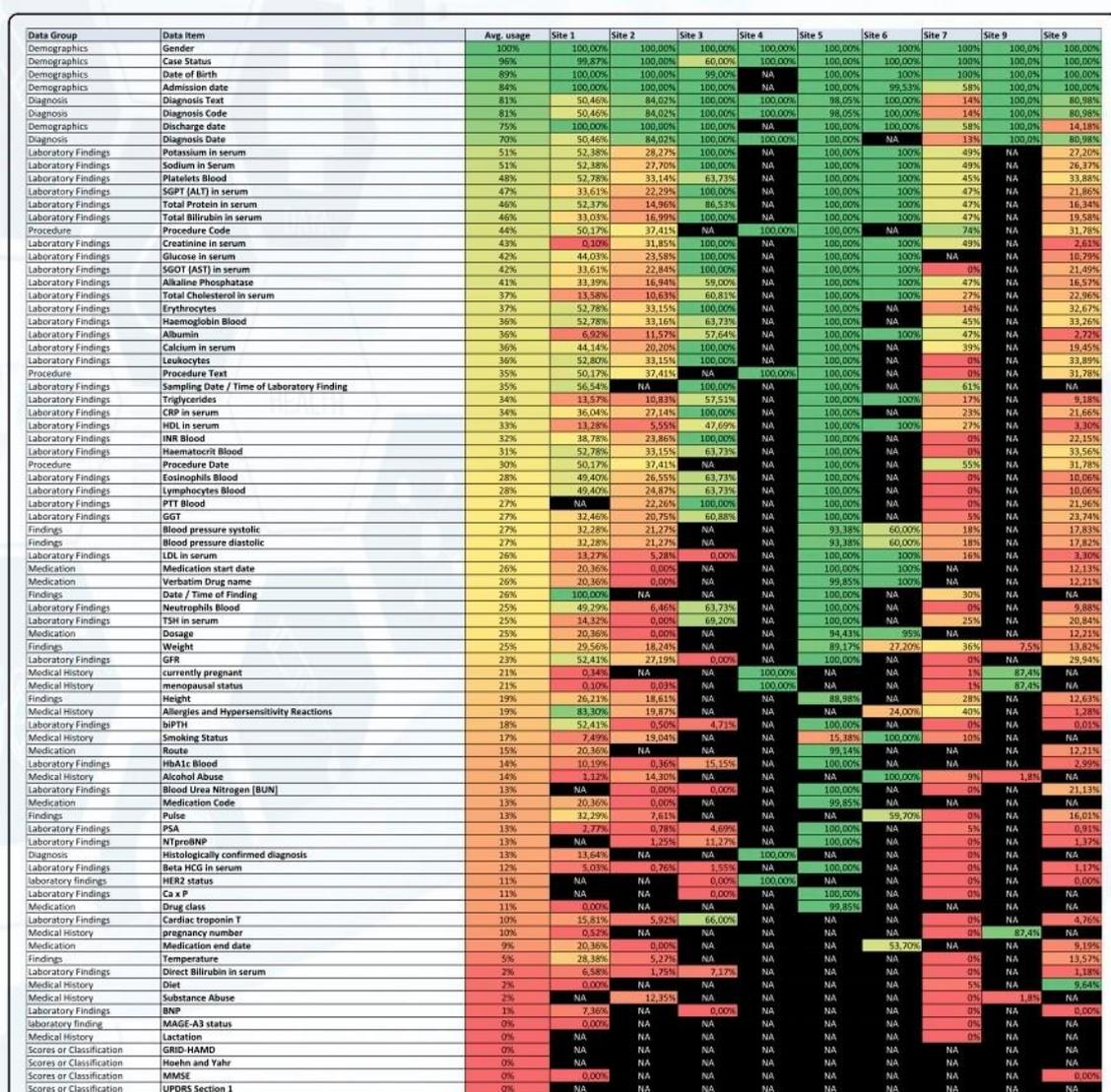


Figure 3 Heat map of the data exports from the data inventory current version. The first two columns describe the ISO 11179 data element concept (data group/data item). The third column shows the average usage of the data element over all sites while the following columns (site 1 to site 9) display the frequency at the individual sites. The Data Inventory is ordered by the average usage sorted in descending order from most available to least. The frequency ranges from 100% (dark green) to 0% (dark red). Data elements that are not available at a site are shown as Not Available (NA) (black).

Figure 9: Heat map showing the data items commonly occurring in clinical trial protocols, and their availability in the pilot site hospital EHR systems

Several surveys, checklists and overviews were developed to prepare deployment of the EHR4CR platform (e.g. 'site readiness', status of ETL processes). Appropriate clinical trials were identified and processed in order to test the platform. Data providers dealt with local ethical requirements for access to real data, set up clinical data warehouses and created mappings from local to central terminologies. Efficiency and effectiveness of the PFS components was tested, a first evaluation of the PRS components (comparison of screening list vs platform) was done with a second one (platform vs manual check) being finished shortly. Due to the adjustment of the project objectives for

scenario three, CTE work was focussed on definition and validation of CTE data elements. Tasks related to general system architecture and platform were performed as well such as an evaluation of the viability of the concept. The impact of the EHR4CR platform on the workflow and on user satisfaction was assessed.

With regard to the protocol feasibility scenario, the proof-of-concept demonstrator has been tested using feasibility queries from twelve different clinical trials. All EFPIA partners participated in this user acceptance test. Overall, 373 free-text eligibility criteria were reviewed by clinical trial experts. 175 feasibility criteria were transformed into a computable representation. Pilot sites mapped approximately 300 codes from their local terminologies into the central EHR4CR terminology, for instance taking into account different national coding systems for medical procedures.

Evaluation of PFS compared the number of patients counts reported using traditional feasibility methods vs. patient counts obtained through EHR4CR platform, vs. manual count of eligible patients obtained through manual review of patient files.

An evaluation of the patient recruitment system (PRS) was also undertaken. The objective was to compare results from the platform with results from manual chart review of patient records. A publication is planned shortly thereafter. A publication about the data inventory from the clinical trial execution and adverse events scenario is in an advanced state and will be submitted shortly. EHR data exports demonstrated that many data items for SAE reporting are currently not available in EHR systems of pilot sites. Therefore the focus was set to identify CTE data elements. Case report forms of 24 clinical trials of different disease areas were analyzed. Through an iterative and consensus-based process, data elements were compiled for all disease areas and with special focus on the reporting of adverse events. Afterwards, the hospitals performed a data element identification and data export step to provide values for availability and completeness of data. The results were compared with the data inventory for patient identification and recruitment. The analysis resulted in 133 unique data elements. Fifty elements are congruent with the previous inventory and 83 elements were identified for clinical trial execution. Demographic and laboratory elements lead the list of the coverage in hospitals EHR systems. For the reporting of serious adverse events only very few elements could be identified in the patient records.

Challenges. Due to several delays in the availability of the Protocol Feasibility (PFS) and Patient Recruitment (PRS) Platforms, not all tests and evaluations were performed as initially planned. The PRS scenario was therefore tested retrospectively. Additionally, the project experienced a significant and unexpected shortfall in budget in the fifth year, which resulted in work package 7 having to close down early the fifth year, and therefore not having the opportunity to undertake any further evaluation work during year five on the PRS or on the prototype implementations of CTE.

The Champion Programme

The intent of the Champion Program is to drive early adoption and to start building a sustainable network of new hospitals together Custodix, as the first EHR4CR service provider, with a group of industry partners. The Champion Programme is designed to provide a low-risk entry for all stakeholders into a new business model approach to efficient use of Real World Data. Thus, the

program is a key step in building the EHR4CR envisaged ecosystem of network of hospitals, service providers and pharma users.

In 2015, industry partners, the European Institute of Innovation through Health Data (i-HD) and Custodix developed a collaboration model that outline principles, contract and budget for the Champion program. The program is an independent cross industry collaboration taking the first deployment step based from the results of the EHR4CR project into a sustainable network of hospitals connected to a new commercial ready platform for EHR data driven services to support clinical trials. The eight involved industry partners, a.k.a. industry Champions, are Amgen, AZ, GSK, Janssen, Roche, Sanofi (as previous EHR4CR Efpia partners), ICON plc, and Boehringer-Ingelheim.

Through a funding mechanism, each industry partner sponsors the connection/setup, by Custodix, of three hospital of preference to the InSite platform. The sponsoring Industry Champion select the Champion Hospitals, in full transparency with the other Industry Champions. The program will also involve governance through the newly established i~HD Institute (also a deliverable from the EHR4CR project). The budget model for industry partners includes each industry Champion to provide in-kind support at various stages of the Champion Program (i.e. provide the necessary resources to propose validation plans and support their execution) and a fee to become members of the i-HD institute.

The scope of the Champion Program is to have a 15 – 30 Champion Hospitals from different EU countries, giving access to at least 2 Million patients in almost real time. The programme aims at including at least one US hospital to demonstrate the global ambition of the program. The current master list of candidate Champion Hospitals includes organisations from Sweden, UK, Poland, Netherlands, Germany, Spain, France, Belgium, Finland, Switzerland and Italy.

The first Champion Program, ending 2017, will offer great opportunities of industry allowing partners to execute Protocol Feasibility Service (S1), and Patient Identification & Recruitment (S2) using all hospitals connected to the platform for on-going trials.

The Champion Program is targeted to provide different business value to participating industry and hospitals. Industry partners will have access to a new innovative tool for better trial design by optimising clinical protocols through direct response from updated EHR data. The protocol feasibility testing service will allow fast iterations of inclusion/exclusion criteria, which will reduce costly corrective measures such as protocol amendments, late addition of new trial countries or sites. The established, and growing, hospital network in place will improve trial success rate and the number of trials failed due to failure to recruit will be reduced.

Through the sponsoring process and transparency within the industry Champion group, industry Champions will be able to improve their relationship with a growing network of hospitals. The pre-competitive collaboration model among stakeholders brings benefits in allowing industry to jointly validate and improve the InSite platform while working with the i-HD in refining the rules of engagement for a sustainable ecosystem. Furthermore, this jointly undertaken initiative offers a cost effective way to reuse EHR data for research as it removes the need for each individual company to establish their own hospital network.

Hospitals participating in the Champion Program, agree to have their InSite clinical data warehouse available for trial design services, are expected to gain a range benefits. Champion hospitals will be

able to attract more clinical research studies and by increased efficiency of the InSite, tools for use within the hospital will speed up identification of trial candidates.

During 2017, with the first results emerging from the Champion Program, the intention is to secure a long term relationship with all actors and to further expand this novel ecosystem for supporting clinical research using EHRs in Europe and beyond. It is clear that the potential and quality of these services increases with the number of participating hospitals. Therefore, the ambition is to grow the network by attracting more hospital sites to join the platform, involve more service providers, and more end-users from both industry and academic centres.

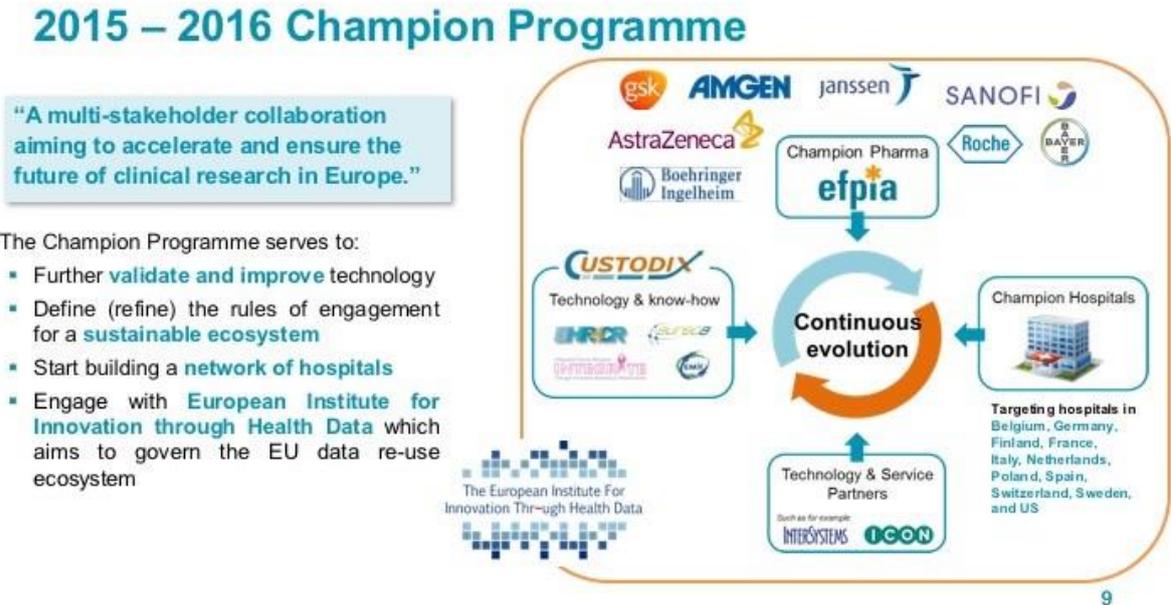


Figure 10: Illustration of the Champion Programme

The European Institute for Innovation through Health Data

The European Institute for Innovation through Health Data (i~HD) has been formed as one of the key sustainable entities arising from the Electronic Health Records for Clinical Research (EHR4CR) and SemanticHealthNet projects, in collaboration with several other European projects and initiatives supported by the European Commission. The vision of i~HD is to become the European organization of reference for guiding and catalysing the best, most efficient and trustworthy uses of health data and interoperability, for optimizing health and knowledge discovery.

i~HD has been established in recognition that there is a need to tackle areas of challenge in the successful scaling up of innovations that critically rely on high-quality and interoperable health data, to sustain and propagate the results of health ICT research, and to specifically address obstacles to using health data that are not being addressed by other current initiatives.

It has been formed after wide consultation and engagement of many stakeholders to fill a recognised gap, to develop products and services that can help to maximise the value obtained by all stakeholders from health data, to support innovations in health maintenance, health care delivery and in knowledge discovery. It will importantly bring multiple stakeholder groups together in order

to ensure that future solutions serve their collective needs and can be readily adopted affordably and at scale (see Figure 11).

i~HD has been established as a European not for profit body, registered in Belgium through Royal Assent. It will be governed by its member stakeholders, public and private, through an elected Board and officers. It will be financed by a mixture of membership subscriptions, fees from providing services such as certification and accreditation, specific project grants and other income from education, training and expert advisory roles.

The following objectives are reflected within the Articles of Association that define i~HD.

Defining and supporting the adoption of best practices in information governance, including complying with legislation and ethics, privacy protection, and codes of conduct, relating to the trustworthy use of health data including capture, processing and sharing.

Championing harmonised health information and standards for capturing, curating, protecting and exchanging health data in a trustworthy, legally compliant and transparent manner using best practices. This is to enable complete and interoperable health records on individuals and populations to deliver benefits to all stakeholders, supporting and guiding the best use of standards and assets for semantic interoperability and privacy protection. These benefits will relate to the care given to individual patients, to the configuration of healthcare and wellness services for populations, and to the reuse of health data for knowledge discovery.

Providing and/or fostering capabilities to enable better quality health data, and the legitimate sharing and uses of health data, including semantic interoperability info-structures and assets, exchange and research platforms and tools, informatics standards and resources to support standards adoption, de-identified health data repositories and research data source catalogues and metadata.

Facilitating, deriving and using intelligence from health data (scientific and clinical intelligence, research, knowledge discovery, service improvement and business intelligence) through advancing the uses of a wide range of potential data sources. These sources primarily are: electronic health and care records and personal health records, citizen sourced and mobile health data, registries and claims databases, cohort studies and biobanks, and clinical trial and electronic case report forms.

Performing and commissioning quality assessments, and conducting quality audits of health data, ICT systems and applications, personnel competence and training and organisational processes relating to the use of health data. Such audits may, for example, relate to the governance of the capture, usage and communication (sharing) of health data, or to the quality of health data at a given site such as a hospital.

Building synergy and consensus: acting as a focal point bringing stakeholders together to share experiences, agree common priorities and approaches for maximizing the benefits of good quality and interoperable health data and the trustworthy reuse of health data. *i~HD* is working towards convergence and cross-fertilization between: healthcare providers, patients and families, health ministries and insurers, EHR system vendors and standards development organisations, pharma and the clinical research community, national and multi-national decision makers.

i~HD will play specific roles in support of the Champion Programme summarised above.

- Providing an essential governance framework for the scale up of EHR4CR and future research platforms across Europe
 - certifying research platforms and service providers
 - establishing codes of practice and privacy protection policies
 - conducting audits and investigating any concerns about security and privacy
 - educating the public of the value of using health data for research and assuring them about the governance protecting their privacy
- Supporting better quality and interoperability of health data
 - establishing a Network of Excellence amongst data providers to improve data quality
 - identifying ICT mature data sources e.g. hospitals
 - facilitating alignment amongst standards bodies, especially in semantic interoperability, ensuring that future standards prioritise clinical and research needs.

Responding to a convergence of needs

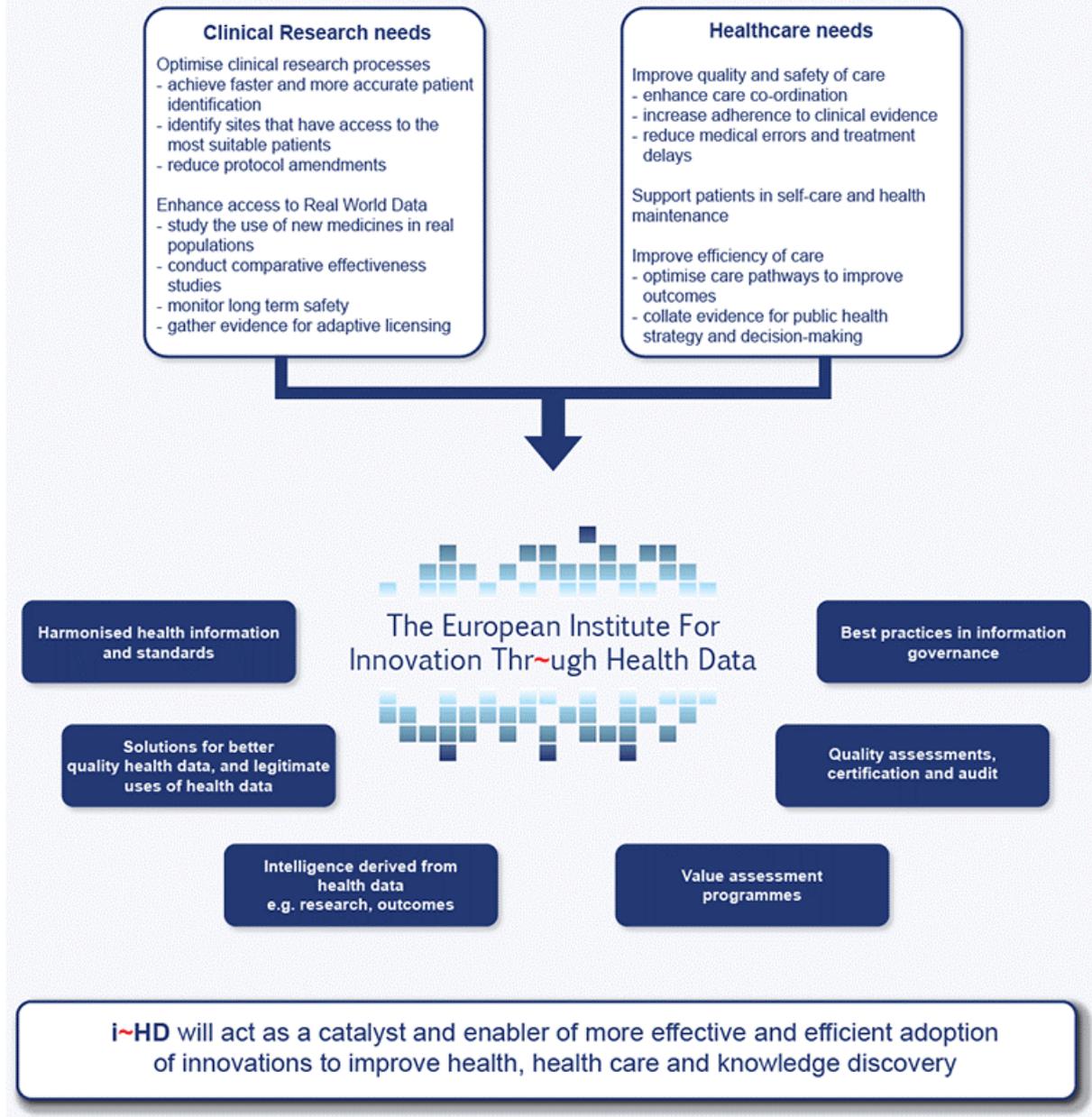


Figure 11: Clinical research and healthcare needs that have triggered the formation of i~HD, and the main areas it will focus on



The European Institute for Innovation through Health Data (i~HD), held its inaugural conference and public launch on 10 March 2016, in Paris. The event brought together over 200 experts from across Europe, including health ministries, insurers, the pharma industry, healthcare providers, patient associations, health professional associations, the health ICT industry and standards bodies.

Dipak Kalra, President of i~HD and Georges de Moor, Chair of its Advisory Board, welcomed participants and introduced the Institute to them. They explained that the priorities of i~HD, a not-for-profit Institute, are developing best practices in privacy protection governing research uses of health data, promoting better adoption of interoperability standards and establishing initiatives to improve the quality and beneficial uses of health data.

Participants learned from keynote speaker Gaël Raimbault why enabling better use of health data is a key target of the French Ministry of Health, which is keen to see greater value derived from national investments in ICT, and regards the reuse of clinical data for research as of strategic importance. Pierre Meulien, the Executive Director of the Innovative Medicines Initiative (IMI1), which is investing over 5 billion Euro in public private research projects, emphasised the ambition of improving the affordability and speed of access to innovations for patients. IMI projects are using electronic health records to speed up clinical trials and using big data to discover how to better target innovative therapies to the particular patients who will respond best to them. Its new Big Data for Better Outcomes programme will also work closely with health care stakeholders to help apply new evidence emerging from big data to improve healthcare systems. Participants also learned about Europe's largest "big data" project in health: the European Medical Information Framework (EMIF, funded by IMI), from its co-ordinator Bart Vannieuwenhuysse.

Terje Peetso, policy officer from the European Commission, DG CONNECT, emphasised the importance that the EC places on improving the interoperability of health data, and empowering citizens to play a greater role in their own health care and wellness. The EC is investing, through its Horizon 2020 programme, in many initiatives to improve information connectivity across health care systems, to provide citizens with great assurances about the privacy protection and trustworthiness of personal health applications and devices, and to tackle the particular healthcare challenges of an ageing society. John O'Brien, a former hospital CEO, explained why hospitals need to better value the health data that they collect, and therefore to ensure the ICT systems that they procure are of a quality that can support the organisation to optimise its performance in delivering patient centred care, maximising outcomes as well as business efficiently.

This inaugural conference also marked the parallel launch of a novel European platform to support multi-centre clinical research. This operational platform, the result of the European project EHR4CR (Electronic Health Records for Clinical Research, a recently completed IMI project), connects securely to the data within multiple hospital EHR systems and clinical data warehouses across Europe, to enable a trial sponsor to predict the number of eligible patients for a candidate clinical trial protocol, to assess its feasibility and to locate the most relevant hospital sites. The role of i~HD, as a not-for-profit Institute, is to provide independent governance oversight of such clinical research platforms and services, as these expand to connect with multiple hospitals across Europe. Several presentations were given on this theme: Mats Sundgren introduced the EHR4CR platform and its pharma-led adoption programme, complemented by Brecht Claerhout who explained the InSite platform and tools that will be used for this. Nikolaus Forgo, Professor of IT Law, explained the new European General Data Protection Regulation and its potential implications for clinical research and the integration of big data. Dipak Kalra and Peter Singleton explained the i~HD governance principles and services that will ensure state-of-the-art in the trustworthy reuse of health data for research. Pascal Coorevits presented the importance of the quality labelling and certification of health ICT products, and how EuroRec will assess the conformity of the new clinical research platform.

Also in the conference spotlight were the challenges and state-of-the-art approaches to improving the quality and semantic interoperability of clinical data, which was discussed within a panel chaired by Veli Stroetmann, comprising health ministry (Michèle Thonnet, Jeremy Thorp), health insurance (Christoph Rupprecht), clinician (Robert Vander Stichele) and patient (Petra Wilson) perspectives. The panel emphasised that interoperability is vital to ensure the coordination of care, especially because of increasing comorbidity, with older generation patients having multiple long-term conditions and multiple treatments that can interact, potentially dangerously, unless care providers have the complete picture on their patients. There is a discord between the actors who use health ICT systems to record information, the actors who want to make use of that information, and those who invest in the ICT systems and thereby determine what is purchased, something that a new Horizon 2020 project VALUeHEALTH is investigating. There was a consensus among the panellists that today's ICT systems demonstrate poor connectivity and poor patient orientation, and that many applications in use are not particularly friendly to the language and workflows of clinicians and patients. The panellists emphasised the importance of making better use of interoperability standards, and declared that the key actors to drive that adoption are the public authorities and health insurance. They also emphasised that there should be better cooperative design of ICT solutions with end users. The audience were informed that i~HD will play an future role in the development and quality labelling of interoperability specifications, bringing together clinical and research domain experts, with patients, to help ensure that future standards will support patient care, learning health systems and clinical research. It will also be running initiatives to support hospitals with improving the quality of their health data.

Two of the conference speakers specifically represented the views of patients and of society. Petra Wilson explained that patients are increasingly involved in the collection of their own health data, for example through monitoring devices, but do not always have access to their own data. They must be much more involved in how their health data are used to inform decision-making, and in those care and treatment decisions themselves. Mary Baker stressed that society needs to be much more committed to promoting wellness and accelerating the discovery and testing of innovative

treatments. Our ageing society is accumulating long-term conditions, and we need to be much more proactive in prevention and early detection. Health data are vital to improving our understanding of disease and the impact on the lives and well-being of patients. Society needs to better trust the security measures that can nowadays be applied to protect privacy, and to recognise the balance in proportionality between safeguarding health data and putting health data to good use.

1.6. Potential impact and main dissemination activities and exploitation of results

The EHR4CR scientific/technical outputs contribute to the following overall IMI objectives:

- to provide socio-economic benefits for European citizens,
- to contribute to the health of European citizens,
- to increase the competitiveness of Europe and help to establish Europe as the most attractive place for biopharmaceutical research and development.

The intended impact of EHR4CR, through the pan-European (and potentially global) scaling up of connectivity between research sponsors and de-identified EHRs is to speed up the process of conducting clinical trials, by:

- Optimising a protocol to maximise the likelihood of recruiting sufficient patients at the start, and so reducing the number of subsequent protocol amendments, thereby removing some causes of delay and cost in the completion of a trial;
- Helping to pinpoint the hospital sites that are most likely to have sufficient patients to be a viable trial centre, thereby avoiding the costs of establishing centres that are unlikely to recruit many or any patients;
- Helping hospital to track down eligible patients so that they can be proactively contacted and potentially recruited, as opposed to common practice today of awaiting their spontaneous presentation to a hospital and remembering their potential eligibility.

The end result of these improvements will be to accelerate the conduct of clinical trials and to reduce their cost, and so to speed up the delivery of new innovative medicines to health care and potentially to reduce the end to end cost of delivering such drugs to the market. This may have an impact on the eventual cost of new innovative medicines as well as their more rapid availability. There are socio-economic benefits to citizens and expected to improve the health of citizens.

If the solutions are primarily deployed in Europe this will improve the economic attractiveness of Europe as a location for clinical trials.

The Champion Programme, including its evaluation, will help to promote the value of the inSite platform and (in due course) equivalents. i~HD will help to promote the value of re-using EHRs for research and the codes of practice that should be followed. The project as a whole has conducted an impressive level of dissemination through multiple channels: academic publications, conference presentations, workshops, interviews with key opinion leaders and decision makers, webinars, web site, videos, social media etc. These dissemination activities are detailed later in this report. The sustainability entities will continue these activities beyond the end of the IMI project.

- To provide socio-economic benefits for European citizens,
- To contribute to the health of European citizens,
- To increase the competitiveness of Europe and help to establish Europe as the most attractive place for biopharmaceutical research and development.

Please outline how the project outputs have/will have the potential to be rapidly and broadly spread and taken up within the scientific/industrial community and healthcare professionals.

1.7. Lessons learned and further opportunities for research

It is not easy to capture in a few words the richness and value of the collaboration that has taken place over the five years between the EFPIA and the public partners. The list below is therefore offered in recognition that it is both incomplete and superficial. It is inevitably also the opinion of the project leadership, who have taken the lead responsibility for writing this report. Our recommendation would be that this matter be explored through focus groups where real depth of understanding, appreciation of value and honest recognition of issues can be discussed in depth.

Positive value

Expertise: in areas such as requirements, interoperability standards, regulatory requirements, business model inputs.

Customer viewpoint: in areas such as requirements, proportionality of problems and complexity of solutions, priority and urgency, business value.

Large-scale industry perspective: noting especially the ways of working, decision making and what decisions matter most, in contrast to the perspectives of hospitals, academia, SMEs etc.

Legacy/risk averse perspective: recognising that it is difficult to introduce innovation in a large, expensive and business critical pipeline.

Project management: finding the balance between industry orientation towards timing, resources and delivery oriented management strategies versus scientific discovery trajectories and software development workflows.

Leadership skills: different styles and expectations of what leadership needs to be about: for example how important is mind share and win-win, how closely to track people and deliveries, how best to balance control and delegated responsibility.

Contact networks: each having potent links, and blind spots.

Issues

Scientific: co-leadership of WPs is a good idea in theory, but industry representatives sometimes lacked the scientific knowledge needed to back up the academic lead in issues of scientific decision making, and sometimes vice versa. Sometimes the customer needs to be assertive!

Financial: imbalance of commitment to meeting in-kind contribution, difficulty in converting in-kind to in-cash, resulting in a (sudden, unexpected) budget shortfall with a negative impact on the work

plan. Unplanned conversion to in-cash is not always the best approach, since the loss of skills cannot always be rectified through a cash injection if the necessary skills cannot quickly be hired, without relevant organisational and project knowledge, and placing an unplanned-for management and governance burden for that cash sum on the Managing Entity or other public partners. We note that the EFPIA contribution has to be matched across the IMI programme, not at the project level, which is not fair to the projects such as EHR4CR. IMI Office seemed unable to help with this: it does rather undermine the relevance and enforceability of a Description of Work and Contract.

Solutions

Since recognition of the issues is not universally understood across the EFPIA partners, the first step in developing better strategies to ensure the effectiveness and risk management of IMI projects is to increase awareness that these are indeed real issues. We therefore recommend workshops for in depth airing of experiences and views, and that better approaches to the IMI projects are developed on the basis of this richer mutual understanding.

Further areas of research

EHR4CR has undertaken some design work on two scenarios (CTE, SAE) that we would still like to implement and evaluate.

We recognise that there are many more areas of R&D that our project has uncovered, and need further research investment. These include:

- Additional clinical research scenarios for which platform tools and services need to be designed, developed and tested, including observational research;
- Semantic interoperability, and the challenge of harmonising eHealth and clinical research standards, harmonising eHealth and life sciences (emerging) standards;
- Data quality;
- Extracting concepts from natural language to augment structured data, for clinical trial feasibility and recruitment, to pre-populate CRFs, and for observational research;
- Interfacing with patient held records, also to augment the data underpinning clinical trial feasibility and recruitment, and for outcomes research;
- Best practices in information governance and privacy protection;
- Interoperability of provenance and other forms of metadata;
- Establishing the trustworthiness of real world data findings.