European Translational Information and Knowledge Management Services

eTRIKS Deliverable report

Grant agreement no. 115446

Deliverable 5.2: ROI Evaluation Report at Two Years

Due date of deliverable: Month 27

Actual submission data: Month 34

Dissemination Level: PU

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

DELIVERABLE INFORMATION

Project	
Project acronym:	eTRIKS
Project full title:	European Translational Information and Knowledge Management Services
Grant agreement no.:	**115446**
Document	
Deliverable number:	5.2
Deliverable title:	ROI Evaluation Report at Two Years
Deliverable version:	1.0
Due date of deliverable:	Month 27
Actual submission date:	Month 34
Leader:	Jay Bergeron, Scott Wagers
Editors:	David Henderson
Authors:	Jay Bergeron
Reviewers:	D. Henderson, Y. Guo
Participating beneficiaries:	U Luxembourg, Imperial College,
Work Package no.:	5
Work Package title:	Program Management
Work Package leader:	Jay Bergeron, Scott Wagers
Work Package participants:	C. Auffray, J. Bergeron, T. Garret, Y. Guo, D. Henderson, N. Julian, C. Marshall, I. Pandis, A. Rowe, M. Saqi, R. Schneider, K. Teburi, S. Wagers
Estimated person-months for deliverable:	1
Nature:	Document associated with IMI Mid Term Review
Version:	1.0
Draft/Final:	1.0
No of pages (including cover):	
Keywords:	eTRIKS, Mid Term Review

Contents

Executive Summary

The eTRIKS mid-term review was originally planned to occur during 4Q2014; personnel availability, however, resulted in a postponement to 2Q2015. The review was conducted at the IMI offices in Brussels on 21-May-2015, with the formal written outcome presented to the Coordinator and eTRIKS Executive Committee (ExCom) on 6-June-2015.

The review committee was comprised of Drs. **Paul Avillach** (Harvard), **Ann Martin** (IMI Scientific Manager, IMI), **Markus Perola** (IMI Scientific Committee, U. of Helsinki) and **Manolis Tsiknakis** (eTRIKS evaluation panel member, FORTH Institute of Computer Science).

The eTRIKS program was presented by **Charles Auffray** (EISBM, WP5,7), **Jay Bergeron** (Pfizer, Scientific Coordinator, WP2,5,6), **Michael Braxenthaler** (Roche, WP3), **Trevor Garrett** (Biosci Consulting, WP5), **Yike Guo** (Imperial College, Academic Lead, WP2,5), **David Henderson** (Bayer, WP5,7), **Paul Houston** (CDISC, WP3), **Ioannis Pandis** (Imperial College, WP2,4,5), **Ghita Rahal** (CC, WP1), **Reinhard Schneider** (U. of Luxembourg, WP4,5), **Kevin Teburi** (Astra Zeneca, Coordinator, WP5), **Scott Wagers** (Biosci Consulting, WP5,6).

In general, the reviewers acknowledged that the eTRIKS project has been proceeding according to its *Description of Work* (DOW) and that both personnel and financial resources have been allocated in a manner that is consistent with the program's objectives. The development of an *open translational research data management platform (eTRIKS v1)* based on the tranSMART data warehouse system in which all of the platform's components are open source/license was lauded as a substantial achievement in support of the research goals of eTRIKS' clients as well as the wider translational research community. The reviewers recognized that operational synergies are arising from the comprehensive set of information services that eTRIKS provides to client projects both within, and beyond, the IMI. Furthermore, the eTRIKS *Public Server* was cited as a welcome extension to eTRIKS' planned deliverables and the *Standards Starter Pack* and *Guidelines for Secondary Reuse of Clinical Data* were commended as broadly applicable contributions to the EU scientific community.

The reviewers cautioned that substantial challenges remain for eTRIKS, reinforcing that eTRIKS' objective to support forty client projects by the completion of the collaboration remains critical to the consortium's success. With thirteen projects engaged at month thirty-two the reviewers expressed serious concern that the forty-project goal is not achievable. The reviewers urged eTRIKS' leadership to align all of the collaboration's efforts towards client engagement/support and to intervene and suspend non-strategic activities that do not directly impact current or potential client projects.

The difficulties associated with concluding cross-collaboration Confidential Disclosure (CDA) and Material Transfer Agreements (MTA) were recognized to be serious impediments preventing eTRIKS from interacting with client projects in the manner that was originally envisioned. Furthermore, the complications that prolong CDA and MTA negotiations are recognized as universal impediments to establishing the types of collaborative ventures that the IMI promotes. The reviewers advocated continued vigilance in establishing broadly applicable CDA and MTA templates that can be generally adopted by EU organizations wishing to engage in scientific collaborations involving data derived from clinical subjects.

No recommendations were made to specifically alter the organizational structure (work/delivery packages) or the primary objectives of the consortium. However, there were several recommendations made by the reviewers to address the operational deficiencies outlined above.

The reviewers insisted that general under-allocation of resources across both academic and EFPIA partners be corrected in the immediate term by eTRIKS' leadership. The reviewers also expressed concern regarding sub-par quality of many of the private deliverables and that more critical evaluations need to be performed with regard to the submitted deliverables.

The reviewers advised that a more rigorous project management approach be instituted within the project and that this approach include specific risk mitigation planning.

The reviewers recommended the addition of two deliverables to the DOW.

1. **End of 2015:** A deliverable to document the contracting difficulties experienced by the program.
2. **End of Project:** A deliverable to document and disseminate expertise regarding the services created for eTRIKS.

Modification of the DOW to include the proposed deliverables will be brought to vote at the next eTRIKS General Assembly meeting.

The reviewers were uncertain concerning the prospects of the eTRIKS network and requested that the sustainability models be more fully detailed.

With regard to forward looking activities, the reviewers enthusiastically endorsed developing the eTRIKS Harmonization Service (EHS) and the accompanying Meta Data Repository (MDR).

Inputs and Outputs from Related Deliverables

Inputs
As the Mid Term Review is a summarization of eTRIKS' results to date associated with objective evaluation by a panel of external experts, all project deliverables are, essentially, inputs. However, the periodic report deliverables 1 and 2, submitted in Sept 2013 and Oct 2014 respectively, are direct inputs to this ROI evaluation.

This deliverable was put forward under the guidance of WP5 with input from all work package group leaders.

Outputs
The following specific outputs are arising from this deliverable

1. Proposal to GA to update the DOW, new deliverable for *"Documenting Issues and Processes Associated with Confidentiality and Material Transfer Agreements"*
2. Proposal to GA to update the DOW, new deliverable for *"Documenting eTRIKS Services and Best Practices for Dissemination at Project Close"*
3. Proposal to WP5 to further detail the sustainability plan including the *"eTRIKS Network"* concept

4. Proposal to WP5 to introduce more rigorous project management approaches including risk mitigation
5. Address resource under-allocation (responsibility of the Coordinator)

Description of Work Achieved

Summary of Accomplishments

eTRIKS offers a comprehensive open data management platform and accompanying best practices to public private partnerships, consortia and collaborative research groups that engage in translational research. The system is designed to facilitate management and analysis of exploratory clinical and multi-OMICs data. eTRIKS has engaged thirteen client programs to date, providing software development, hosting, data curation and data import services. Additionally, eTRIKS has trained over 100 investigators in the use of the eTRIKS platform and detailed CDISC-based data standards to facilitate the reuse and integration of translational datasets. eTRIKS has assembled and promotes guidelines to foster the ethical use of human data for exploratory research and maintains an open access portal containing forty-three curated public domain translational studies. Although eTRIKS has been challenged in closing Confidential Disclosure (CDA) and Material Transfer (MTA) Agreements, the collaboration has responded by establishing a mandate allowing the coordinator (Astra Zeneca [AZ]) to authorize confidentiality agreements with client projects and developing an MTA template to facilitate multi-party negotiations.

The eTRIKS platform is based on the popular open source *tranSMART* translational research data warehouse. Modified by eTRIKS to incorporate the open source PostgreSQL relational data base, this fully open source tranSMART platform, initially released as eTRIKS v1.0/tranSMART 1.1 and currently revised and published as eTRIKS v2.0/ tranSMART 1.2, has become a prominent translational research system for both academic institutions and commercial entities alike.

To augment the eTRIKS platform, mature operational practices have been developed for application release, hosting, data curation/loading and client training/support. Business development/outreach practices have been established to promote eTRIKS to prospective clients. The result is a comprehensive suite of products and services that enable investigators to curate and process complicated and variable clinical and molecular data elements in a manner that is consistent, comprehensive, economical and open to inspection.

Motivated to sustain these best practices beyond the completion of the collaboration, eTRIKS conceived and launched the *eTRIKS Network* which assembles cooperatives of eTRIKS-enabled service providers to respond to the translational information needs of large collaborative research projects on a fee for service basis. From training to hosting to program management, the eTRIKS Network provides both comprehensive and targeted information services tailored to fit the needs of translational research programs. By month thirty-four, the eTRIKS network had signed data curation and hosting contracts with two end term projects utilising the company ITTM (*Information Technology for Translational Medicine S.A.*), a commercial spinoff arising from eTRIKS and the U. of Luxembourg, as a provider of eTRIKS services.

Program Structure

eTRIKS is overseen by Astra Zeneca (Coordinator) and Imperial College London (Managing Entity) with the assistance of an Executive Committee (ExCom) having representation from Imperial College, U. of Luxembourg, EISBM/CNRS, Astra Zeneca, Johnson and Johnson and Bayer. The ExCom is accountable for ensuring that the project executes per the vision of the partners as put forward in the

eTRIKS *Full Project Proposal* and *Description of Work (DOW)*. The ExCom is supported by a Program Management Office (Biosci Consulting) and a Scientific Coordinator (Pfizer).

All eTRIKS participants share equal influence with regard to changes to the program via General Assembly (GA) meetings that are held at the annual meeting and otherwise as required. For example, amendments to the *Description of Work* require a General Assembly majority vote.

eTRIKS is comprised of seven work packages (WP) aligned with the key efforts of the consortium. Work packages actively collaborate to promote the development of comprehensive solutions via recurring cross-WP meetings. The *"Delivery Package"* formalizes cross-WP working relationships for sizable deployments to clients, including multi-faceted support for select IMI projects such as U-BioPRED, ABI-RISK and OncoTrack. Each WP is led by representatives from Pharma and the Beneficiaries. As a group, the WP leaders meet by phone monthly and face to face thrice annually to ensure coordination of effort and satisfactory progress towards goals.

eTRIKS also participates in a wider translational research informatics community that includes the tranSMART Foundation (a U.S. non-profit responsible for coordinating tranSMART development and the larger tranSMART community) and the CTMM TRaiT consortium (a Netherlands-focused public private partnership, similar to eTRIKS, that promotes translational research information best practices and also uses tranSMART as a base software system). ETRIKS' academic leader and scientific coordinator serve in leadership roles with respect to the tranSMART Foundation. Moreover, eTRIKS and CTMM TRaiT personnel interact closely, meeting twice monthly to promote cooperative interaction.

eTRIKS Project Challenges

Timely prosecution of CDA and MTA agreements
Closing CDA and MTA agreements has been highly problematic given the need for individual review and authorization by each Participant within both eTRIKS and the client project. The iterative rounds of refinement have led to turn-around-times of many months for CDAs and months/years for MTAs. In order to expedite CDAs, eTRIKS partners signed a mandate allowing the coordinator (AZ) to authorize CDAs on behalf of all eTRIKS partners provided that the CDA conforms to a pre-authorized template. MTAs have proven especially difficult given the variability in legal philosophy across partners and national borders. eTRIKS has developed an MTA template for use by new projects. However, to date, agreements between specific data handlers (organizations *and* individuals) have been put in place to promote data transfer. The difficulties associated with MTAs have prompted eTRIKS to consider using alternative IMI *Data Processing* agreements or to simply abandon cross-consortium MTA agreements altogether.

Multi-Instance Implementation
Related to the prosecution of CDAs and MTAs, project-specific data use constraints have forced eTRIKS to create a distinct software instance corresponding to each supported project. A multitenant environment in which security and access would be maintained by eTRIKS was considered an ideal outcome at the start of the project. The multitenant instance was to serve as a publically accessible source for IMI data following project completion. The public server certainly can become such an open platform for IMI projects. However, the IMI requirement for post-project data dissemination can be interpreted in a variety of ways leading to uncertainty in how best to create and manage a single comprehensive post-project data repository. Nevertheless, eTRIKS is evaluating alternatives for post

project data sustainability (e.g. with groups such as Elixir) and the EFPIA partners are considering an IMI-2 data sustainability project.

Scaling Best Practices

The number of projects that eTRIKS has been able to support has been limited predominantly by the number of eTRIKS staff members who can apply informatics best practices on behalf of clients. These practices include client engagement, data curation/import, systems training, coordination of legal agreements, custom software/analytics development and consulting. These services are provided comprehensively to certain clients, such as UBioPRED. However, it became clear that providing such comprehensive support to all potential clients would severely limit the number of projects for whom eTRIKS could provide services. Recognizing that many prospective eTRIKS clients have very specific information management requirements and that informatics expertise was often available among the client's staff, eTRIKS developed a targeted engagement model designed to provide select data management services tailored to address the priority needs of individual clients. Moreover, eTRIKS seeks to enable client staff in the performance of these best practices allowing the clients themselves to sustain the best practices with their existing staff members. The targeted delivery model makes more efficient use of the limited resources available and has allowed eTRIKS to support 13 client projects by month 32, where a maximum of perhaps eight projects might have been supported with more comprehensive support packages.

Resource Allocation

eTRIKS is operating with under allocation of its resources, which is partially due to the low resource use in the ramp-up phase in the first year. ExCom and project management are working to ensure that the available finances are optimally distributed to support the activities in the second half of the running time.EFPIA contributions are, however, trending low and the ExCom is currently engaging the pertinent partners to develop remediation plans.

Strategy

As eTRIKS approached its project mid-point, the partners sought ways to reinvigorate the strategic roadmap based on the eTRIKS *Description of Work* and the experience of our colleagues and current/prospective customers. With an emphasis on research and operations, the following strategic elements will be eTRIKS priorities during the remainder of the project.

Operational Excellence

Expand outreach to the community of translational research projects. Support many new client projects (goal set at 40) by focusing on enabling these clients to use and support the eTRIKS platform for their own purposes. Lead the adoption of consistent translational research informatics practices in the EU based on the holistic approach developed by eTRIKS.

Knowledge Management

The loosely-defined *Knowledge Management* (KM) infrastructure has been a key element of the proposed eTRIKS platform that, as of month 32, has not been substantially resourced. To address KM goals, the "eTRIKS Harmonization Service" (EHS) subproject was launched in Feb 2015 as a cooperative effort involving work packages 2, 3, 4 and 6 to speed data curation and the application of standards to client data. The EHS will promote data consistency and harmonization across studies and

provide the ability to tailor the representation of study data for specialized uses. The EHS implementation will include a Meta Data Repository (MDR) to facilitate standards association.

Sustainability

Development of the eTRIKS Network will continue with a focus on growing customers and delivering fee based informatics services by suppliers committed to eTRIKS best practices. BioSci Consulting is leading the business development activities necessary to sustain the business practices established by eTRIKS long after the eTRIKS project ends. eTRIKS will continue to work closely with the tranSMART Foundation with respect to development and delivery of new tranSMART software. eTRIKS will cooperate with the IMI translational informatics working group to identify and implement alternatives for long-term data retention for IMI-translational projects.

Return on investment

Return on investment (ROI) is a financial concept defined mathematically as:

$$ROI = (Profit - Cost)/Cost$$

When dealing purely with the sale of a product the calculation of ROI is straightforward. When the impact of an investment is more long term or more broadly distributed it becomes more difficult to calculate ROI. When you do not know the Profit you have to make estimation. It becomes even more difficult when the return is more in terms of a strategic or social benefit. There is precedent for calculating a social ROI [1]. A social ROI uses the same formula with a 'monetised' social benefit in place of Profit.

The main point of calculating an ROI is to provide a quantification that will help in decision-making. Of course the more estimations one makes the less reliable the quantitation will be. However, physicists make estimations routinely, using order of magnitude, or Fermi estimates. The principle is that you estimate the order of magnitude of the number you are uncertain about. It is easier to make an estimate based on orders of magnitude. To increase the accuracy you break down the estimate into a series of estimates that affect that input to the estimate you are trying to obtain and put in real numbers when they are available. By combining a series of estimates you have a better chance of arriving at a more accurate estimate. These types of estimates are being applied to help make business decisions [2].

Many of the benefits from the project are difficult to quantify as the impact of the benefit is wide and data is not available for all the various benefits and can only be estimated. For the most part, however, it is possible to have estimates that are better than an order of magnitude accurate. The main point of this calculation is to determine if the eTRIKS project has had a positive ROI, will continue to have a positive ROI and how to maximize that ROI in the final half of the project.

Table 1 summarizes the estimates used to monetize the benefits from eTRIKS. The left hand column includes the assumptions that determine the inputs. The calculated costs are highlighted in grey.

Table 2 summarizes the return on investment at the ½ way point and the project return at the end of project.

---

[1] Millar & Hall (2012) Social Return on Investment (SROI) and Performance Measurement. In Public Management Review, DOI:10.1080/14719037.2012.698857, p.4

[2] http://lesswrong.com/lw/h5e/fermi_estimates/

Table 1

Variable	Value	Basis for estimation
Deploy a translational research knowledge platform	768.000	This is based upon the estimate that it would require 4 people working for 2 years to develop and deploy a customised system. The person month cost was estimated at 8.000 euros, which is a reasonable estimate for a higher valued bioinformatician and developer. The original budget for knowledge management for U-BIOPRED was of a similar magnitude.
Resource for ETL	96.000	For ETL it is estimated that the typical project would need 1 person working full time for a year at 8.000 euros/month
Number of projects supported at 1/2 way point of eTRIKS	15	This is the number of projects supported in some form or another by eTRIKS at year 2.5.
Number of projects supported with full ETL support	4	These are the projects where eTRIKS was primarily responsible for ETL
Cost of a statistician to perform exploratory analyses	2.000	This is based upon an estimate of 1 week of effort to obtain data and perform statistical analyses to answer simple statistical exploratory analyses
No. of exploratory queries per project	10	This is a FERMI based estimation using the rationale that 100 exploratory queries would be too many, so the next order of magnitude lower, 10, was chosen
Cost of software development	540.000	This is based upon 3 FTEs working full time for 2.5 years at an FTE cost of 6.500/month
Reduction in resource realised when standards are used	50%	Estimate based upon the fact that even if standards are used there will be some ETL effort, but that will be reduced significantly because of readily available mappings and a reduced need to spend effort in gaining an understanding of data types
Reduction in effort to understand data re-use principles and regulation	8.000	Assumed to be 2 months of effort to understand and set up processes for data re-use compliance. A reduction of 50% of that effort would be realised if the code for the re-use of medical ethics is used. Person month cost 8.000 euros
Number of projects benefitting from the code of medical data re-use	10	The code of re-use has been downloaded a number of times. It is not possible to know for certain if it is being used. Again based on FERMI estimation principles the number 10 was chosen
Number of projects expected to benefit from standards in the next 2.5 years	225	This is based upon a 20% up take rate for 1.000 new non IMI projects in the next 2.5 years and uptake by the 25 additional project supported by eTRIKS

Number of projects expected to benefit from code of medical re-use in the next 2.5 years	100	This code is already in use and being downloaded. Based upon Fermi estimation principles 100 was chosen as it will clearly be used in more than 10 projects but 1.000 seems unrealistic
Number of project supported by the end of eTRIKS	40	This is the goal number of projects eTRIKS aims to support by the end of the project
Software development	1.080.000	This is calculated as a doubling of the development estimate at the 1/2 way point

Table 2

1/2 point return summary		5 year return summary	
Curation supported projects	3.456.000	Standards benefit	10.800.000
Non curation supported projects	9.504.000	Medical Re-use guidelines benefit	800.000
Exploratory queries	300.000	Support of 40 projects	31.104.000
Software development	540.000	Software development	1.080.000
Medical Reuse guidelines	80.000		
Total return at 1/2 way point	13.880.000	Return at end of project	43.784.000
Total spend at 1/2 way point	7.831.287	eTRIKS total budget	20.890.728
Differential	6.048.713	Differential	22.893.272
ROI	77%	ROI	110%

Conclusions

The eTRIKS project is delivering a positive return on investment and is project to deliver a 110% return on investment at the end of the project. This highlights one of the most important value propositions of working in a collaborative project – it is more cost efficient. This is not always the outcome. It has been highlighted that there is often a perception that collaborative consortium based projects are less productive than expected.

The positive ROI is driven largely by the fact that eTRIKS is providing an open source platform to projects that would otherwise be costly to deploy again and again with each project developing their own system. This emphasizes the efficiency gain that is realized from working in a standardized manner

with a common set of tools. This type of efficiency gain does require consistent effort from not only those deploying the system, but also from those using the system. Conforming to standards takes effort. As this analysis demonstrates it is a worthwhile effort. However, what is not reflected in this analysis is the value gains from combining and re-using datasets which efforts to transform and load data such as eTRIKS make possible.

Content of Deliverable

The following elements comprise this deliverable:

1. ROI Evaluation Report at Two Years (this document). The ROI evaluation report will serve as the primary artifact associated with this deliverable.

Adobe Acrobat
Document

2. eTRIKS interim review consolidated report.

List of Abbreviations

AZ:	Astra Zeneca
CC:	Computing Center of the Institute of Nuclear and Particle Physics, CNRS
CDA:	Confidential Disclosure Agreement
CDISC:	Clinical Data Interchange Standards Consortium
CNRS:	Centre National de la Recherche Scientifique
DOW:	eTRIKS Description of Work
EHS:	eTRIKS Harmonization Service
EISBM:	European Institute for Systems Biology and Medicine, CNRS
eTRIKS:	Eu Translational Research and Information Knowledge management Services
ExCom:	eTRIKS Executive Committee
GA:	General Assembly
IMI:	Innovative Medicines Initiative
ITTM:	Information Technology for Translational Medicine
KM:	Knowledge Management
MDR:	Metadata Repository
MoU:	Memorandum of Understanding
MTA:	Materials Transfer Agreement
tranSMART:	Open source/license translational data warehouse
WP:	Work Package