Deliverable 7.3

Project ID	654241
Project Title	A comprehensive and standardised e-infrastructure for analysing medical metabolic phenotype data.
Project Acronym	PhenoMeNal
Start Date of the Project	1st September 2015
Duration of the Project	36 Months
Work Package Number	7
Work Package Title	Privacy and Ethics
Deliverable Title	D7.3 Evaluation report for the introduction of a data provider form
Delivery Date	M8
Work Package leader	ICL
Contributing Partners	ICL, EMBL-EBI

Authors: Robert Glen, Timothy Ebbels

Table of Contents

1. Executive Summary 3
2. Project Objectives 3
3. Detailed report on the deliverable 4
   3.1. Background 4
   3.2. Discussion of Data Provider Forms and Guidelines 4
   3.3. Conclusions 7
4. Delivery and Schedule 7
5. Background Information 7
6. Annexes 8
   6.1. Background information on ELSI requirements and recent changes 8

1. Executive Summary

Research data derived from patient samples is sensitive and subject to Ethical, Legal and Social Implications (ELSI). Submitting such data for use and subsequent re-use thus requires consideration of ELSI regulations. In order to enforce patient rights and maximise the opportunities for research, guide users, and, comply with local and international laws and ethical considerations, a data provider form provides the rigour required to consider the implications of acquiring, donating and using clinically derived data as well as being a record of the conditions under which the data are made available. In addition, guidance is required on selection and completion of the appropriate form with the required information. This will ensure that all data collected and held within the project will comply with all local laws, regulations and ethics and all personal information will be processed in accordance with accepted data protection principles. As a general principal, the final responsibility for data will remain with the host institution/data provider. A guidance document on handling sensitive human data has been produced (Annex 1). Technical solutions to secure deposition and access of data are explored, using the exemplar of the European Genome-Phenome Archive (EGA, https://www.ebi.ac.uk/ega/home). This document should be read in conjunction with deliverable D7.2 - report on policies and procedures for sensitive human data management.

2. Project Objectives

No.	Objective	Yes	No
1	Develop appropriate policies, procedures and management accountability and structures to provide a robust governance framework for information management.	X	
2	Raise awareness of information governance within the consortium and assure on going compliance.	X	
3	Provide a forum for information exchange on best practice in clinical data sharing and disclosure.	X	

3. Detailed report on the deliverable

3.1. Background

Two PhenoMeNal workshops were held on ELSI (20th November 2015, 27th January 2016: see Deliverables D7.1 and D7.2) at the Imperial College London (ICL) under the PhenoMeNal objectives:

- Develop appropriate policies, procedures and management accountability and structures to provide a robust governance framework for information management.
- Raise awareness of information governance within the consortium and assure on going compliance.
- Provide a forum for information exchange on best practice in clinical data sharing and disclosure.

This formed a discussion and consensus on the appropriate and cost effective way of adopting/developing data provider criteria: a data provider form and advice on completing the form. The outputs from these workshops formed the basis of the deliverables D7.2 (submitted M6) and the current deliverable D7.3.

It is also important to note that the ELSI environment is constantly changing including recent developments around EU privacy law (European Commission - Press release. Agreement on Commission's EU data protection reform will boost Digital Single Market. Brussels, 15 December 2015, http://europa.eu/rapid/press-release_IP-15-6321_en.htm). In particular in December 2015 an exemption was granted to the use of clinical and medical data, which will be able to be used and shared under consent (more details in Annex 6.1). We will ensure that PhenoMeNal will keep such developments under review.

3.2. Discussion of Data Provider Forms and Guidelines

The developments of the ELSI guidelines and data provider forms were discussed in detail and the approach taken in the EU project BioMedBridges (http://www.biomedbridges.eu) was seen as the most appropriate, advanced, and an excellent starting point that, with suitable modification, would serve as the template for the PhenoMeNal project.

A significant amount of work has been performed in ELSI by BioMedBridges. The project has now ended and has been superseded by CORBEL (Coordinated Research Infrastructures Building Enduring Life-science Services, which will build on the output of BioMedBridges and is a four-year €14.5 million EU project that will harmonise user access to biological and medical technologies, biological samples and data services, required by cutting-edge biomedical research). Building on BioMedBridges /CORBEL work is seen as an appropriate choice at this stage.

The underlying information on the use of data and the relationships between parties has been explored and formulated in a number of templates. These are based on the BioMedBridges Ethical Governance Framework document at http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/BioMedBridges%20Ethical%20Governance%20Framework_V1-1.pdf.

Using this as a starting point, template agreements were developed which cover many of the common transactions between data providers and users.

The following templates from http://www.biomedbridges.eu/deliverables/52-0 were seen as key exemplars for use in PhenoMeNal:

- Participant information sheet and consent form (adult): (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/bmb_pis_and_consent_form_adult_v2_final.docx)

- Participant information sheet and assent form (child): (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/bmb_pis_and_assent_form_child_v1_110714.docx)

- Data provider agreement - non-personal data: (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_data_provider_agreement_-_non-personal_data.pdf)

- Data provider agreement - personal data: (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_data_provider_agreement_-_personal_data.pdf)

- Data transfer agreement - non-personal data (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_data_transfer_agreement_-_non-personal_data.pdf)

- Material transfer agreement - non-personal biosamples (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_material_transfer_agreement_-_non-personal_biosamples.pdf)

- Material transfer agreement - personal biosamples (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_material_transfer_agreement_-_personal_biosamples.pdf)

- Provider agreement - human biosamples (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_provider_agreement_-_human_biosamples.pdf)

- Data transfer agreement - personal data (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_data_transfer_agreement_-_human_data.pdf)

- Provider agreement - non-personal biosamples (http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/deliverables/template_provider_agreement_-_non-personal_biosamples.pdf)

In order to secure data deposition and access, the guidelines and data forms should be incorporated into a technical solution that adheres to the guidelines and offers a safe and secure environment for data deposition. We discussed in detail a template for such a solution currently offered by EMBL- EBIs' European Genome Phenome Archive (EGA) where all data must be affiliated to biomed research/consortium projects. Here a formal application procedure with data access agreement with an associated committee, Data Access Committee (DAC) is required for both data deposition and data access. Each EGA dataset is governed by at least one DAC, e.g. the Wellcome Trust Case Control Consortium. Since conception in 2008, the EGA now has more that 2PB of data and in excess of 6000 users. To upload data requires an 'Intention to submit' form with a description of the controlled access requirements. The submission must be compliant with laws and ethical regulations of the source countries. All data is Pretty Good Privacy (PGP) encrypted before upload and is then stored in the EGA secure 'drop box' area. To securely archive at the EGA, data is decrypted with PGP and moved to a private EGA vault, encrypted with the Advanced Encryption Standard (AES256), and stored in the archive. To access data, a user submits a 'Data Access Agreement' form and applies to the appropriate DAC. The DAC in question can then permit access to the applicant via online DAC Administrator tools, provided by the EGA. Access to the DAC Admin tools requires the use of an RSA key, provided by EGA to each DAC. EGA user accounts are for one individual only e.g. a PI is not allowed to share login etc. with others - each person must have an individual account approved by the DAC. The EGA has been subjected to a total of four security audits, passing all. Importantly, everything is traceable within the EGA ecosystem. This provides an opportunity for PhenoMeNal to use the EGA data access portal and framework for federated rights-managed access where data is held locally and accessed remotely, e.g. via cloud systems. Technical discussions are continuing as the project proceeds.
It should be noted that PhenoMeNal is developing workflows composed of open-source components developed by the world-wide metabolomics community. One of the major achievements of PhenoMeNal will be to scale those components to work with big data and to ensure that they interoperate. The components themselves will remain under the control of their original authors unless they decide otherwise. Due to this working principle, the PhenoMeNal consortium cannot guarantee that the components in principal safe to use in unprotected compute environments or that no backdoors exist that allow unauthorised access to the data submitted to the PhenoMeNal pipelines. For this reason, our general guidance to users with sensitive data will be to use the PhenoMeNal workflows on their own, local, secure cloud with appropriate safeguards and only to distribute patient derived data under strict ELSI considerations within a secure framework such as EGA.

3.3. Conclusions

In combination, the Ethical Governance Framework document and the appropriate selection of a template allows the creation of a collaboration agreement or Materials Transfer Agreement (Data). For example, the Data Transfer Agreement (DTA) for Non-Personal Data governs the transfer and use of non-personal data (animal or fully anonymized) that is made available by a provider to the entity that wishes to use this research data for its own research purposes (recipient). It is designed for cases, where no cooperation agreement exists between the contractors. This will be a common situation in PhenoMeNal where data will be transferred and deposited at other institutions within the EU. It should be noted that a DTA should be concluded between legal entities, which are to be bound by the contractual provisions, not between individual scientists involved in the transfer and the related research, since they would not be able to guarantee the implementation of the contractual obligations. PhenoMeNal will adopt this approach in data transfer agreements while properly informing participants of the ELSI guidelines and constraints outlined in the BioMedBridges Ethical Governance Framework.

The EGA provides a technical solution to data deposition, archiving and access, which will be used as a template for PhenoMeNal.

4. Delivery and Schedule

The delivery is delayed: No.

5. Background Information

Patient and research subject data is very sensitive, and it is paramount to establish a robust governance framework for overall information management including sensitive data. The PhenoMeNal e-infrastructure will be able to cope with data generated from comprehensive clinical, genotypic, 'omics and analytic sources including medical records, electronic health records, clinical measurements, genotypic data, phenotypic data from tissue and biofluid analysis, image and pathology data. Primarily, all data collected and held within the project will comply with all local laws, regulations and ethics. All personal information will be processed in accordance with accepted Data Protection Principles outlined above. Responsibility for data will be with the host institution/data provider.

This deliverable refers to task 7.4: Evaluate the introduction of a data provider form to be completed by each data provider to the project, the intention of which is to ensure that all ethical aspects of making data available within PhenoMeNal has been addressed, similarly to what implemented in the BioMedBridges project.

Person Month per Participant	ICL	EMBL-EBI	IPB	UB	UL	UOXF	SIB	UU
	1.5		0.2			0.4		

6. Annexes

6.1. Background information on ELSI requirements and recent changes.

http://www.nature.com/news/european-medical-research-escapes-stifling-privacy-laws-1.19054#auth-1**.** The Regulations underpinning the legal framework on the protection of personal data in the EU is under review. It will replace the current EU Directive and its remit includes the use of personal data in research (given previous patients approval). A draft was produced by the European Commission in January 2012 and is being reviewed by the European Parliament and Council of Ministers. Since June 2015, the Commission, Parliament and Council of Ministers have been negotiating a final draft of the Regulation – a process known as trilogue. Negotiations concluded in December 2015, meaning the final text can be voted on and legal checks progressed, therefore a final version may be approved in 2016. This article in Nature reviews the current changes anticipated in EU policy with respect to the handling of medical data. One amendment has removed an exemption for researchers from a rule that all personal data remain anonymous in perpetuity. This would have created problems for scientists who need to unmask data to assess the progression of diseases and the long-term outcomes of treatments. Another policy required researchers to obtain fresh consent from donors each time their data or tissues are used in a different study, something that medical scientists say is unworkable. But the compromise agreement between the European Parliament, Council and Commission now permits medical researchers to unmask data in special circumstances and to re-use an individual's data and samples for multiple studies in different diseases, provided that person agrees to a general type of consent that covers all ethically and scientifically approved studies.
Additional relevant discussions are available at the following URLs:

- **Data Protection Legislation:** http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Personal-information/Data-protection-legislation/.
- As a preamble to the terminology and legal position to the use of Clinical Trial data, the following is useful: https://www.cov.com/~/media/files/corporate/publications/2005/10/oid64167.ashx. In particular, this discussion highlights the differences between EU member states in the allowed use of patient derived data.
- BioMedBridge: http://www.biomedbridges.eu/
- EGA (European Genome Phenome Archive): https://www.ebi.ac.uk/ega/home