



Deliverable D4.1

Report on Materials Requirements in Systemic Toxicology

Grant Agreement	HEALTH-F5-2010-267042
Acronym	ToxBank
Name	ToxBank – Supporting Integrated Data Analysis and Servicing of Alternative Testing Methods in Toxicology
Scientific Coordinator	Douglas Connect (DC)
Administrative Coordinator	Istituto Di Ricerche Farmacologiche Mario Negri (IRFMN)

Contract No.	HEALTH-F5-2010-267042
Document Type:	Deliverable Report
WP/Task:	4.1
Document ID:	ToxBank D4.1
Version:	1.0
Date:	24 th March 2012
Status:	Draft

Organisations:	Douglas Connect (DC) Istituto Di Ricerche Farmacologiche Mario Negri (IRFMN) Leadscope (LEAD) Ideaconsult (IDEA) Pharmatropo (PHT) National Institute Biological Standards UK (NIBSC) In Silico Toxicology (IST) Karolinska Institutet (KI)
Authors:	Glyn Stacey (NIBSC-HPA)

Distribution:	Partnership, EC
---------------	-----------------

Purpose of Document:	To Report the requirements for data in the ToxBank database on biological materials (cells and reagents)
----------------------	--

Document History:	1 - First draft, 24 th March 2012 2- Second draft, 30 th March 2012
-------------------	--

--	--

Table of Contents

1. Executive Summary	5
1.2. Requirements gathering	5
1.3. ToxBank overview	5
2. Introduction	7
3. Requirements Analysis	7
3.1. Development Process	7
3.2. Draft Outline of Biomaterials Requirements	7
3.3. Consultation	10
3.4. Interpretation	10
3.5. Compilation of Biomaterials Requirements	11
4. Conclusions	14
5. Ongoing Development	14
6. Acknowledgements	15
Appendix 1: Outline Biomaterial Requirements PPT	
.....	16

1. Executive Summary

1.1. The ToxBank data warehouse

The ToxBank data warehouse will provide a web accessible shared repository of know-how and experimental results to support the SEURAT-1 cluster in developing a replacement for *in vivo* repeated dose toxicity testing. The information within the ToxBank data warehouse is uploaded from the research activities of the cluster partners as well as relevant data and protocols from other sources, such as public databases. The data will be collected to enable a cross-cluster integrated data analysis leading to the prediction of repeated dose toxicity. The warehouse should continue to provide access to this knowledge after SEURAT-1 for academic and industrial uses, as a potentially self-sustainable operation.

1.2. Requirements gathering

Prior to designing the ToxBank data warehouse, the ToxBank consortium implemented a detailed requirements gathering exercise. As part of this process, ToxBank partners visited around 20 partner's sites and conducted interviews with individual scientists. These discussions covered a variety of activities including, cell differentiation, cell engineering, biomarker identification, dose response analysis, toxicity testing, 'omics experiments, chemical analysis, and cell banking. The interviews focused on understanding in detail what specific steps were performed across a variety of tasks performed, and this can only be accurately recorded by observing the work. Notes were taken along with examples of documents used. This information was collected to ensure any system design both meets the needs of scientists across the entire cluster at the same time as fitting within current workflows. The interviews along with other requirements gathering exercises resulted in over 1,000 separate notes and 40 tasks outlined. The ToxBank team organized and analyzed the notes, descriptions of tasks, and associated documents as a group and developed a design for the ToxBank data warehouse directly from this analysis. The design was subsequently tested using a paper prototype where the major components of the user interface were sketched out on paper. This allowed for further testing of the usability of the system, resulting in a refinement of the user interface where problems were encountered using this mock-up to perform different scenarios.

1.3. ToxBank overview

ToxBank is being developed to manage and provide access to all protocols and experimental data across SEURAT-1 to support an integrated data analysis.

Once a new protocol has been developed, documented, and reviewed within the partner's organization, it can be uploaded to the ToxBank data warehouse by the laboratories' principal investigator. ToxBank will provide guidelines concerning the content and organization of this document. The protocol will be loaded through the ToxBank user interface where additional information will be entered and associated with the protocol. This includes summaries of the protocol, identification of the protocol's owner, authors of the protocol, and a specification of who should have access to the protocol. In addition, keywords based on a cross-cluster keyword hierarchy, will be assigned to support searching. Study data is loaded in a similar manner;



however, a protocol must have been already loaded that defines how individual steps of the study were performed and what data is generated. The data should be in a defined and standardized format agreed across the cluster. Once any new protocols or study data are loaded into the system, a regularly scheduled email alerting scientists across SEURAT-1 who have registered an interest in a specific type of information is sent out.

The protocols and data loaded can be accessed via a simple free text search. This will return summaries of any information matching the query. The protocols or study data can then be viewed or downloaded directly along with links to related information, such as the Gold Compound wiki.

Where the investigator does not have permission to view the specific protocol or study data, only the summary information will be displayed. The investigator is then free to contact the principal investigator who loaded the content to request access rights. ToxBank will provide documents to support any bilateral agreements between the two parties. Once an agreement is in place, the principal investigator who loaded the information would modify permission levels accordingly.

ToxBank Vision

The focus of the first phase of the ToxBank data warehouse project is the development of the unified data access. As this is being implemented over the next year, the ToxBank consortium will continue to collect requirements for the integrated data analysis to be implemented as phase 2 of the project.

The benefits of this approach include:

- The approach provides access to existing and new protocols and data, as well as facilities for uploading information through a simple web interface
- The use of standardized data templates and controlled terms supports cross-cluster experimental consistency and will enable an integrated analysis
- The approach supports protocol development and collaboration which is close to current work activities, especially the SEURAT-1 focus on experimental development
- This approach will link public databases and in-house data

2. Introduction

The following report outlines the requirements gathering process that was performed for biomaterials in order to design the ToxBank data warehouse. This included collecting data from interviews conducted across the entire SEURAT-1 cluster. This data was carefully organized and analyzed and used to define how the ToxBank data warehouse will operate conceptually. A preliminary report was presented at the SURAT-1 meeting in Cascais March 2011. Further developments of the biomaterials requirements were presented as a poster at the February 2012 SEURAT-1 conference.

3. Requirements Analysis

3.1. Development Process

Prior to designing and building the data warehouse, the ToxBank consortium initiated an extensive biomaterials requirements gathering exercise. Delivery of an effective solution for scientists across the SEURAT-1 cluster, would require close interaction with the SEURAT-1 life scientist partners. A methodology referred to as contextual inquiry/design was utilized for construction of the main database structure. However, in the case of biomaterials the responsible partner NIBSC/UKSCB had established considerable experience with the development of a stem cell database

The Process can be divided into 4 stages:

- Step 1. Draft outline for the biomaterials requirements
- Step 2. Consultation with scientists
- Step 3. Interpretation and consolidation
- Step 4. Delivery to systems design team

Step 1. Draft outline for the biomaterials requirements.

A draft outline for the biomaterial requirements was constructed and discussed with other Toxbank partners.

Step 2. Consultation with scientists

Key scientists with most relevant knowledge were invited to give feedback on the draft outline for the biomaterials requirements.

Step 3. Interpretation. Compare response from different contributors and rationalize responses into an annotation of the original

Step 4. Delivery of biomaterial requirements to systems design team. Coordinate with ToxBank designers to translate biomaterials requirements into the new structures for the data-warehouse.

3.2. Draft Outline of Biomaterial Requirements

The draft outline for the biomaterials requirements was constructed in the first quarter of 2011. The development of the draft was reported at ToxBank tele-conference calls and presented at the ToxBank session of the March 2011 SEURAT-1 meeting for further discussion. Unlike the system used in ToxBank for the preparing main structure of the data-warehouse the WP leader NIBSC had significant existing knowledge in the development of cell line and stem cell data-bases (see www.hescreg.eu/ , www.ukstemcellbank.org.uk/) and detailed knowledge of key parameters of importance for cell culture and stem cell characterization. When beginning work to understand the needs of those interested in biological materials in toxicology a number of questions were considered of relevance to the needs of the cell culture experimentalist and these were reported at the SURAT-1 meeting as follows:

- What data is needed in planning the experiment?
- What is the experimental protocol and how will it be defined?
- What data output will the experiment generate?
- How will the data be recorded and stored?
- What processing will be carried out on the data?
- What analysis is needed on the data?
- What information from other sources will be required for analysis?

In addition the types of data required in planning a cell culture experiment were considered important and were also reported as follows:

1. Cell identifiers
2. Supplier specific data: cells, SOPs
3. Published information: known characteristics of specific cells
4. Tissues and primary cells
5. Specialist cell culture requirements
6. Culture conditions, species differences, genotypic stability
7. (National) legislative issues.
8. Different “-omics” types

Utilising experience in construction of the hESCReg database and the USCB website catalogue the following core elements of cell line descriptors were identified :

- **Name of cell line**
- **Supplier reference (critical identifier for traceability)**
- **Pluripotent Stem Cell Line Nomenclature**
- **Cell origins (tissue: type, disease mutation present)**
- **Derivation method:**

- hESC isolation method
- iPSC method and constructs
- **Ethical issues:**
 - Fully informed consent (statement from originator – EC and hESCreg database)
 - Donor constraints

Suppliers of stem cell lines and other cells need to operate in an optimal way to provide sufficient data on the cells they release and core information of relevance to ToxBank would include issues surrounding the ethics of using human tissues and cells, scientific characterisation and access arrangements. These were identified under the following headings:

- **Supplier identity and contacts**
- **Minimal data set generated (NB specific to each supplier)**
 - Quality Control
 - Safety testing
 - Characteristics
- **Access arrangements**
- **General description of charges and application process**
- **Summary of terms and conditions, agreements and supplier constraints**

Types of published data included:

- **Scientific literature**
 - Derivation
 - Differentiation
- **SOPs:**
 - Culture, preservation and Differentiation (weblinks)
 - NB multiple SOPs for same cell type and multiple sources.
 - Origin ref and version numbers as will change frequently

Specialist cell culture data would also be a feature of SEURAT-1 and the key biomaterials aspects were drafted as follows:

- IPSC growth and differentiation
- Hepatic cell lines
- High throughput
- Materials requirements for cellular barrier assays (4.2.9-2c)
- Materials requirements for 3D architectures, bioreactors, specialist cell culture applications and miniaturised sub-systems (4.2.9-2, b & d)

Information types for tissues and cells used in SEURAT-1 were identified as follows:

- Sourcing
- Harvest/isolation, preservation, culture and stability
- QC assessment criteria for release and use
- SOPs for cell preparation
- Ethics and other issues including transport described within supplier info.

It was decided that it would also be important to include requirements to identify the type of experimental procedure and how the respective protocol will be defined. Key elements to be included in the biomaterials requirements were:

- Culture and QC of cells pre-assay
- Delivery of cells in form ready for assay
- Method for recording and evaluating cell responses
- Mechanism to enable comparison of data generated at different times and in different centres: references materials (aliquotted compounds from single bulk prep stored in a stable state)
- Feedback on methodologies from users

These outline biomaterials requirements were compiled as a powerpoint presentation (appendix 1) for dissemination.

3.3. Consultation

The powerpoint presentation (appendix 1) was presented to ToxBank partners and other SEURAT-1 scientists for discussion and feedback at the first SEURAT-1 annual meeting (March 2011). This and other discussions at the SEURAT-1 meeting were used to identify a panel of SEURAT-1 senior biologists to be recruited to provide feedback and detailed biomaterials requirements. These scientists were subsequently recruited by telephone and email from DETECTIVE, HemiBio, NOTOX and ScrnTox. Having distributed the power point draft biomaterials requirements document feedback was obtained by telephone conference calls and email. In addition the biomaterials requirements were discussed at the 3-day ToxBank data-warehouse workshop meeting in Milan July 2011 (see deliverable 1.1).

3.4. Feedback and Interpretation

The feedback from the recruited scientists in DETECTIVE, HemiBio, and ScrnTox was used to annotate the draft outline materials requirement document. In some cases further clarification of feedback was necessary by telephone conference calls or email. Returned annotated ppt documents and other feedback was compiled as a single set of annotated ppt slides which were prepared as a poster presented at the February 2012 2nd SEURAT-1 annual meeting in Lisbon with coauthors L Healy (NIBSC), G Myatt (LeadScope) and B Hardy (DouglasConnect).

A number of respondents identified certain specific biomaterials requirements that were not yet developed but would be available from month 18 of their projects which were also included in the annotated slides presented in Lisbon. In particular responses on - omics data were very general and there were no proposals for the types of data that ToxBank should hold.

3.5 Compilation of Biomaterials Requirements

The Tables below give the consolidated biomaterials requirements consolidate from responses from DETECTIVE, HemiBio and ScrnTox and other discussions with ToxBank partners at the first SEURAT-1 annual meetings and the ToxBank workshop in July 2012.

A) Cell Lines Descriptors

- Name of cell line
- Supplier reference (critical identifier for traceability)
- PSC Line Nomenclature. Proposed nomenclature as published by Luong et al., (2011) Cell Stem Cell.
- Cell origins: tissue: type, disease mutation present, cell line passage number/information on genotypic and phenotypic stability
- Derivation method:
 - hESC isolation method
 - iPSC method and constructs
- Ethical issues:
 - Fully informed consent (statement from originator - EC and hESCreg database)
 - Donor constraints

B) Supplier Specific Information

- Supplier identity and contacts
- Minimal data set generated (standards established in IISCB (2009) Stem Cell Reviews and Reports. NB these will be specific to each supplier and will be evaluated in later ToxBank deliverables.

- Quality Control
- Safety testing
- Characteristics
- Access arrangements (fees, application process and ethics review)
- General description of charges and application process
- Summary of terms and conditions, agreements and supplier constraints

C) Publicaitons

- **Scientific literature/ and other resources**
 - Derivation
 - Differentiation
- **SOPs:**
 - Culture, preservation and Differentiation (weblinks to supplier data)
 - NB multiple SOPs for same cell type and multiple sources.
 - Origin reference and version numbers as will change frequently (this may be too dynamic to track within Toxbank systems)

D) Types of Information Partners wish to be available

- IPSC growth and differentiation (**HeMiBio and Screentox input**)
- IPSC stability – Epigenetic data on DNA methylation patterns and histone modification (**will be made available from DETECTIVE and HemiBio**). MicroRNA profiles considered vital (**HemiBio to provide later in project**). Transcriptome data (**HemiBio to provide later in project**).
- Hepatic cell lines (**NOTOX to be approached again for input**)
- High through put (**Screentox input to be sought as their HTP work develops**)
- Materials requirements for cellular barrier assays (4.2.9–2c) **skin/blood brain etc (the Hemibio project will be able to provide information on endothelial cell barrier model from 18months)**
- Materials requirements for 3D architectures, bioreactors, specialist cell culture applications and miniaturised sub–systems (4.2.9–2, b & d) – no specialists in

- cluster – catalogue (**bioreactor requirements will be developed in the HeMiBio project**).
- Culture conditions, species differences, genotypic stability
 - National legislative issues on stem cell use.
 - Lists of reagents: antibodies, PCR primers, factors for PSC line for differentiation to hepatocytes (**will be available from Hemibio including miRNA expression in 2011**)
 - Information on development of bioreactors (Hemibio from 18months)
 - Detective:
 - Genomics: Affym. Gene expr arrays
 - Proteomics: 2D Page and selective characterisation by MALDI-TOF, QTOF MS
 - Metabonomics: GC-MS, NMR spectroscopy
 - Functional readouts
 - Electrophysiology data for cardiomyocytes (LAT, ARI, Q-T interval, beat freq)
 - Impedence measurement (Xcelligence, Roche)
 - High Content Imaging
- (HeMiBio and DETECTIVE projects, are currently preparing SOPS for all the cytomic and molecular assays that will be carried out on primary hepatocytes with the purpose of probing/monitoring their functionality)**

E) Tissues and Primary Cells

- Sourcing
- Harvest/isolation, preservation, culture and stability
- QC assessment criteria for release and use
- SOPs for cell preparation: stem cells, feeder cells, genetic modification and hepatocyte selection for tox assays (**Protocols will be available from DETECTIVE and HeMiBio in 2011**).
- Ethics and other issues including transport described within supplier info.

F) Defining Experimental protocols

- Culture and QC of cells pre-assay (**coordination between ToxBank and ScrnTox**)
- Delivery of cells in form ready for assay
- Method for recording and evaluating cell responses
- Mechanism to enable comparison of data generated at different times and in different centres: reference preparations (aliquotted compounds from single bulk prep stored in a stable state)
- Qualification data on research protocols

G) Other data issues to be considered

- SOPs: published research protocols to be distinguished from SOP – SOPs will develop with time and require qualification data
- User feedback on cells/methods: this will require a special area in ToxBank for interactive exchange (now under development as Wiki pages in the ToxBank data system).

4. Conclusions

The core biomaterial requirements circulated to selected SEURAT-1 partners have met with general agreement and a range of supplementary data types were proposed and have been incorporated (see 3.4).

It is clear that the biomaterials requirements will develop as SEURAT-1 progresses and ToxBank will need to sustain its interaction with other consortia to ensure it maintains data of interest in the database.

Working with key partners in the ToxBank team (notably LeadScope and DouglasConnect) these biomaterials requirements will now be incorporated into the main ToxBank data warehouse to provide valuable information resource for SEURAT- scientists.

5. Ongoing Development

Both DETECTIVE and HemiBio identified a number of datasets that would be desirable to host on the ToxBank system (see table below). These will begin to become available in 2012.

Communications between ToxBank and other partners will be maintained to accommodate these developing needs for access to data as the project develops and the data becomes available.

The biomaterials section of the ToxBank data warehouse is now being developed as Wiki pages for reagents and cell lines with links to suppliers.

Biomaterials requirements to be considered as SEURAT-1 progresses

- IPSC stability – Epigenetic data on DNA methylation patterns and histone modification (**will be made available from DETECTIVE and HemiBio**). MicroRNA profiles considered vital (**HemiBio to provide later in project**). Transcriptome data (**HemiBio to provide later in project**).
- Hepatic cell lines (**NOTOX to be approached again for input**)
- High through put (**Screntox input to be sought as their HTP work develops**)
- Materials requirements for cellular barrier assays (4.2.9–2c) **skin/blood brain etc (the HemiBio project will be able to provide information on endothelial cell barrier model from 18months)**
- Materials requirements for 3D architectures, bioreactors, specialist cell culture applications and miniaturised sub-systems (4.2.9–2, b & d) – no specialists in cluster – catalogue (**bioreactor requirements will be developed in the HeMiBio project**).
- Lists of reagents: antibodies, PCR primers, factors for PSC line for differentiation to hepatocytes (**will be available from HemiBio including miRNA expression in 2011**)
- HeMiBio and DETECTIVE projects, are currently preparing SOPs for all the cytomic and molecular assays that will be carried out on primary hepatocytes with the purpose of probing/monitoring their functionality
- SOPs for cell preparation: stem cells, feeder cells, genetic modification and hepatocyte selection for tox assays (**Protocols will be available from DETECTIVE and HeMiBio in 2011**).

6. Acknowledgements

All DETECTIVE, HemiBio and ScrnTox scientists who contributed to the development of the biomaterials requirements are thanked for their time and helpful discussions.

Appendix 1: Draft Outline Biomaterials Requirements Power Point Presentation

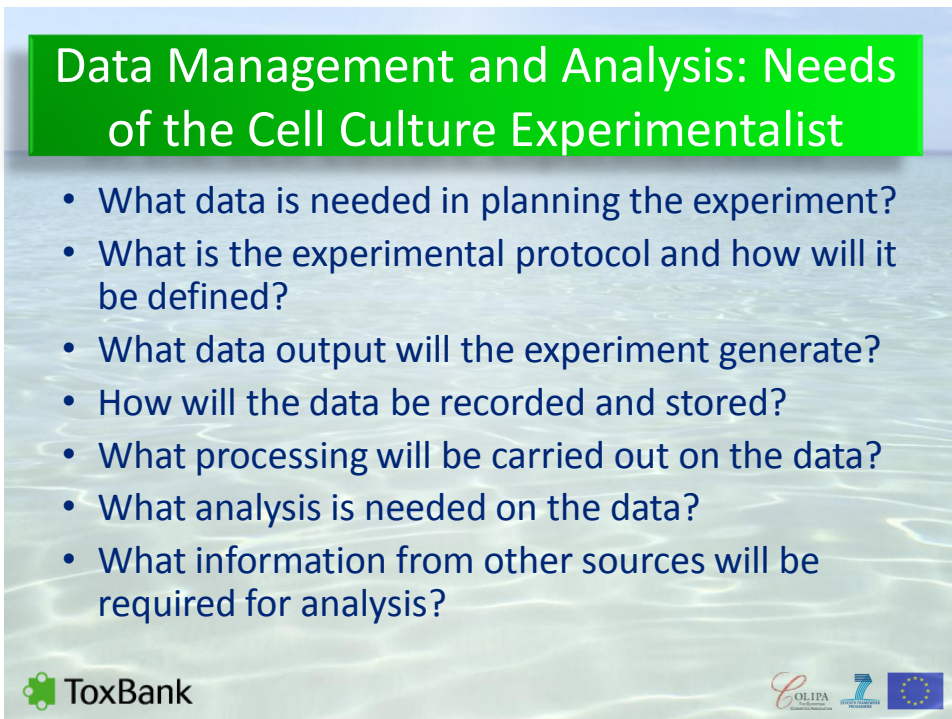


**Draft Feedback on
Biological Materials
Requirements**

Glyn Stacey
UKSCB, NIBSC-HPA

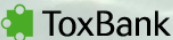



Seurat Meeting, Cascais, 2nd March 2011

 ToxBank   



**Data Management and Analysis: Needs
of the Cell Culture Experimentalist**

- What data is needed in planning the experiment?
- What is the experimental protocol and how will it be defined?
- What data output will the experiment generate?
- How will the data be recorded and stored?
- What processing will be carried out on the data?
- What analysis is needed on the data?
- What information from other sources will be required for analysis?

 ToxBank   

What data is needed in planning the experiment?

1. Cell identifiers
2. Supplier specific data: cells, SOPs
3. Published information: known characteristics of specific cells
4. User feedback on cells/methods: enable conti
5. Tissues and primary cells?
6. Specialist cell culture requirements
7. Culture conditions, species differences, genotypic stability
8. (National) legislative issues.
9. -omics X5!
10. Details provided from Hemibio and Detective

Cell Line Descriptors

- Name of cell line
- Supplier reference (critical identifier for traceability)
- iPSC Nomenclature? (Cell Stem Cell in press)
- Cell origins (tissue: type, disease mutation present,
- Derivation method:
 - hESC isolation method
 - iPSC method and constructs
- Ethical issues:
 - Fully informed consent (statement from originator – EC and hESCreg database)
 - Donor constraints

Supplier Specific Information

- **Supplier identity and contacts**
- **Minimal data set generated (standards - ISCBI 2009; CSC 2010) NB specific to each supplier**
 - Quality Control
 - Safety testing
 - Characteristics
- **Access arrangements**
- **General description of charges and application process**
- **Summary of terms and conditions, agreements and supplier constraints**

Publications

- **Scientific literature**
 - Derivation
 - Differentiation
- **SOPs:**
 - Culture, preservation and Differentiation (weblinks)
 - NB multiple SOPs for same cell type and multiple sources.
 - Origin ref and version numbers as will change frequently

What types of information would you like to be available about the following?

- IPSC growth and differentiation – , HeMiBio and **Screentox input**
- IPSC stability
- Hepatic cell lines - **NOTOX input?**
- High through put – **Screentox input?**
- Materials requirements for cellular barrier assays (4.2.9-2c) **skin/blood brain etc**
- Materials requirements for 3D architectures, bioreactors, specialist cell culture applications and miniaturised sub-systems (4.2.9-2, b & d) – no specialists in cluster – catalogue

Tissues and primary cells:

- Sourcing
- Harvest/isolation, preservation, culture and stability
- QC assessment criteria for release and use
- SOPs for cell preparation (stem cells) for tox assays.
- Ethics and other issues including transport described within supplier info.
- Discuss with HeMiBio, NOTOX, Detective, Screentox partners

What is the experimental protocol and how will it be defined?

- Culture and QC of cells pre-assay
- Delivery of cells in form ready for assay
- Method for recording and evaluating cell responses
- Mechanism to enable comparison of data generated at different times and in different centres: references materials (aliquotted compounds from single bulk prep stored in a stable state)

What data output will the experiment generate?

- SOPs (Hemibio)Published information NB will develop with time
- User feedback on cells/methods
- Details to be provided from Hemibio, Detective, SCR&Tox, Notox