

D7.1

Data input and input interface

Project number:	668858
Project acronym:	PrECISE
Project title:	PrECISE: Personalized Engine for Cancer Integrative Study and Evaluation
Start date of the project:	1 st January, 2016
Duration:	36 months
Programme:	H2020-PHC-02-2015

Deliverable type:	OTHER
Deliverable reference number:	PHC-668858 / D7.1 / 1.0
Work package contributing to the deliverable:	WP 7
Due date:	April 2017 – M16
Actual submission date:	27 th April, 2017

Responsible organisation:	BCM
Editor:	Pavel Sumazin
Dissemination level:	PU
Revision:	V1.0

Abstract:	This report describes graphical user interfaces developed for genomic and clinical data entry for PrECISE.
Keywords:	Graphical user interface, secure databank



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 668858.

This work was supported (in part) by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0324-2. The opinions expressed and arguments employed therein do not necessarily reflect the official views of the Swiss Government.

Editor

Pavel Sumazin (BCM)

Contributors

Pavel Sumazin (BCM)

Zsolt Torok (ABT)

Disclaimer

The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The users thereof use the information at their sole risk and liability.

Executive Summary

ABT has developed a graphical user interface to securely deposit genomic and phenotypic data, as well as analytic products into a secure database environment that fosters collaboration within PrECISE, and can be released to the general public at the completion of the project. In this report we describe their developed interfaces.

Together with this short report, we provide the official user manual for the SmartBiobank.

Contents

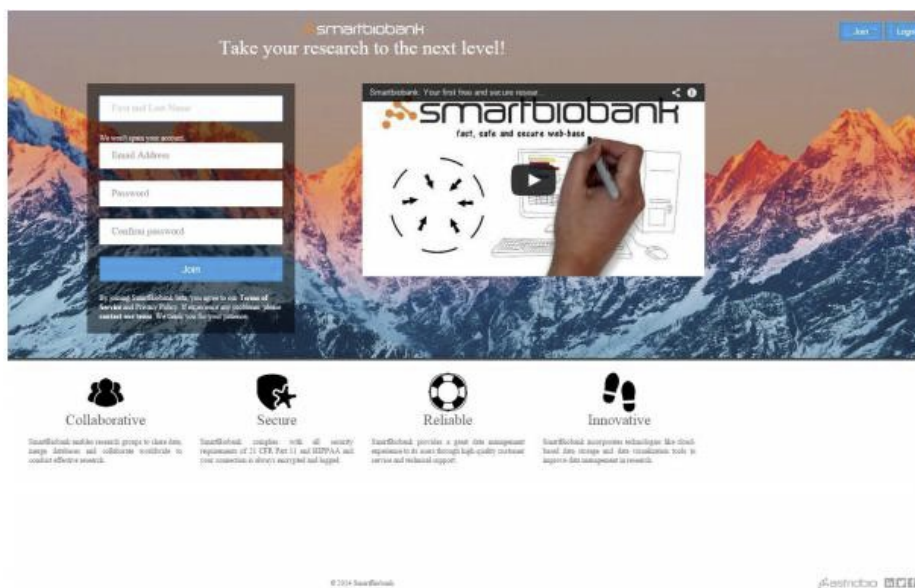
Chapter 1	SmartBiobank	1
1.1	User registration.....	1
1.2	Setting up a new project within PrECISE.....	1
1.3	User dashboard.....	2
1.4	Code validation.....	2
Chapter 2	Security	3
2.1	Project properties.....	3
2.2	Project owners.....	3
2.3	Administrators.....	3
2.4	Biobank users.....	3
2.5	Clinical users.....	3
2.6	Laboratory users.....	4
Chapter 3	Biospecimen management	5
3.1	Biospeciman location.....	5
3.2	Biospecimen identifier.....	5
Chapter 4	Tracking user activity	6
Chapter 5	Research tools	7
5.1	Descriptive statistics.....	7
5.2	Charts.....	7
5.3	Visualizing relationships between data.....	8
5.4	Search & Query.....	8
5.5	Data visualization.....	8
Chapter 6	Data exchange	10
Chapter 7	SmartBiobank instruction manual	11
Chapter 8	Summary and Conclusion	12

Chapter 1 SmartBiobank

SmartBiobank is an online biobank software designed to help clinicians, lab biologists and researchers to integrate translational research elements into one single system. SmartBiobank can store, manage, analyze, and report on the analysis of biospecimen data, clinical data as well as experimental data. It supports the establishment of population or clinical biobanks as well as patient registries. SmartBiobank has been customized for PrECISE data entry, analyses, and their sharing. We envision that at the end of the project PrECISE SmartBioBank could be made available to the general public.

1.1 User registration

Each PrECISE user can register to use SmartBiobank. Users that entered data are able to control and release it to other PrECISE users that apply to view or edit the data. The registration page is depicted below.



1.2 Setting up a new project within PrECISE

Users are able to setup a new project within PrECISE (see user interface below). They can then control permissions of other users that are interested in data deposits, data analysis, and visualization within this project. Project owners can also invite participants to view data and analyses

Add new SmartBiobank project

Choose a project name and a domain. [Details](#)

Project name

Project domain .smartbiobank.com

1.3 User dashboard

Each project is given a user dashboard (below). This dashboard allows to control participant permission, view and upload bio specimen, organize bio specimen, describe project details, design questionnaires based on data for users or wider audiences, control profiles, share and visualize data, query data based on metadata description, and query relationships within the data. The dashboard contains tiles that lead to these functions. You can drag and drop tiles to reorganize them according to your preferences. The Settings, Data exchange, Biospecimen organizer and Favorites categories are empty and invisible by default but they appear when you drag a tile over the page. Once the page is refreshed, the Hidden tiles category disappears. Pull down the Quick settings menu on the upper right corner of the Dashboard to select a new color theme or to log out. There is also a Logout tile among the hidden tiles. The default dashboard is given below. The accompanying instruction manual includes details about available functions.

SmartBiobank welcomes you

Biobanking

- Participants
- Biospecimens
- Biospecimen organizer

Project design

- Project details
- Questionnaires
- Profiles
- Data exchange
- Settings

Research Center

- Data visualization
- Search & Query
- Descriptive statistics
- Variable Relationship

[Show hidden tiles](#)

PRECISE
This SmartBiobank project does not have any description.
[Modify access](#)

Getting started

- [SmartBiobank User manual](#)
- [Ask the experts](#)
- [Visit SmartBiobank™ blog](#)
- [Start another SmartBiobank™ project](#)

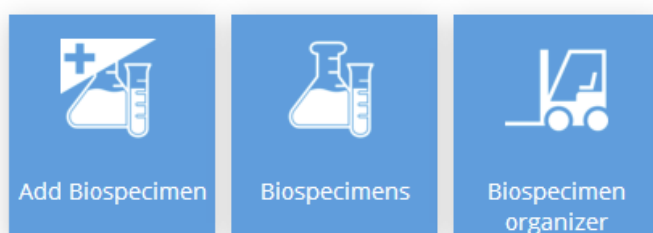
1.4 Code validation

All participants and biospecimen entries have identification codes. However, consistency or precision may require to apply a strict pattern for identification codes which may be set as a *regular expression*. See the accompanying instruction manual for further details.

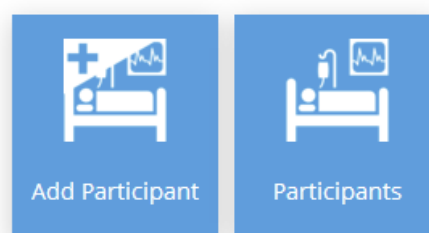
Chapter 2 Security

Security is managed by project owners and administrators, who may assign data deposit, download, analysis and visualization rights to other users. Project owners and administrators have an access to a subpage of Project details called Project users where they can manage who can have an access to the project. There are five different roles defined in the system including Project owners, Administrators, Biobank users, Clinical users, and Laboratory users. These users have access to varying sections of the database, with access as depicted below.

Biobanking



Biobanking



2.1 Project properties

Project security is controlled through users (described in sections below) and project properties. Properties are the essential elements of a project's profile, as they contain all information that define both the biobank group (participant group or biospecimen type) and each biobank entry. For participants, it can be name, data of birth, gender, etc. For biospecimens, it can be volume, holder type, etc. Biospecimen,

2.2 Project owners

Project owners are the ones who have registered a project. They are the only users who require full registration. They have access to all functions and they can invite new participants or change their roles. Only owners can remove the project. Project owners cannot be deleted from the project.

2.3 Administrators

Administrators are invited users with full access to the SmartBiobank™ functions, similarly to the project owner. They can help project owners by creating profiles and questionnaires, managing participants etc.

2.4 Biobank users

Biobank users have limited access, their activity is restricted to data input or edit. Additionally they can view some informational pages like *Project details*. This allows them to check some settings and assist the project according to these settings. They cannot import or export data and they cannot modify questionnaires or questions. If you are working with sensitive data, you may want to further restrict user access (see the last two user groups below).

2.5 Clinical users

Clinical users are biobank users only permitted to work with participant data.

2.6 Laboratory users

Laboratory users are biobank users only permitted to work with biospecimen data.

Chapter 3 Biospecimen management

SmartBiobank includes extensive features to manage anonymized patient data and biospecimen. These include management of physical biospecimen location.

3.1 Biospeciman location

The Refrigerators module allows users to click and add *refrigerators* on the sidebar. Enter the name of the refrigerator you want to add and set its temperature. Location and Description are optional fields but may help in identification if your project uses a large number of refrigerators. Refrigerators allow users to describe the physical structure of a refrigerator and its contents.

3.2 Biospecimen identifier

The location of biospecimen boxes and samples in boxes can be added to the refrigerator by following a drill-down selection: Biospecimen organizer → Biospecimen boxes → Add biospecimen box route. Biopsecimen are given unique names or codes, and are placed in labeled biospecimen boxes, in labeled refrigerators. Box sizes can range from 1×1 to 15×15 slots (the number of rows and columns does not have to be equal). Biospecimen can have unique profiles that describe properties including extraction date, specimen types, and patient ID. See the accompanying instruction manual for details about available functions. An example box is given below. These can be modified by the user as needed. Forms and derivatives can be used to populate default parameters for biopsecimen sets.

11. box

Click on a slot to select it. You can add/remove biospecimen to/from a slot.

	A	B	C	D	E	F	G	H	I
1	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY
	Biospecimen: <input type="text"/>				Store	Disable	Reserve		
2	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY
3	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY

Chapter 4 Tracking user activity

To facilitate transparency and security SmartBiobank allows project managers and administrators to track all user activity. All activity is maintained in event logs. Events are displayed by time and e-mail address of its actor. The logged events include logins and login attempts, page loads, creation, edits and deletions of biobank entries, profiles, questionnaires, questions and properties and storage, search related events: saving, loading and removing criteria, and any errors that occurred during an event:

Event log

Time	Event level	Actor	Area	Action	Event message
03/17/2014 3:38:58.0 PM	info	orsolya.ludmany@astridbio.com	Event log	index	Event log at Mon Mar 17 15:38:58 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:38:53.0 PM	info	orsolya.ludmany@astridbio.com	Dashboard	dashboard	Dashboard at Mon Mar 17 15:38:53 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:55.0 PM	info	orsolya.ludmany@astridbio.com	Question	save	Save Question [Name] at Mon Mar 17 15:34:55 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:47.0 PM	info	orsolya.ludmany@astridbio.com	Question	create	Add new question at Mon Mar 17 15:34:47 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:40.0 PM	info	orsolya.ludmany@astridbio.com	Question	create	Add new question at Mon Mar 17 15:34:40 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:24:27.0 PM	info	orsolya.ludmany@astridbio.com	Questionnaire	show	Show Profile [Patient data] at Mon Mar 17 15:24:27 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:24:24.0 PM	info	orsolya.ludmany@astridbio.com	Profiles	index	Profile management at Mon Mar 17 15:24:24 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:22:42.0 PM	info	orsolya.ludmany@astridbio.com	Dashboard	dashboard	Dashboard at Mon Mar 17 15:22:42 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:12:57.0 PM	info	orsolya.ludmany@astridbio.com	Questionnaire	show	Show Profile [Patient data] at Mon Mar 17 15:12:57 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:12:37.0 PM	info	orsolya.ludmany@astridbio.com	Profiles	index	Profile management at Mon Mar 17 15:12:37 CET 2014 by orsolya.ludmany@astridbio.com

Showing 1 to 10 of 738 entries [First](#) [Previous](#) **1** [2](#) [3](#) [4](#) [5](#) [Next](#) [Last](#)

These also include changes to metadata, and structured elements such as Profiles, Questionnaires, Properties, Questions, Participants and Biospecimens. By carefully following the log, you can back track values to the creation of the entry. See the example below.

Diastolic blood pressure

Question text	Diastolic blood pressure	Display name	Diastolic
Description	▲ not set	Hint	▲ not set
Mandatory	no	Important	yes
Type	Integer	Created by	orsolya.ludmany@astridbio.com
Created	03/18/2014 4:24 PM	Last modified	03/18/2014 4:28 PM

Integer specific properties

Minimum value	▲ not set
Maximum value	▲ not set
Unit of measure	Hgmm

Changelog

actor	date	domain	action	property	old value	new value
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Display name	Diastolic	Diastolic
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Type	Real number	Integer
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Important	false	true
orsolya.ludmany@astridbio.com	03/18/2014 4:27 PM	Question	modified	Description		minimum pressure
orsolya.ludmany@astridbio.com	03/18/2014 4:28 PM	Question	modified	Description	minimum pressure	minimum pressure

Showing 1 to 5 of 5 entries [First](#) [Previous](#) **1** [Next](#) [Last](#)

Chapter 5 Research tools

SmartBiobank™ offers many tools to make use of the data that each project stores. These tools control the research process using the feedbacks of the software and enhance your research results using statistical and visualization tools. The Research center category on the Dashboard lists functions that visualize data with the help of charts. Many charts can be accessed from more than one page. Charts can be saved as an image file or can be printed directly using the upper-right drop-down menu of the charts.

5.1 Descriptive statistics

Descriptive statistics can be used to let you monitor and improve the quality of your project by assessing the usefulness of the questions you work with, and to provide visualization tools to understand your results better. Its main page collects and displays all box plots and pie charts in the project, which helps you overview your data or visually compare results.

Descriptive statistics

Adequacy of questions

Unattached questions (4/69)

- [Date of analysis](#)
- [Expected date of delivery](#)
- [Protein concentration](#)
- [Sample taken](#)

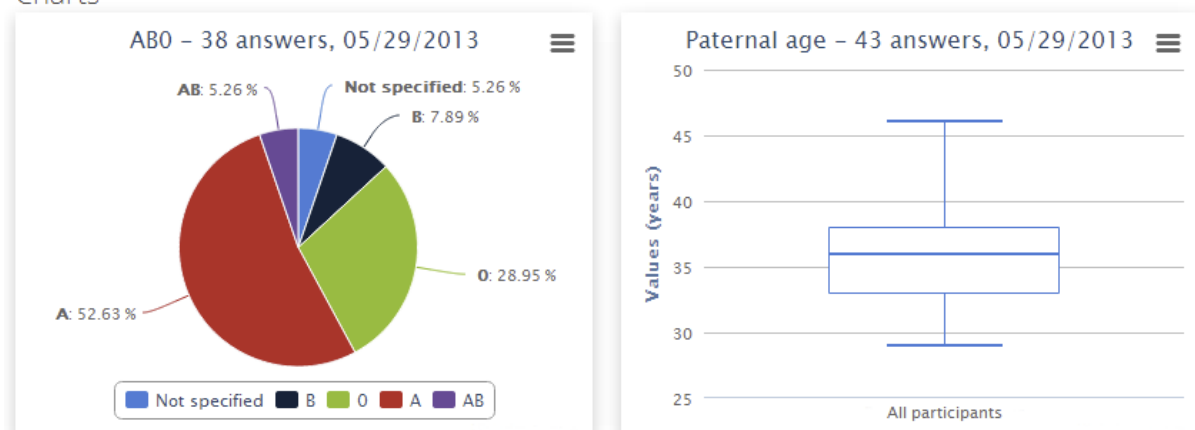
Unanswered questions (3/69)

- [Further information](#)
- [Specify disease of maternal stepsibling\(s\)](#)
- [Specify disease of paternal stepsibling\(s\)](#)

Constants (5/69)

- [Child with disease among maternal stepsiblings](#) (false)
- [Child with disease among paternal stepsiblings](#) (false)
- [Down syndrome in father's family](#) (false)
- [Hereditary diseases of father](#) (false)
- [Pregnancy-independent maternal diseases](#) (cardiovascular)

Charts



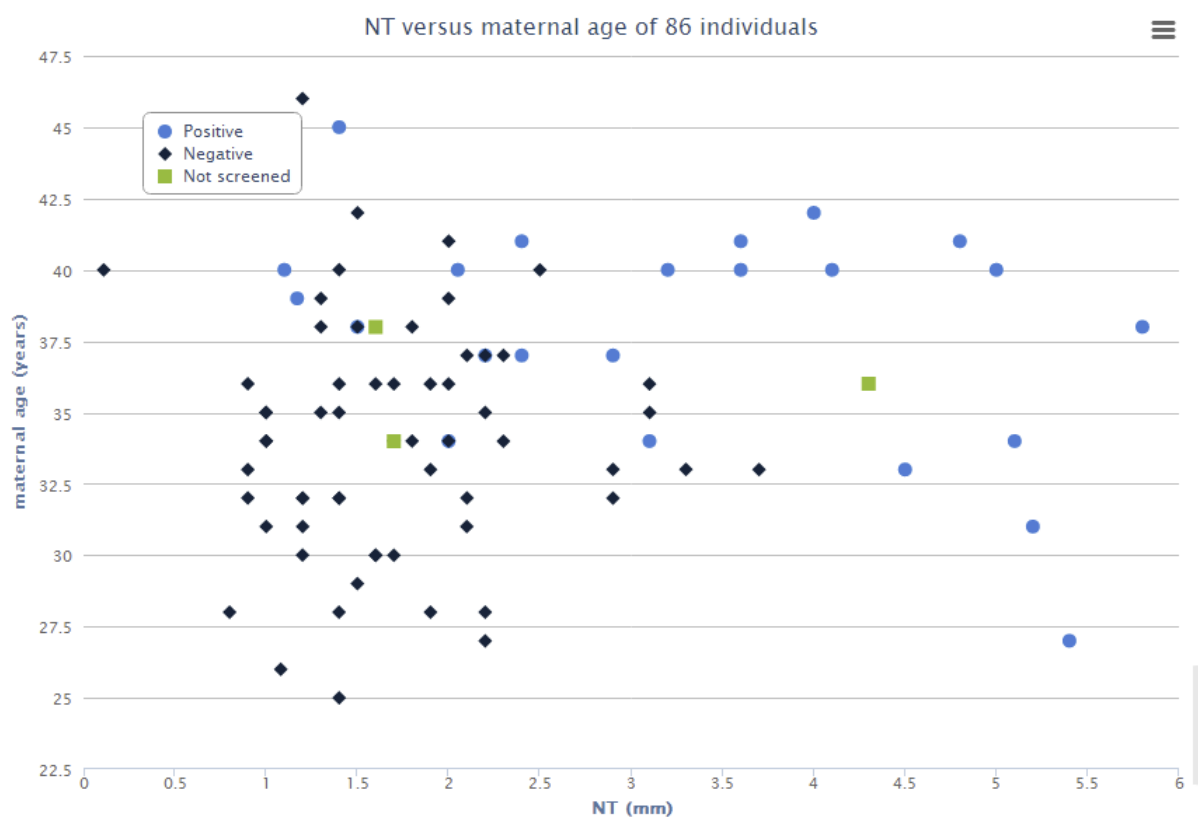
5.2 Charts

Intuitive charts allow users to visualize statistics and research products. Available charts include Box plots, and other standard graphs.

5.3 Visualizing relationships between data

Visualization functions allow users to recognize relationships between certain variables in their project or test your hypothesis. Variables of a profile or questionnaire can be visualized as a scatter chart in the Variable relationship page. Users can choose a profile or questionnaire and set the variables for the two axes. You can use drop-down type data when available to highlight groupings with different colors. If no drop-down type data is available, all dots will be displayed with the same color.

The chart below shows an example. This project was created to monitor the screening for Down syndrome. Axis X plots Nuchal Translucency that is used for ultrasound screening. Axis Y refers to maternal age. The dots are colored according to the screening result: Positive or Negative confirmed with chromosome test (further tests were missing in three cases). The chart verifies assumptions based on practical observation: higher maternal age and higher nuchal translucency value relate to a higher risk for Down syndrome.



5.4 Search & Query

On the Search & Query page, users can search for participants or biospecimens based on their data in related profiles or questionnaires. In this part of SmartBiobank™, the difference between questions and properties is not important and both are handled the same way, and hence we will refer to both as questions. This sample search is based on the answers for questions in biobank entries. The accompanying instruction manual includes details about available functions.

5.5 Data visualization

The Data visualization page offers statistical comparison of previously defined groups of biobank entries. First, users have to define at least two groups of biobank entries on the Search and Query page, see 7.1. To make further work easier, rename your searches to

know what kind of groups they contain. Then users save searches, go to the Data visualization page and select search groups from the list in the Groups to visualize box. Then select questions and properties you want to compare for these groups. See example below.

Data visualization

Data visualization name Untitled data visualization 03/12/2014 1:30:59.375 PM [Rename](#)

Visualization for

Groups to visualize Please select at least one group to visualize properties. Groups can be created at Search & Query page. Save a query and it will available as a group here!

Properties to visualize Please select at least one property to visualize it!

Chapter 6 Data exchange

Data exchange allows users to save work by enabling them to upload questions and data from a table to SmartBiobank™ – including uploading and mapping genetic data to participants. They also include protocols for saving, uploading and downloading clinical questionnaires, descriptions of Participants & Biospecimens, medical reports, participant data, and genomic and molecular data. Each type of data has a custom tool, and logging product. See example below, and the accompanying instruction manual for details about available functions.

Import Report

Import overview

	All		Imported		Error	
	Profile	Questionnaires	Profile	Questionnaires	Profile	Questionnaires
Patients	30	30	19	18	11	12
Samples	0	0	0	0	0	0
Summary	30	30	19	18	11	12

Started at	11/14/2013 9:11 AM
Started by	orsolya.ludmany@astridbio.com
Status	Finished successfully

Revoke data import

When you revoke your import **ALL** Patient, Sample, Filled questionnaires from your project that created within this import **will be deleted!** Be careful, once you revoke your import, there is no going back, all related data will be deleted!

By security reason, only the user created the import or the project owner can revoke an import!

[Revoke import](#)

Messages

[Download messages](#)

[▼ Show all](#)

- Import process started at 11/14/2013 9:11:22.868 AM
- Patient AB01 from sheet Main profile imported successfully!
- Patient AB02 from sheet Main profile imported successfully!
- Patient AB03 from sheet Main profile imported successfully!
- Patient AB04 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Anemia.
- Patient AB05 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Asthma.
- Patient AB06 from sheet Main profile imported successfully!
- Patient AB07 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Diabetes | High blood pressure | Obesity | Seizures.
- Patient AB08 from sheet Main profile cannot be imported: Invalid boolean value to question "Currently smoking?": null.
- Patient AB09 from sheet Main profile imported successfully!
- Patient AB10 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Obesity.

Chapter 7 SmartBiobank instruction manual

In order to instruct and streamline the use of the PrECISE SmartBiobank software, we designed an instructional manual, which is included in an appendix.

Chapter 8 Summary and Conclusion

ABT has designed a customized SmartBiobank software for PrECISE. This software allows to secure data repository, storage, access, download, analyses, and visualization. We are confident that this software will serve PrECISE investigators and facilitate data and analysis sharing within the project. This software will also help disseminate PrECISE products to the general projects following the successful completion of the project.

Together with this short report, we provide the official user manual for the SmartBiobank.

SmartBiobank™.com

User Manual

Version 2.0Astridbio Technologies Inc

Valid for SmartBiobank™ v1.3.1-greta build 37FB7



Table of contents

1	Introduction.....	4
2	Getting started	5
2.1	Registration	5
2.2	Login	6
2.3	New project	7
2.4	Inviting other participants.....	8
3	Glossary	9
4	Settings.....	10
4.1	Dashboard	10
4.1.1	Tiles and functions.....	11
4.2	Project details.....	14
4.2.1	Code validation patterns	14
4.2.2	Profile acceptance and completion – setting threshold values	15
4.2.3	Project users	16
4.3	My Profile	17
4.4	Appearance	18
5	Project design	20
5.1	Create profiles	20
5.1.1	Create properties	26
5.2	Create questionnaires	31
5.2.1	Create questions.....	32
6	Data input = Biobanking	37
6.1	Adding participant and biospecimen data	37
6.1.1	Submitting questionnaires	38
6.1.2	Tables of questionnaire data.....	39
6.2	Biospecimen management.....	39
6.2.1	Add refrigerator.....	39
6.2.2	Add biospecimen box	40
6.2.3	Biospecimen tracking	41
6.3	Listings.....	44
6.3.1	Listing biobank entries.....	45
6.3.2	Listing questions in Questionnaires and properties in Profiles.....	45

6.3.3	Listing questions on the Questions page.....	46
6.3.4	Listing questionnaires.....	47
6.3.4	Listing biospecimen organizers	47
6.4	Tracking user activity.....	48
6.4.1	Event log.....	48
6.4.2	Change log	49
7	Research center.....	50
7.1	Descriptive statistics.....	50
7.1.1	Adequacy of questions	51
7.1.2	Charts.....	51
7.2	Variable relationship	54
7.3	Search & Query.....	55
7.4	Data visualization	60
8	Data exchange	62
8.1	Import questionnaires.....	62
8.2	Copy questionnaires.....	64
8.3	Export questionnaires	65
8.4	Import Participants & Biospecimens	66
8.4.1	Previous data imports	68
8.5	Export data	69
8.6	Import genomic data.....	70

1 Introduction

Thank you for choosing SmartBiobank™, the online research data management software designed to help clinicians, lab biologists and researchers to integrate translational research elements into one single system. This introductory chapter includes general information about how to use this help system and how to get additional assistance if needed.

Intended Use

SmartBiobank™ is intended for life science research only. It is not for use in diagnostic procedures.

How to use this User Manual

This guide follows the logic of building a research project, starting from registration and ending with detailed data analysis. We recommend that you read the whole guide before building your first project.

We have also made it easy to find any particular topic or information that you are looking for. The table of contents is hyperlinked and will take you to the specified section by clicking on them. Clicking on a hyperlink will take you to the appropriate page with relevant information. Hyperlinks are also used to link various sections of the manual.

We designed this guide with plenty of screenshots. There are also three specific tools that we employed for easier understanding:

***Examples** are used in several cases where a first time reader might not understand correctly the applicability of a particular function.*

***Hints** give you ideas about a particular function to help you exploit the full potential of SmartBiobank™.*

***Notes** call your attention to features that might not be obvious but are important characteristics of the software.*

Additional Help

Technical Support

Contact technical support at info@astridbio.com .

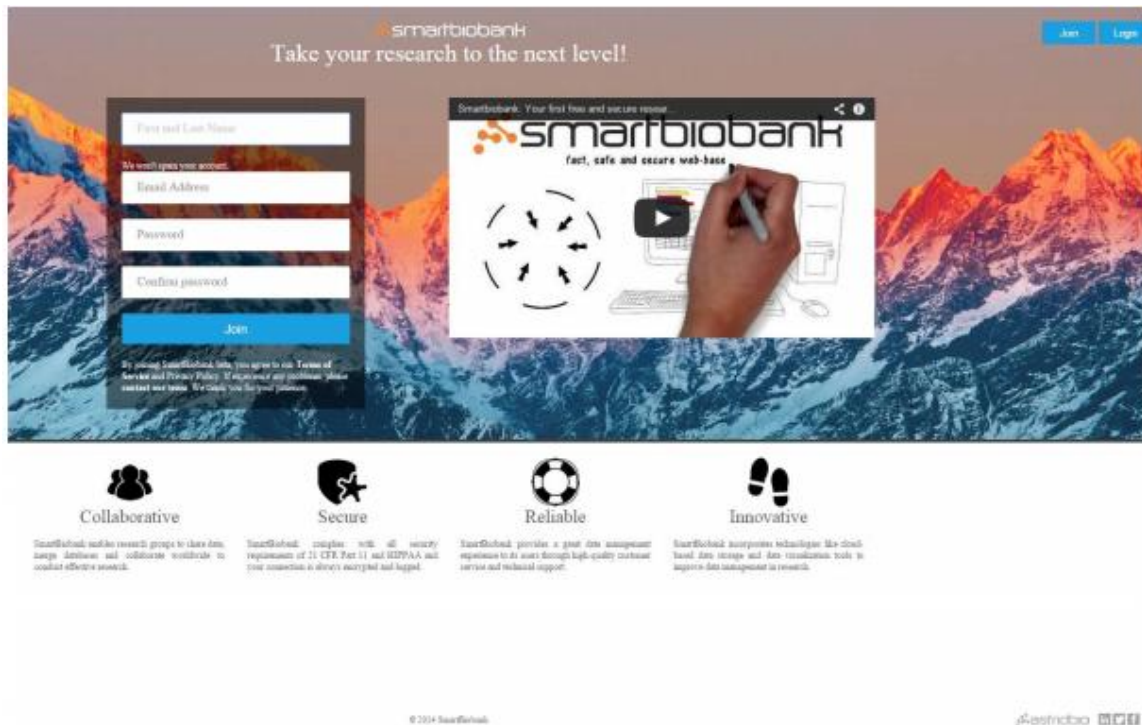
The company behind the software

Visit the AstridBio [website](#) for information about the company that developed SmartBiobank™.

2 Getting started

2.1 Registration

To start working on the PRECISE SmartBiobank™ project you need to register at <https://join.astridbio.com/>



You are required to fill in all fields. Instead of signing in on the homepage of SmartBiobank™ every time you want to log in, you can use your project domain name assigned by SmartBiobank™ where you and other project participants will be able to access your project:

<https://precise.astridbio.com/>

This domain name can be entered as a URL and you can access your project directly!

Passwords must be 6-10 characters long and must contain at least one numerical character. Please note that you will not be able to change your email address and project domain later.

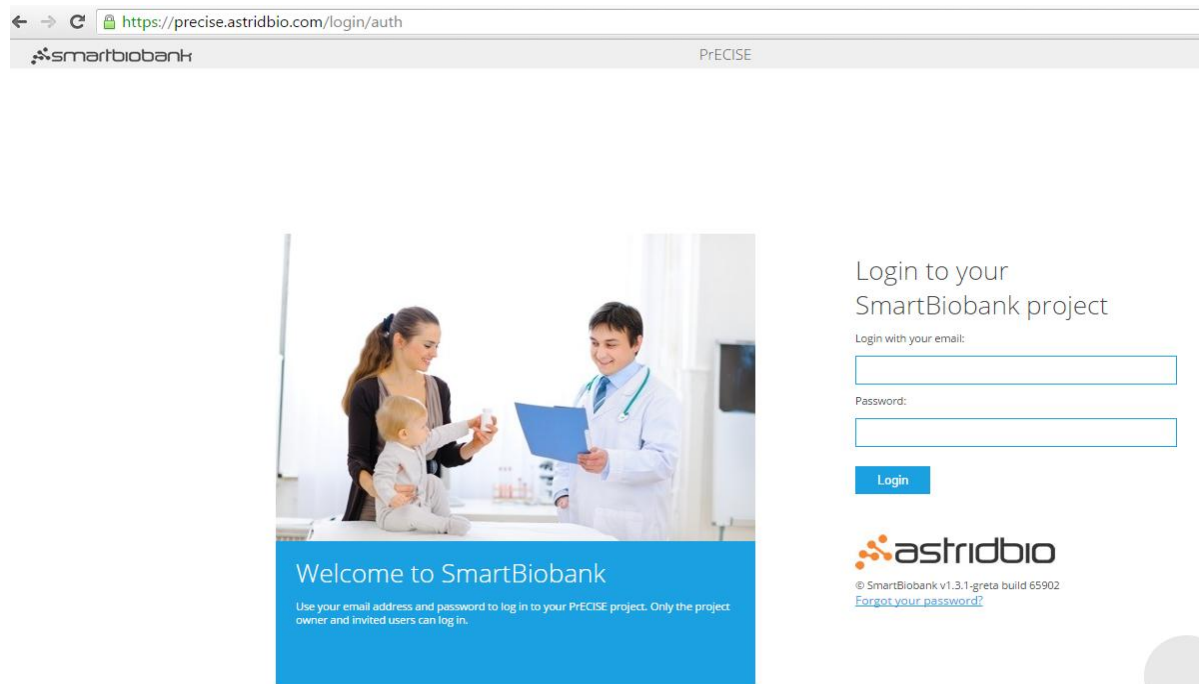
Before saving your data, please carefully read our **Terms of Service** and only proceed if you agree to them completely.

After completing your registration with SmartBiobank™, you will receive an activation email. Follow the instructions in the email to access your project from the project domain.

2.2 Login

After completing the registration steps, you can proceed with logging into SmartBiobank™. You can log in through the special domain name: <https://precise.astridbio.com/>

You can bookmark your Project URL and access it easily from there.



You can change account settings on the *My Account*, such as your password.

Note: The software starts a security countdown for each loaded page and initiates logout after 20 minutes of user inactivity. Users have 10 more minutes to refresh their session with a single mouse click, after which they need to reenter their password to continue.

2.3 New project

You can manage any number of projects with one registration. Follow these steps to start a new project:

- Log-in at the *My Account* page (see **2.2**)
- Click on the *Add project* option on the left side of the screen
- Choose a new *Project name* and domain and click *Save*

Add new SmartBiobank project

Choose a project name and a domain. [Details](#)

Project name	<input type="text"/>	?
Project domain	<input type="text"/> .smartbiobank.com	?
	<input type="button" value="Save"/>	<input type="button" value="Cancel"/>

Your new project will appear in the list, and you will receive an e-mail notification that will contain the link to your project.

2.4 Inviting other participants

Other participants in the project do not need to register to use SmartBiobank™, they can be invited through the *Project details* page using their email address.

Invite users to the project

To invite people to your project, please enter their email addresses (separate the addresses with a comma or space)

Email addresses	<input type="text" value="info@astridbio.com"/>
Role	<input type="text" value="Biobank user"/>
Invitation text	<input type="text"/>
	<ul style="list-style-type: none">Clinical userLaboratory userBiobank userAdministrator
	<input type="button" value="Invite"/> <input type="button" value="Cancel"/>

Following the instructions of the invitation email, new users can start helping in your project right after they register. They may also start a new project of their own.

SmartBiobank - Communi x
http://join.smartbiobank.com/signup/invite/?email=info@astridbio.com&project=2

smartbiobank
Take your research to the next level!

Join Login

SmartBiobank: Your first free and secure research

smartbiobank
fast, safe and secure web-base

Collaborative
SmartBiobank enables research groups to share data, merge databases and collaborate worldwide to conduct effective research.

Secure
SmartBiobank complies with all security requirements of 21 CFR Part 11 and HIPAA and your connection is always encrypted and logged.

Reliable
SmartBiobank provides a great data management experience to its users through high-quality customer service and technical support.

Innovative
SmartBiobank incorporates technologies like cloud-based data storage and data visualization tools to improve data management in research.

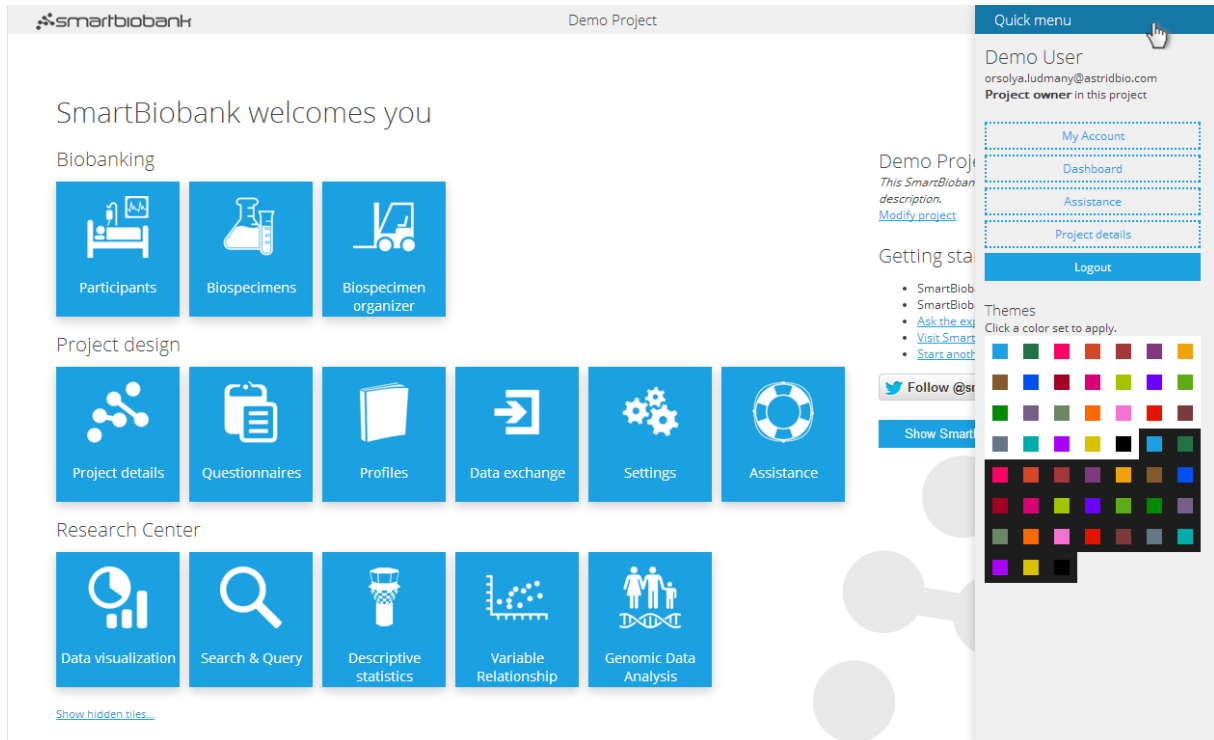
© 2014 SmartBiobank
astridbio

3 Glossary

Project	A project can be considered as one copy of the software. A project has its own domain name and settings. You can access all of your projects through the <i>My Account</i> page, but otherwise projects are not linked to each other.
Project name	A title that identifies your project. It is displayed on several pages of the project.
Project domain	The domain name where you can directly access your project.
Dashboard	The main page of your project. Navigate to all project functions from the Dashboard.
Tile	An icon that serves as a link to access a function.
Category	A Dashboard feature to group tiles. You can rearrange tiles according to your preference.
Biobank entry	An object that corresponds to either a participant (patient/individual who is used in your study) or a biospecimen (DNA, RNA or related biological samples). Participant entries and biospecimen entries are collectively referred to as biobank entries.
Profile	Profiles are lists of Properties that collect permanent data on biobank entries. Profiles are used to specify what data needs to be collected, they do not contain the individual data on biobank entries. Profiles can be thought of as blank forms that can be used repetitively to fill out specific information about biobank entries. Profiles are used to specify Properties of biobank entries that does not need to be changed or modified through the course of the project. You need to create profiles to add biobank entries to your project, so profile creation is the first step in building your project. Most projects require the co-existence of several different profiles.
Property	Properties are special permanent characteristics about participants and/or biospecimens that can be attached to profiles. Each property should have a unique name. For participants, it can be name, date of birth, gender, etc. For biospecimens, it can be volume, holder type, etc. Property names cannot be identical to question titles.
Questionnaire	Questionnaires contain the questions aimed to collect variable data, or data that is measured during the project. One questionnaire may be filled multiple times, the answers can be visualized and analyzed, while changes in answers can be followed. The same question can be listed in more than one questionnaire. Note that Questionnaires differ from Profiles in the sense that they list questions that contain unique and variable data that is vital for the success of the research project.
Question	Questionnaires are built from questions. The answers to the questions make up the database of SmartBiobank™. Apart from creating new questions, it is possible to reuse existing questions and attach them to more than one Questionnaire.

4 Settings

4.1 Dashboard



The main page, called Dashboard, contains tiles that lead to the functions. Tiles are organized into eight groups: *Biobanking*, *Project design*, *Research center*, *Settings*, *Data exchange*, *Biospecimen organizer*, *Favorites* and *Hidden tiles*.













You can drag and drop tiles to reorganize them according to your preferences. The *Settings*, *Data exchange*, *Biospecimen organizer* and *Favorites* categories are empty and invisible by default but they appear when you drag a tile over the page. Once the page is refreshed, the *Hidden tiles* category disappears.












Pull down the *Quick settings* menu on the upper right corner of the Dashboard to select a new color theme or to log out. There is also a *Logout* tile among the hidden tiles.

4.1.1 Tiles and functions

Below is a brief description of the tiles found on the Dashboard.

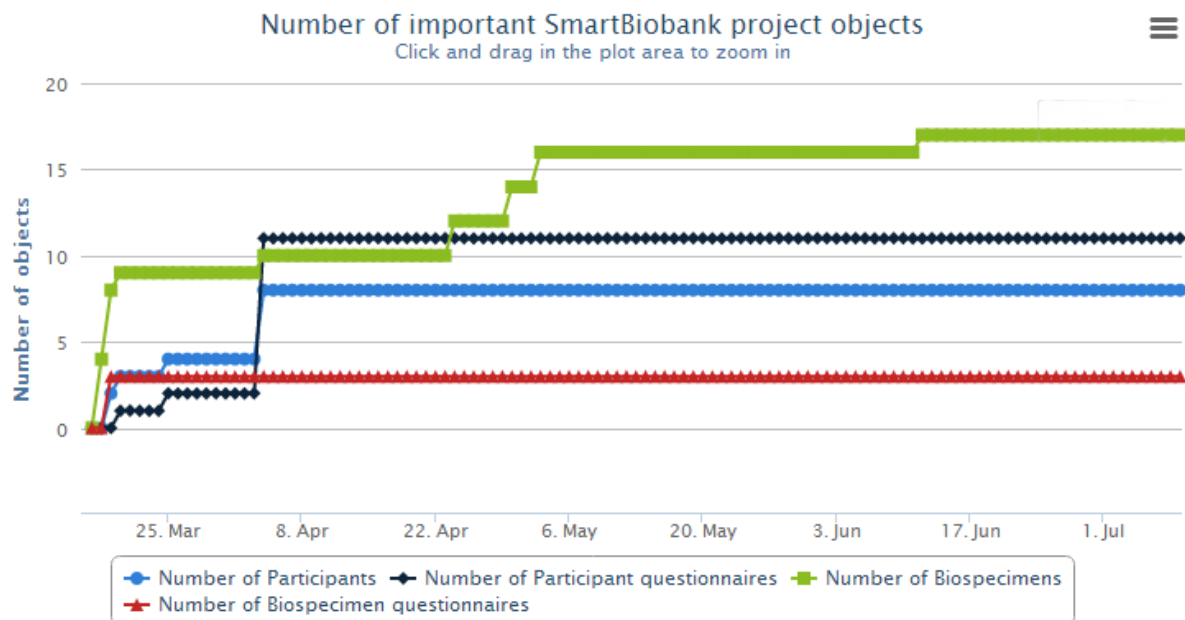
 <p>Profiles</p>	<p>Lists all participant and biospecimen profiles in your project. You can also create new profiles here. See 3. and 5.1.</p>	 <p>Questions</p>	<p>Lists all questions and characteristic properties in your project. You can also add new ones that need to be attached to a questionnaire or profile later. Read more about questions in section 5.2.1. The overview pages of certain types of questions and properties generate charts automatically. See 7.1.</p>
 <p>Questionnaires</p>	<p>Lists all questionnaires in your project. You can also add new questionnaires here. See g, 5.2.1 and 6.1.1.</p>	 <p>Add Questionnaire</p>	<p>Quick link to add a new questionnaire.</p>
 <p>Participants</p>	<p>Lists all participants in your project. You can access the Add participant function.</p>	 <p>Add Participant</p>	<p>Quick link to add a new participant. See 6.1.</p>
 <p>Biospecimens</p>	<p>Lists all biospecimens in your project. You can access the Add biospecimen function.</p>	 <p>Add Biospecimen</p>	<p>Quick link to add new biospecimen. See 6.1.</p>
 <p>Biospecimen organizer</p>	<p>The following functions are available under this tile (see them individually below): Refrigerators & boxes, Refrigerators, Biospecimen boxes. See 6.2.</p>	 <p>Refrigerators & boxes</p>	<p>It visualizes all refrigerators in your project with tiles. Clicking on a refrigerator tile displays biospecimen boxes stored inside the selected refrigerator. You can access the Add refrigerator and Add biospecimen box functions. This tile can also be found among the Hidden tiles on the Dashboard by default.</p>

 <p>Refrigerators</p>	<p>Lists all refrigerators in your project. You can access the Add refrigerator function. Open a refrigerator to access the list of stored boxes. This tile can also be found among the Hidden tiles on the Dashboard. See 6.2.1.</p>	 <p>Biospecimen boxes</p>	<p>Lists all biospecimen boxes in your project. You can access the Add biospecimen box function. Open a box to view its content. This tile can also be found among the Hidden tiles on the Dashboard by default. See 6.2.2.</p>
 <p>Settings</p>	<p>The following functions are available under this tile (see them individually below): Appearance, My profile, Project details.</p>	 <p>Appearance</p>	<p>You can change the appearance of the software (color, chart style etc). This tile can also be found among the Hidden tiles on the Dashboard by default. See 0.</p>
 <p>My profile</p>	<p>You can change your account information. This tile can also be found among the Hidden tiles on the Dashboard by default.</p>	 <p>Data visualization</p>	<p>Compares previously defined groups of biobank entries by custom criteria. See 0.</p>
 <p>Event log</p>	<p>Events are listed here. The following information is displayed: date, actor, location, action and event message. See 6.4.1.</p>	 <p>Project details</p>	<p>Displays some basic data of your project. You can modify some settings or invite new users if you have the required user roles.</p>
 <p>Descriptive statistics</p>	<p>Lists all charts based on answers given to one question. Also lists those questions that have effectiveness problems based on their answers. See 7.1.</p>	 <p>Variable Relationship</p>	<p>The program is able to generate scatter charts out of the selected numerical data. See 7.2.</p>
 <p>Search & Query</p>	<p>Search & Query: Search participants or biospecimens by their profile or questionnaire answers. You can add more criteria to your search session. See 0.</p>	 <p>Data exchange</p>	<p>The following functions are available under this tile: Import questionnaires, Export questionnaires, Copy questionnaires, Import data, Export data, Previous Data imports, Import VCF, Export genomic data. See 8.</p>

 <p>Import questionnaires</p>	<p>You can import empty questionnaires from an Excel spreadsheet. This tile can also be found among the Hidden tiles on the Dashboard by default. 8.1</p>	 <p>Copy questionnaires</p>	<p>You can copy empty questionnaires or profiles from other projects into your own. No research data will be copied! This tile can also be found among the Hidden tiles on the Dashboard by default. 8.2</p>
 <p>Export questionnaires</p>	<p>You can export empty questionnaires into an Excel spreadsheet. This tile can also be found among the Hidden tiles on the Dashboard by default. 0</p>	 <p>Import data</p>	<p>You can import participant and biospecimen profile data or questionnaire data into already existing profiles or questionnaires. See 8.4</p>
 <p>Export data</p>	<p>You can export participant and biospecimen profile data or questionnaire data into Excel files. See 8.5.</p>	 <p>Previous data imports</p>	<p>Lists previous data imports. Imports may be revoked. See 8.4.1.</p>
 <p>Import VCF</p>	<p>You can import genomic data from VCF files and link them to participants. See 8.</p>	 <p>Genomic Data Analysis</p>	<p>You can run a simple analysis based on the genomic data stored in your project. See Hiba! A hivatkozási forrás nem alálható.. It requires previously defined groups, see 0.</p>
 <p>Export genomic data</p>	<p>You can export your genomic data into MAP and PED files. See Hiba! A ivatkozási forrás nem található..</p>	 <p>Send feedback</p>	<p>You can send us feedbacks, suggestions or bug reports.</p>
 <p>Logout</p>	<p>Click to exit SmartBiobank™. This tile can be found among the Hidden tiles on the Dashboard by default.</p>		

4.2 Project details

Project details contain basic information about the project such as the numbers of participants and biospecimens (*biobank entries*). A chart at the bottom of the page shows how these numbers have changed over time.



At the beginning of a project, you need to enter some basic details about the project. Project owners and administrators can customize their project further by adding a description to the dashboard and a welcome message to the login page (see 2.2).

4.2.1 Code validation patterns

All participant and biospecimen entries have an identification code, which may be entered without restrictions by default. However, consistency or precision may require to apply a strict pattern for identification codes which may be set as a *regular expression*. The generic code is written as follows:

$$[character\ range]\{quantifier\}$$

Therefore if you need a code to contain exactly one alphabetical character (from A to Z) followed by three numerical characters (from 0 to 9), you need to enter the following regular expression:

$$[A - Z]\{1\}[0 - 9]\{3\}$$

Theoretically any pattern can be used in SmartBiobank™, but a detailed review about regular expressions is beyond the scope of this manual. (If you want to know more about regular expressions

please contact us directly.) If a validation pattern is set, an entry with a conflicting code cannot be submitted. You may even set a specific message for users who forgot to enter a correct code.

Do not set a code pattern if you use different types of codes for the same profile type. This is the case when parent biospecimen codes differ from aliquot biospecimen codes in character number, for instance.

4.2.2 Profile acceptance and completion – setting threshold values

When project owners and administrators create questionnaires, they can set how many times they require that questionnaire to be submitted. Profile acceptance and completion threshold values serve as a progress indicator, allowing you to know at a glance if enough questionnaires have been submitted to start data evaluation, or if some entries lack submissions.

Every participant and biospecimen entry contains a status value (*profile status*) that represents the completeness of the assigned questionnaires. In addition, *questionnaire status* is indicated inside the entry, under the assigned questionnaires.

Note: Status does show how many questions have been answered, but how many times questionnaires have been submitted – with or without answers.

Completed profiles mean that all required questionnaires have been submitted and further input is not required. Accepted profiles on the other hand lack some submissions, but submitted data is sufficient for the project. Default values are 95% and 80%, respectively.

Example: A questionnaire called Body Temperature is required to be submitted at least 10 times. We use the default settings and start to measure the temperature of participants every day, submitting a questionnaire after every measurement. Profile status is “incomplete” until the 8th day when it changes to “accepted”. Continuing this process profile status changes to “completed” after the 10th measurement, since the threshold was 95%. We can choose to continue our measurements and inputs, but it does not affect the status.

If more than one questionnaire is assigned to a profile, profile status is calculated based on all questionnaire status.

4.2.3 Project users

Project owners and administrators have an access to a subpage of *Project details* called *Project users* where they can manage who can have an access to the project. There are five different roles defined in the system.

Project owners are the ones who have registered a project. They are the only users who require full registration. They have access to all functions and they can invite new participants or change their roles. Only owners can remove the project. Project owners cannot be deleted from the project.

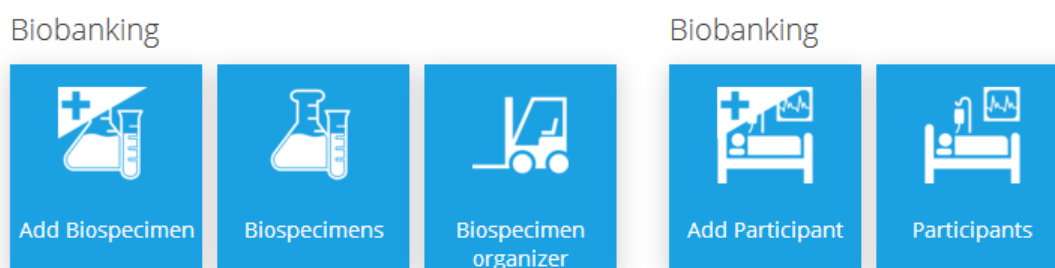
Administrators are invited users with full access to the SmartBiobank™ functions, similarly to the project owner. They can help project owners by creating profiles and questionnaires, managing participants etc.

Biobank users have limited access, their activity is restricted to data input or edit. Additionally they can view some informational pages like *Project details*. This allows them to check some settings and assist the project according to these settings. They cannot import or export data and they cannot modify questionnaires or questions. If you are working with sensitive data, you may want to further restrict user access (see the last two user groups below).

Clinical users are biobank users only permitted to work with participant data.

Laboratory users are biobank users only permitted to work with biospecimen data.

The picture below shows the difference between the accessible tiles and functions of the laboratory (left) and clinical user (right).



4.3 My Profile

You can manage your account settings at the *My profile* page, which is located inside *Settings* or among the hidden tiles on the *Dashboard*. You can also access it from the *My Account* page. Every account detail can be modified except for the email address since it also functions as a login name. After changing your password you will have to use the new one in all of your projects.

The page lists all projects you are registered to, along with your user role(s).

User profile

Email (cannot be changed) ?

First name

Last name

Owner profile

Organization

Department

Address

Country ▼

City

Postal code

Phone

Change password

New password ?

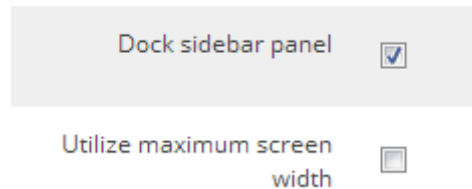
Re-enter new password ?

4.4 Appearance

Users may customize the visual appearance of the software on the *Appearance* page.

Sidebar and panel size

Sidebar and screen width

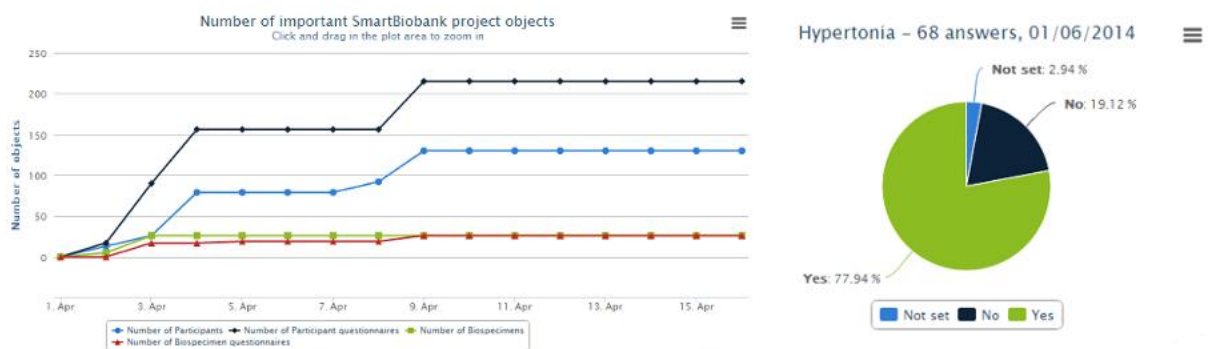


The docked sidebar is always visible on the left side of the page. If you uncheck the button, the sidebar will slide out of the view when not used. If you enable maximum content width, lists will occupy the whole width of your screen, irrespectively of the size of your monitor.

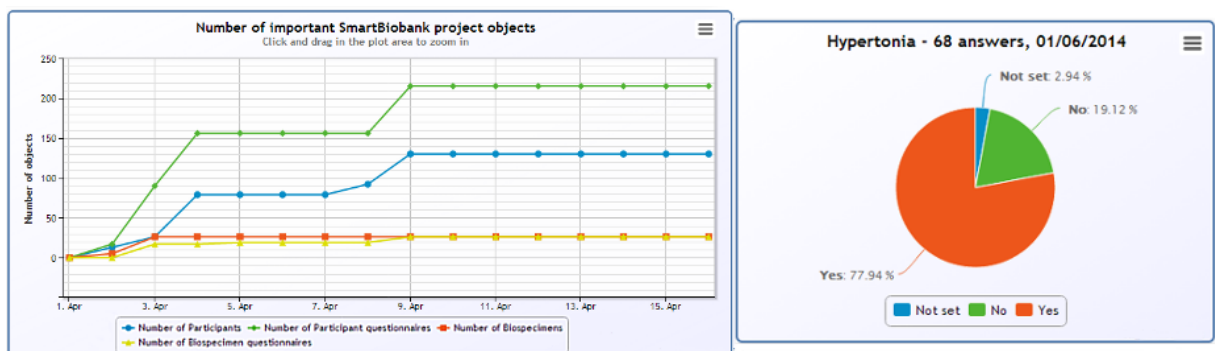
Chart styles

Whether for visual or practical purposes, you may want to change the appearance of charts. You can choose from the following styles:

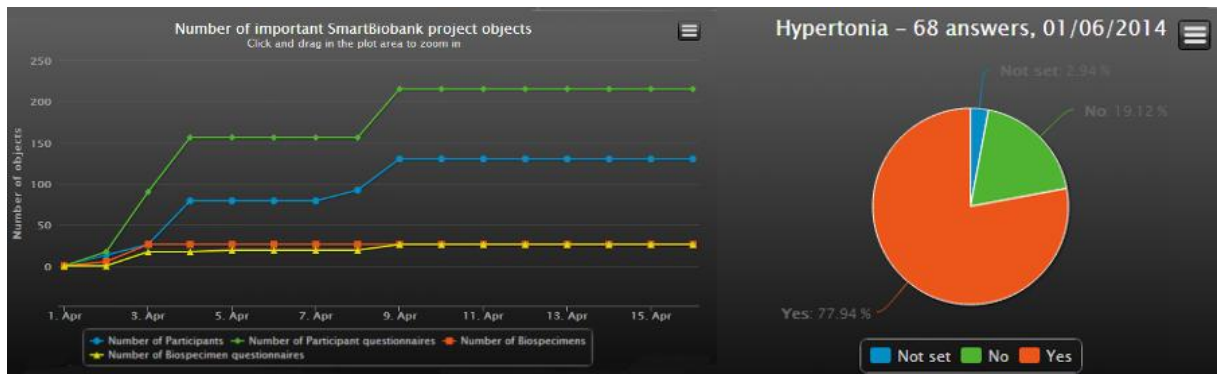
- *Light style*



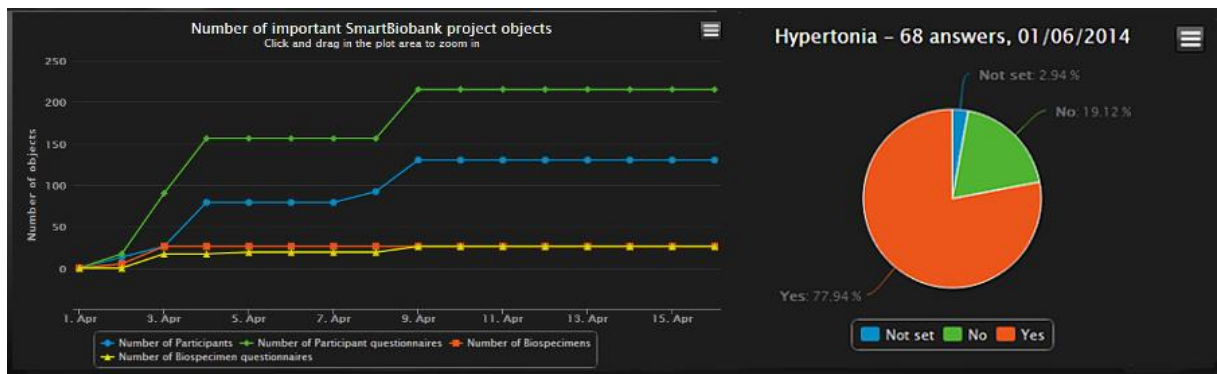
- *Grid style*



- *Dark style*



- *Dark (no gradient)*



Continuous data visualization

Continuous data can be plotted as a line or area chart. Only the timeline (located on the *Project details* page) contains continuous data.

Themes

You can choose from 52 different color themes. Just click on the color of your choice and enjoy the new looks in all of your projects.

5 Project design

5.1 Create profiles

Profile creation is the first step before adding biobank entries to your project. Profiles are lists of Properties that collect permanent data on biobank entries. Profiles can be thought of as blank forms that can be used repetitively to fill out specific information about biobank entries. Several profiles can be added to a project, and different profiles may contain different properties.

Example: You can create a profile for participants and another for the control group. You also need to add different profiles if you want to register participants with different diseases, since you will want to input different data. This is also the case with biospecimens – blood sample and urine sample will need to be registered under different profiles.

To create a new profile click on *Profiles* on the *Dashboard*. If you use this function for the first time, an *assistance* page will appear that will help you get started. You can start creating a new profile by clicking on the sidebar (*Add new participant profile* or *Add new biospecimen profile*) or by following instructions from the *assistance* page.

The screenshot displays the SmartBiobank interface for a new project. The top navigation bar includes the SmartBiobank logo, the project name 'demo project', and the user name 'Hello Demo User!'. The left sidebar is blue and contains a 'Profile management' section with two options: 'Add participant profile' and 'Add biospecimen profile'. Below this is a 'Dashboard' section with a note: 'All profiles are listed here. Click on a profile to add or attach properties. Details'. The main content area is titled 'Getting started with your project' and includes a welcome message: 'Thank you for creating a new SmartBiobank project. If you're familiar with SmartBiobank, you can start using your project in the old fashioned way. If you are unfamiliar with SmartBiobank this page will help you to start your project.' The 'Where should I start?' section offers two paths: 'Design your project' (with sub-steps: 1. Define at least one participant or biospecimen profile, 2. Define at least one questionnaire) and 'Collect your data' (with sub-steps: 1. Create at least one participant or biospecimen, 2. Fill questionnaires). The 'Next step' section provides three buttons: 'Add participant profile', 'Add biospecimen profile', and 'Import profiles'. The page also features a lifebuoy icon and a network diagram.

You have to choose a unique name for your new profile that is different from all other profile and questionnaire names. You may set a description for your new profile that appears in the profile overview page. Since SmartBiobank™ supports medical research with the ICD-10 coding system, you may select a code for participant profiles.

Hint: Start typing either the code or the name of the disease to select the code you want. If you wish to search by a part of the disease name, start your search with an asterisk (*) then type the word fragment you want.

Examples of different biospecimen profiles are different types of biological specimens like blood, urine, RNA, protein, etc. For biospecimen profiles, you also need to set an optimal storage temperature.

While you can add properties any time, it is advisable to add them during profile creation to ensure accuracy during entry.

Note: Properties contain data that define the given biobank entry – either participant or biospecimen and are attached to Profiles. Add properties that only refer to constant information that is relevant to the participant group or biospecimen type, and differentiate one entry from another by using unique titles.

Below is an example of a saved biospecimen profile titled *Sample information*. This is the Overview page that lists the profile data such as date of creation, and the list of attached properties.

Sample information

Profile overview. You can attach properties here. [Details](#)

Overview

Profile name	Sample information
Description	
Applied to	Biospecimens
Created by	orsolya.ludmany@astridbio.com
Created	03/13/2014 1:43 PM
Last modified	03/13/2014 1:43 PM

Sample informations properties

Biospecimen ID	Identifier	
Date of sampling	Date	
Gestational age	Real number	
Sample volume	Real number	
Date of centrifuging	Date	
Date of refrigeration	Date	

[Click to add property](#)

Next, is an example of a sample screenshot that shows how a new biospecimen is added to the appropriate profile. Notice how the properties from the previous screenshot appear.

Add new biospecimen

Enter biospecimen ID and select profile to fill questionnaire. [Details](#)

Biospecimen ID

Biospecimen profile

Date of sampling (MM/dd/yyyy)

Gestational age weeks

Sample volume ml

Date of centrifuging (MM/dd/yyyy)

Date of refrigeration (MM/dd/yyyy)

Below is how the *Overview* page looks like after saving the form. Notice that the list contains the profile title (Sample information in this case).

Biospecimen #1045

completed (100%)

Biospecimen ID	1045
Biospecimen profile	Sample information
Date of sampling	06/22/2011
Gestational age	26.0 weeks
Sample volume	12.0 ml
Date of centrifuging	06/22/2011
Date of refrigeration	06/22/2011
Patient (copied)	1045
Biospecimen stored	not stored
Created by	orsolya.ludmany@astridbio.com
Created	03/13/2014 1:55 PM
Last updated	03/13/2014 1:55 PM

Questionnaires

DNA concentration

Fill DNA concentration

03/13/2014: FILL79330

You have several options for adding a property to the profile. You can either add predefined properties that are provided by SmartBiobank™ or create new properties of your own. If there are previously created properties or questions in your project, then select *Properties in this project* option. You can then reattach those properties to more profiles or questionnaires as convenient.

Note: *Even though questions are normally used in questionnaires, you can attach them to profiles; also, you can attach previously created properties to questionnaires (see 5.2). A good example for a universal question may be “Further information”.*

You can create new properties by clicking on the *Create New property* option. Read more about property creation in section **5.1.1**. Make sure to click *Save* after finishing the process.

Patient data

Profile overview. You can attach properties here. [Details](#)

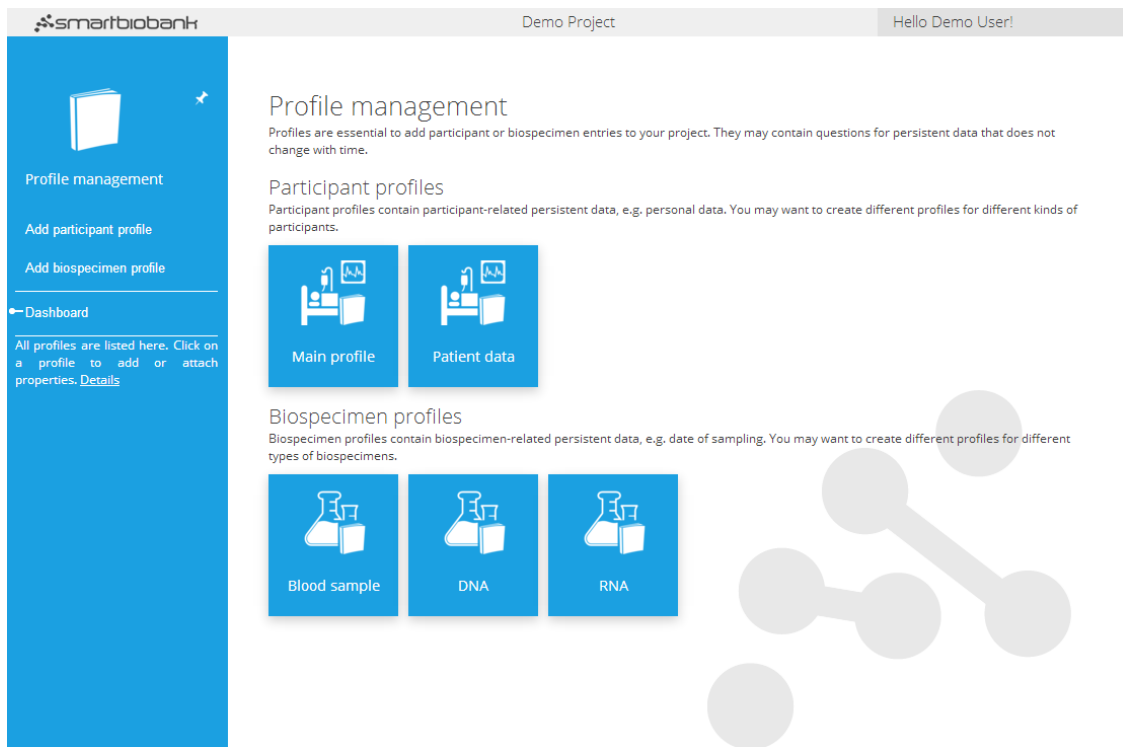
Overview

Profile name	Patient data
Description	
Applied to	Participants
Created by	orsolya.ludmany@astridbio.com
Created	02/05/2014 1:02 PM
Last modified	02/21/2014 2:00 PM

Patient datas properties

Participant ID	Identifier	!
Gender	Drop Down select	> !
Date of birth	Date	> !
Smoking	Boolean	>
Physical activity level	Drop Down select	>
Family medical history	Drop Down multiple select	>
Job role	Drop Down select	>
Work hours a week	Integer	>
Activity level	Drop Down select	>
Click to add property		

If you have already created profiles, the corresponding tiles will appear on the Profiles page. You can click on the sidebar buttons to create a new profile. Please see screenshot below:



Note: If a question of property have been attached to multiple profiles or questionnaires, their chart appearing on the overview page will contain data from all sources. (See 7.1.2) If you want to see the profile- or questionnaire-specific charts, open the property/question overview page through the profile or questionnaire.

You can read more about charts in section 7.1.

Note: profiles look just like questionnaires; the difference is not in the appearance rather in its function.


You can check what biobank entries have been submitted using a profile. While on a Profile overview page, click on Show answers on the sidebar to view all the entries with the first five answers.








5.1.1 Create properties

Properties are the essential elements of a profile, as they contain all information that define both the biobank group (participant group or biospecimen type) and each biobank entry. For participants, it can be name, data of birth, gender, etc. For biospecimens, it can be volume, holder type, etc. Property names cannot be identical to question titles.



Create a new property

Choose a unique name for your profile and add a description if needed. You can attach properties before or after saving your profile.

 After this property is created, it will be added to the Patient data profile.

Question text	<input type="text"/>	
Display name	<input type="text"/>	
Description	<input type="text"/>	
Hint	<input type="text"/>	
Mandatory	<input type="checkbox"/> This question is mandatory.	
Important	<input type="checkbox"/> This question is important.	
Type	<input type="text" value="Short text"/>	

Properties related to the type

Minimum length	<input type="text"/>	
Maximum length	<input type="text"/>	

Property text: It appears in profiles on the left side of the answer box. Please note that property texts and question texts cannot be identical. If you need to have the same property in more profiles, or need a question (created for a questionnaire) for a profile, you need to attach them through the *Properties in this project* option. For example, if you are using the same patient data for more than one profile, instead of creating new “Name” questions for every participant profile, you can attach the “Names” via the property text option.

Patient data

Profile overview. You can attach properties here. [Details](#)

Overview

Profile name	Patient data
Description	
Applied to	Participants
Created by	orsolya.ludmany@astridbio.com
Created	03/17/2014 3:13 PM
Last modified	03/17/2014 4:19 PM

Patient datas properties

 Participant ID	Identifier	!
 Name	Short text	>
 Gender	Drop Down select	> !
 Date of birth	Date	> !
Click to add property		

Display name: Usually a short version of the Property text - it appears in tables, charts and the overview page of the property. If you leave *Display name* blank, the software will use *Property text* as a default display name.

Description: Extra information related to the property. This piece of information only appears on the property overview page.

Hint: Sometimes further guidance is necessary to help users give correct information. Hints appear next to the answer box as a small question mark. Move your mouse over the symbol to read the hint.

Mandatory: There may be information that must be entered for the project to be effective. You can mark these by checking the *Mandatory* question box. Profiles cannot be submitted until all mandatory data is entered.

Hint: *If you want to attach a new mandatory property to an existing profile, it will be implemented in new entries. This means, however, that you will have to provide the necessary data for the new property when you edit older entries.*

Important: You can indicate whether the property is important or not for the project.

Type: You can enter nine different types of data into your profiles. Some types allow data visualization that may be important for your project. You may also set some restrictions that differ for each data type. Type of data helps you identify what values (characters, numbers, etc) are acceptable for a property.

- Short text: This is the default type. Use it to enter short information (e.g. names). Answers cannot be visualized. Allows you to set a minimum and/or maximum length for the answer.
- Long text: Similar to the Short text type but more characters are visible. Answers cannot be visualized. You may set a minimum and/or maximum length for the answer.
- Integer: Only numerical values are accepted. Values are visualized on the property overview pages with box plot, and they can be used for scatter chart visualization on the *Variable relationship* page. You can enter a unit of measure and a minimum and maximum range that will help avoid entering values in an inappropriate unit of measure.

Note: SmartBiobank™ does not offer unit conversion.

The allowed value range must also include pathological values in research data. Integers are not a good choice to enter codes.

Crown rump length (CRL) mm 

Hint: Integer type is not appropriate for some data that are normally considered numbers. For example: we want to know how many children the participant has. In this case additional statistics provided by the box plot, such as median, are not relevant information. Use drop-down type question to have pie charts (showing percentages) based on the entered data. See drop-down select for further information.

- Real number: Only numerical values are accepted. This type is intended for entering decimal fractions with a dot as the decimal mark. Values are visualized on the property overview pages with box plot, and they can be used for scatter chart visualization on the *Variable relationship* page. You can enter a unit of measure and a minimum and maximum range that will help avoid entering values in an inappropriate unit of measure.

The allowed value range must also include pathological values. If large integers are entered to a real number type question, it is automatically converted to normalized notations.

Sample volume ml

- Date: Dates follow the MM/dd/yyyy format and users are also required to enter the slashes manually. Dates cannot be visualized. You can set the earliest and latest acceptable dates.

Date of sampling (MM/dd/yyyy)

Note: Invalid dates are automatically converted. For example a date of 02/30/2013 would appear as 03/02/2013 after saving the profile.

- Boolean: Yes/no or Unknown/yes/no options can be selected from a drop-down menu. Answers are visualized in pie charts, but will not appear on the scatter chart (see 7.1.2.)

Previous pregnancy with Rh+ child

No ▲

No

Yes

- Drop-down select: A single-select or multi-select feature, with all options entered manually during property creation.

Hint: Consider setting the first option to “Not known” or “Not specified” since there may be cases when none of the options apply or there is not enough information to give a proper answer.

Answers are visualized in pie charts and they can highlight groupings on scatter charts (see 7.1.2.)

ABO

A ▲

A

B

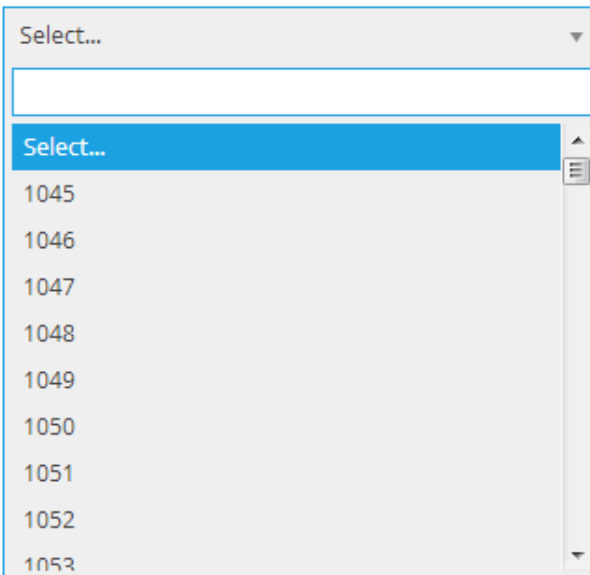
AB

O

Not specified

- Participant link: This type serves to link participants to other biobank entries. Link a biospecimen to a participant by adding a Participant link type property to the biospecimen profile, or alternatively to a biospecimen questionnaire. During the process of adding a new biospecimen, a drop-down menu appears in the profile or questionnaire with all the participants in your project. Choose a participant to link it to the selected biospecimen. It is also possible to link different participants, e.g. to be able to record relationships between participants.

Participant



Select...

Select...

1045

1046

1047

1048

1049

1050

1051

1052

1053

- Biospecimen link: Similar to Participant link, it may be used to link already added biospecimens to other biobank entries. It can be used to connect related biospecimens.
- Genetic data: This type can hold genetic data imported from VCF files. Since upload is not done manually, adding this type of property to a profile is not recommended. Consider adding a questionnaire with only a genetic data type question. See **8.6** section for more details.

Hint: You can copy profiles from another project. Read about these options in section 8.

The software analyses some basic attributes of the answers and warns you if it detects ineffective properties. See **7.1.1**.

Viewing the property overview page, you can check what profile the property has been attached to. Click on *Show attached questionnaire* on the sidebar to view all the attached profiles and questionnaires.

5.2 Create questionnaires

Questionnaires contain the questions aimed to collect variable data, or data that is measured during the project. One questionnaire may be filled multiple times, the answers can be visualized and analyzed, while changes in answers can be followed. The same question can be listed in more than one questionnaire. Note that Questionnaires differ from Profiles in the sense that they list questions that contain unique and variable data that is vital for the success of the research project.

Questionnaires serve as essential tools to monitor changes in variables. There is no limit how many times a participant or biospecimen questionnaire can be filled, although administrators and project owners can indicate their requirements (number of expected questionnaires filled). First, navigate from the *Dashboard* into the *Questionnaires* page or click on the *Add questionnaire* tile (hidden by default). After clicking on *Add questionnaire* on the sidebar, you have to choose a name that is different from any other profile and questionnaire name in the project. You may also set a description for the questionnaire. From the drop-down menu select how many times you expect this questionnaire to be filled per biobank entry then link the questionnaire to the biobank entry (either participants or biospecimens).

Questionnaire creation is very similar to profile creation. Click on the *Add question* option to create a new question, choose a predefined question or browse previously created ones.

Hint: *The difference between properties and questions is only administrative, not technical, therefore both are listed on the Questions page, and they are visualized similarly.*

You can read more about question creation in section **5.2.1**. There is no explicit limit on how many questions a questionnaire can contain, but too many questions within the same questionnaire may decrease manageability.

Hint: *When a questionnaire is submitted about a participant or a biospecimen, it is recorded with the submission date. If examination dates* in your project are usually not the same as submission dates, it is advisable to add a question to the questionnaire asking for the examination date. This would be necessary in case of a high amount of archive data is entered on the same day.*

**Examination dates refer to dates when patients or biospecimen samples were examined and the examination data was recorded.*

You can check what biobank entries have submitted answers to a questionnaire. Click on *Show answers* on the sidebar to view all the entries with the first five answers.

5.2.1 Create questions

Probably the most important elements in SmartBiobank™ are questions that make data input possible. Well-planned questions help your project in many ways, such as keeping it organized and making preparing statistics easy.

The process of question creation is the same whether you start from the *Questions* page or from within the Profile or Questionnaire. If starting from the *Questions* page, you have to attach the created question either in the profile or the questionnaire afterwards.

Hint: The difference between properties and questions is only administrative, not technical, therefore both are listed on the Questions page, and they are visualized similarly.

The elements of the question creation block are the same as the property creation module, only some of the element names are different. The list below is nearly identical to the Create properties chapter, feel free to skip it if you are already familiar with them.

Question text: It appears in questionnaires on the left side of the answer box. Please note that one question text may be set once per project. However, the attaching question function eliminates the need for creating the same questions for all of the profiles or questionnaires. For example, instead of creating new “Name” questions for every participant profile, you only need to attach them.

Display name: Usually a short version of the Question text - it appears in tables, charts and the overview page of the question. If you leave Property name blank, the software will use Question text as a property name.

Question description: Extra information related to the question. This piece of information only appears on the question overview page.

Hint: Sometimes further guidance is necessary to help users give correct answers to a question. Hints appear next to the answer box as a small question mark. Move your mouse over the symbol to read the hint.

Mandatory: There may be questions that must be answered for the project to be effective. You can mark these by checking the Mandatory question box. Questionnaires and profiles cannot be submitted until all mandatory questions are answered.

Hint: If you want to attach a new mandatory question to an existing questionnaire, it will affect new entries. This means however that you will have to provide answer for the new question when you edit older entries.

You cannot check the Mandatory question box later during question editing if questionnaires have already been submitted with a missing answer.

Important: You can indicate whether the question is important for the project.

Type: There are currently nine different types of questions in SmartBiobank™ that allows you to create highly customized questionnaires. Your objectives may require specific question types as only a few of them provide data that can be visualized with charts. You may also set some restrictions that differ for each question type.

- Short text: This is the default question type. Use it to enter short information (e.g. names). Answers cannot be visualized. Allows you to set a minimum and/or maximum length for the answer.
- Long text: Similar to the Short text type but more characters are visible. Answers cannot be visualized. You may set a minimum and/or maximum length for the answer.
- Integer: Only numerical values are accepted. Values are visualized on the question overview pages with box plot, and they can be used for scatter chart visualization on the *Variable relationship* page. You can enter a unit of measure and a minimum and maximum range that will help avoid entering values in an inappropriate unit of measure.

Note: SmartBiobank™ does not offer unit conversion.

The allowed value range must also include pathological values. Integers are not a good choice to enter codes.

Crown rump length (CRL) mm 

Hint: Integer type is not appropriate for some data that are normally considered numbers. For example: we want to know how many children the participant has. In this case additional statistics provided by the box plot, such as median, are not relevant information. Use drop-down type question to have pie charts (showing percentages) based on the entered data. See drop-down select for further information.

- Real number: Only numerical values are accepted, with a dot as the decimal mark. Values are visualized on the question overview pages with box plot, and they can be used for scatter

chart visualization on the *Variable relationship* page. You can enter a unit of measure and a minimum and maximum range that will help avoid entering values in an inappropriate unit of measure.

Note: SmartBiobank™ does not offer unit conversion.

The allowed value range must also include pathological values.

Sample volume ml

- Date: Dates follow the MM/dd/yyyy format and users are also required to enter the slashes manually. Dates cannot be visualized. You can set the earliest and latest acceptable dates.

Date of sampling (MM/dd/yyyy)

Note: Invalid dates are automatically converted. For example a date of 02/30/2013 would appear as 03/02/2013 after saving the questionnaire.

- Boolean: Yes/no or Unknown/yes/no options can be selected from a drop-down menu. Answers are visualized in pie charts, but will not appear on the scatter chart (see 7.1.2.)

Previous pregnancy with Rh+ child

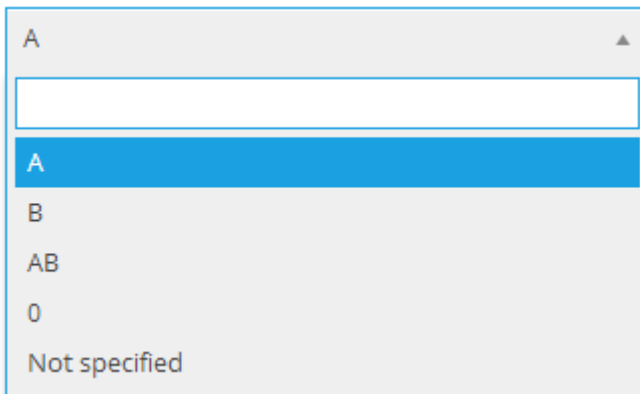
No
Yes

- Drop-down select: A single-select or multi-select feature, with all options entered manually during question creation.

Hint: Consider setting the first option to “Not known” or “Not specified” since there may be cases when none of the options apply or there is not enough information to give a proper answer.

Answers are visualized in pie charts and they can highlight groupings on scatter charts (see 7.1.2.)

ABO



A

A

B

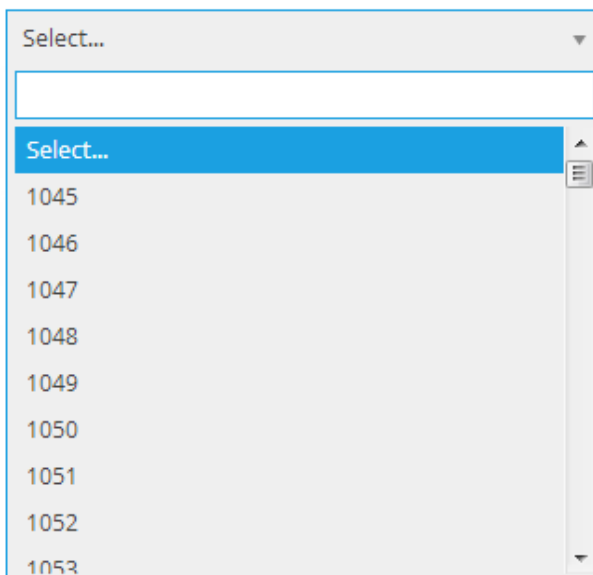
AB

O

Not specified

- Participant link: This question type serves to link participants to other biobank entries. Link a biospecimen to a participant by adding a Participant link type question to the biospecimen profile, or alternatively to a biospecimen questionnaire. During the process of adding a new biospecimen, a drop-down menu appears in the profile or questionnaire with all the participants in your project. Choose a participant to link it to the selected biospecimen. It is also possible to link different participants, e.g. to be able to record relationships between participants.

Participant



Select...

Select...

1045

1046

1047

1048

1049

1050

1051

1052

1053

- Biospecimen link: Similar to Participant link, it may be used to link already added biospecimens to other biobank entries. It can be used to connect related biospecimens.

- Genetic data: This question type can hold genetic data imported from VCF files. Since upload is not done manually, consider adding a separate questionnaire with only a genetic data type question. See **8.6** for more details.

***Hint:** Creating questions can be time-consuming but SmartBiobank™ offers some methods to simplify this process, for example you can copy questions from another project. Read about these options in section 8.*

The software analyses some basic attributes of the answers and warns you if it detects ineffective questions. See **7.1.1**.

You can check what questionnaire the question has been attached to. Click on Show attached questionnaire on the sidebar to view all the attached profiles and questionnaires.

6 Data input = Biobanking

After customizing the settings and creating all the necessary profiles and questionnaires, your project is ready for data input.

6.1 Adding participant and biospecimen data

You can add new participant or biospecimen (collectively known as biobank entry) directly from the *Dashboard* or from the participant or biospecimen list.

Add new participant
Enter participant ID then select profile to fill questionnaire. [Details](#)

Participant ID	<input type="text"/>	?
Participant profile	Patient data	?
Gender	Select...	
Date of birth	<input type="text"/>	(MM/dd/yyyy)
Smoking	Not set	
Physical activity level	Select...	
Family medical history	Select Some Options	
Job role	Select...	
Work hours a week	<input type="text"/>	h/week ?
Activity level	Select...	

Enter a unique code for the entry. If more profiles are available, select the suitable one for the related answer boxes to appear. Click on Save after filling out the boxes and the new entry appears in the participant or biospecimen list. See **6.3.1**.

After single-clicking on an entry, a small menu appears with *View*, *Edit* and *Delete* options. You can also use double-click for a quick access to the overview page.

Participant #P001		incomplete (50%)
Participant ID	P001	
Participant profile	Patient data	
Gender	Male	
Birth	03/17/1976	
Smoking	Yes	
Physical activity level	Sedentary	
Family medical history	1. Diabetes 2. High blood pressure	
Job role	Manager	
Work hours	60 h/week	
Activity level	▲ not set	
Created by	orsolya.ludmany@astridbio.com	
Created	02/21/2014 2:57 PM	
Last modified	02/21/2014 2:57 PM	

The overview page contains all profile data as well as the date of creation and last modification. The profile status is displayed in the upper right corner. See **4.2.2**. Assigned questionnaires and linked participants/biospecimens are listed at the bottom of the page.

6.1.1 Submitting questionnaires

Available questionnaires are listed in the participant and biospecimen entry *Overview* pages. Questionnaire status is displayed below the questionnaire name, and previously submitted questionnaires (marked with the submission date) are listed on the right.

Questionnaires

Body Temperature

[Fill Body Temperature](#) Body Temperature questionnaire isn't filled related to Participant #H0002.

Regular test

[Fill Regular test](#) [04/04/2013: FILL0358](#) [04/04/2013: FILL0359](#) [Table view](#)


Click on *Fill* and answer the questions then click *Save*. Click on previous submissions to view and edit them, or click on *Table view* to compare submitted questionnaires of the same template.

Click on *View* to access questionnaire overview page where you can compare some entry-specific data to global data. See **7.1.2**.

6.1.2 Tables of questionnaire data

When more than one questionnaire is filled out about a biobank entry, Table view offers a quick and convenient overview of the answers. This also makes possible to point out trends.

Regular test

Question	11/07/2013 1:55 PM	02/13/2014 2:33 PM	03/20/2014 4:18 PM
Weight	85.0 kg	76.0 kg	76.0 kg
Temp.	36.6 °C	37.0 °C	36.8 °C
Sickness since last check	No	No	No
Reflexes	Normal	Normal	Normal
Medication since last check	 not set	1. Cold medicine	1. Food supplement
Lymph nodes	Normal	Normal	Normal
Further information		No.3	No.2
Checkup date	11/04/2013	02/14/2014	02/13/2014
Blood pressure	Normal	Normal	Low

Showing 1 to 9 of 9 entries [First](#) [Previous](#) **1** [Next](#) [Last](#)

The Table view option can be found inside the biobank entries at the attached questionnaires.

6.2 Biospecimen management

Track your biospecimens with virtual boxes and refrigerators. You can access them in the *Biospecimen organizer* page.

6.2.1 Add refrigerator

Click on *Refrigerators* and click *Add refrigerator* on the sidebar. Enter the name of the refrigerator you want to add and set its temperature. Location and Description are optional fields but may help in identification if your project uses a large number of refrigerators.

Add new refrigerator

Choose a unique name for your refrigerator then set a temperature. [Details](#)

Name	<input type="text"/>	?
Location	<input type="text"/>	?
Description	<input type="text"/>	?
Temperature	<div>Room temperature (15-25°C) ▲</div> <input type="text"/> <ul style="list-style-type: none">Room temperature (15-25°C)Cold (2-8°C)Frozen (-20°C)-80°CLiquid N2	?

6.2.2 Add biospecimen box

Double-click your new refrigerator in the list or select and click *View* to access the refrigerator overview page. You can click *Add biospecimen box* here, or by following the *Biospecimen organizer* → *Biospecimen boxes* → *Add biospecimen box* route.

Enter a unique name or code for your biospecimen box and select the refrigerator where the box is stored. Box sizes can range from 1×1 to 15×15 slots (the number of rows and columns does not have to be equal). After setting the number and indexing method of columns and rows, click *Save*. Double-click the box to access its overview page with the visualized slots at the bottom of the page.

11. box

Click on a slot to select it. You can add/remove biospecimen to/from a slot.

	A	B	C	D	E	F	G	H	I
1	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY
	Biospecimen: <input type="text"/> Store Disable Reserve								
2	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY
3	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY	EMPTY

Click on a slot to enter the ID of the biospecimen you want to store. You may use a barcode reader for code input. You can disable slots, or reserve them for later use.

Hint: You can resize boxes even if biospecimens are stored inside them. Biospecimens falling outside of the new dimensions will not be lost, but stay in the box “invisibly”. To avoid invisible storage, make sure to remove biospecimens from rows and columns you want to delete.

6.2.3 Biospecimen tracking

Biospecimen tracking is a useful tool in SmartBiobank™ if you are looking to track smaller biospecimen samples like sub-samples of DNA, RNA or serum from a larger Blood sample. If you want to track smaller biospecimen samples, you will first need to create a profile for parent biospecimens and add properties for biospecimen data.

New biospecimen profile

Choose a unique name for your profile and add a description if needed. You can attach properties before or after saving your profile.

General properties

Profile name	<input type="text" value="Blood sample"/>	Choose a unique name for your new profile such as the name of the participant group or the biospecimen type.
Description	<input type="text" value="Parent sample"/>	You have the option of adding a short description to the profile.
Optimal temperature	<input type="text" value="Not selected"/>	

Profile properties

Biospecimen ID	Identifier	
Sampling date	Date	
Volume	Real number	
Participant link	Participant link	
Click to add property		

Save

You can then add new biospecimens using the parent biospecimen profile and create a new profile for aliquots (subsamples of biospecimens samples).

New biospecimen profile

Choose a unique name for your profile and add a description if needed. You can attach properties before or after saving your profile.

General properties

Profile name	<input type="text" value="DNA"/>	Choose a unique name for your new profile such as the name of the participant group or the biospecimen type.
Description	<input type="text" value="Derived specimen"/>	You have the option of adding a short description to the profile.
Optimal temperature	<input type="text" value="Not selected"/>	

Profile properties

Biospecimen ID	Identifier	
Extraction date	Date	
Derived from	Biospecimen link	
Derivative volume	Real number	
Click to add property		
<input type="button" value="Save"/>		

These new biospecimen profiles for the aliquots will contain a Biospecimen Link type property.

When you create a new entry for the biospecimen aliquots, link them back to the original biospecimen. Based on the purpose of your project, you may create different profiles for aliquots used in different processes, or you may create one aliquot profile and add a property to indicate the process. You may also choose to create a biospecimen questionnaire to track changes (e.g volume decrease) in the original biospecimen sample.

Derivative

Questionnaire overview. You can attach questions here. [Details](#)

Overview

Questionnaire name	Derivative
Description	
Applied to	Biospecimens
Created by	orsolya.ludmany@astridbio.com
Created	02/05/2014 1:02 PM
Last modified	03/19/2014 6:00 PM

Derivatives properties

☰ Actor of takeout	Short text	➤
☰ Starting volume	Real number	➤
☰ Volume taken out	Real number	➤
☰ Remaining volume	Real number	➤
☰ Action	Drop Down select	➤
☰ Isolated	Drop Down select	➤
☰ Derived sample	Biospecimen link	➤
☰ Further information	Long text	➤
Click to add question		

With these new biospecimen profiles and questionnaires containing relevant questions, biospecimen tracking and chain-of-custody is achievable as seen below.

Biospecimen #AB040 completed (100%)

Biospecimen ID	AB040
Biospecimen profile	Blood sample
Sampling date	02/06/2014
Volume	100.0 ml
Related participant	A0004
Biospecimen stored	not stored
Created by	orsolya.ludmany@astridbio.com
Created	03/19/2014 5:46 PM
Last updated	03/19/2014 5:46 PM

Questionnaires

Derivative

Fill Derivative	02/21/2014: FILL77163	02/21/2014: FILL77162	02/21/2014: FILL77161	02/21/2014: FILL77167
	02/21/2014: FILL77166	02/21/2014: FILL77165	02/21/2014: FILL77164	02/21/2014: FILL77170
	02/21/2014: FILL77169	02/21/2014: FILL77168	Table view	

Laboratory results

Fill Laboratory results	02/21/2014: FILL77175
--------------------------------	-----------------------

Linked profiles

DNA

Derived from	AB040-D01	AB040-D02	AB040-D03	AB040-D04
	AB040-D05			

RNA

Derived from	AB040-R01	AB040-R02	AB040-R03	AB040-R04
	AB040-R05			

6.3 Listings

Many items in SmartBiobank™ are organized into lists. Lists can contain a variety of information; they usually include the name of the item and the creation date. List elements may be sorted by the headings. Click on the arrows next to the column title. Default order is determined by the creation date of the entry.

You can select how many entries you want to see on one page in the upper left corner; while you can browse between pages in the lower right corner. Click on an item to access some additional functions, such as *View* or *Delete*.

The search box in the upper right corner of the list is a simple function that performs the search in the sortable columns.

In the following chapters we show the specific characteristics of the different lists in SmartBiobank™:

6.3.1 Listing biobank entries

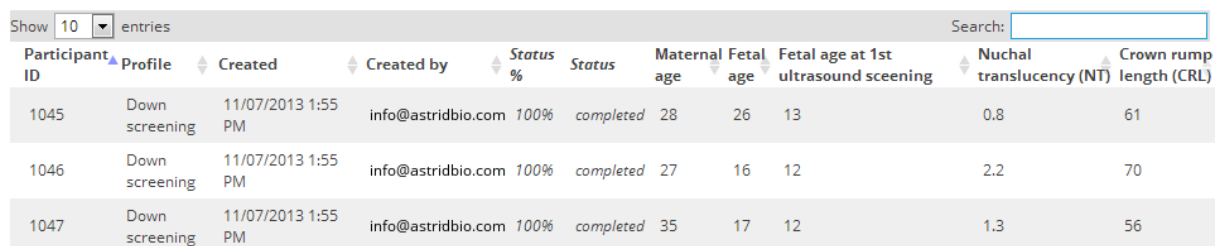
There is a difference between columns of the two biobank entry type.

- Participants: participant ID, profile name, creation date, creator, status,
- Biospecimens: biospecimen ID, profile name, location, creation date, creator, status,

You can sort and search using any of the above categories, other status of the profile or questionnaire. Read more about status in section 4.2.2.

You can narrow the list by selecting only one profile. To do this, click on the profile name on the sidebar. If only one profile is selected, the list displays the first five profile answers (except long texts) that can be used for search (see the below picture).

Participants from Profile Down screening



Participant ID	Profile	Created	Created by	Status %	Status	Maternal age	Fetal age	Fetal age at 1st ultrasound scening	Nuchal translucency (NT)	Crown rump length (CRL)
1045	Down screening	11/07/2013 1:55 PM	info@astridbio.com	100%	completed	28	26	13	0.8	61
1046	Down screening	11/07/2013 1:55 PM	info@astridbio.com	100%	completed	27	16	12	2.2	70
1047	Down screening	11/07/2013 1:55 PM	info@astridbio.com	100%	completed	35	17	12	1.3	56

6.3.2 Listing questions in Questionnaires and properties in Profiles

Properties are used to define profiles and questions are used to define questionnaires. This needs to be kept in mind while working on SmartBiobank™.

Questions are listed in individual Questionnaires, and Properties are listed in Profiles. You can drag and drop questions to modify order. Mandatory questions are marked with asterisks and important questions are marked with exclamation points.

Sample information

Profile overview. You can attach properties here. [Details](#)

Overview

Profile name	Sample information
Description	
Applied to	Biospecimens
Created by	orsolya.ludmany@astridbio.com
Created	03/13/2014 1:43 PM
Last modified	03/13/2014 1:43 PM

Sample informations properties

Biospecimen ID	Identifier	!
Date of sampling	Date	
Gestational age	Real number	
Sample volume	Real number	
Date of centrifuging	Date	
Date of refrigeration	Date	
Click to add property		

6.3.3 Listing questions on the Questions page

The *Questions* page lists all questions and properties in the project and displays some additional information such as: Question text, Display name, Question type, Creation date, Last modification date, Number of questionnaires attached and the Number of answers.

Show entries Search:

Question text	Display name	Type	Created	Created by	Questionnaires	Answers
Fetal age	fetal age	Real number	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	3 questionnaires	216 answers
Further information	Further information	Long text	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	3 questionnaires	231 answers
Maternal age	maternal age	Integer	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	3 questionnaires	216 answers
Sample volume	Volume	Real number	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	3 questionnaires	97 answers
Body Weight	Weight	Real number	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	2 questionnaires	100 answers
Date of sampling	Sampling	Date	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	2 questionnaires	59 answers
AB0 type	AB0	Drop Down select	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	1 questionnaire	51 answers
Age of the patient	Age	Integer	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	1 questionnaire	100 answers
Allergies	Allergies	Boolean	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	1 questionnaire	4 answers
Anti-D antibody	Anti-D antibody	Drop Down select	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	1 questionnaire	92 answers

Showing 1 to 10 of 77 entries First Previous **1** 2 3 4 5 Next Last

6.3.4 Listing questionnaires

All questionnaires are listed on the *Questionnaires* page. They display data that specify the following: Questionnaire name, which Profile type it is applied to (whether it is a Participant or Biospecimen), Creation date, Last modification date, Name of the person who created the template, Number of times it is filled and Number of attached questions.

Questionnaire list

Questionnaire name	Applied to	Created	Last modified	Created by	Filled	Questions
Anti-D antibody test	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	92 times	2 questions
Children	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	37 times	9 questions
Diseases, medicines	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	39 times	13 questions
DNA concentration	Biospecimens	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	27 times	3 questions
Environment	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	39 times	5 questions
General and pregnancy data	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	73 times	7 questions
Genomic Data	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	3353 times	1 question
SCA symptom checker	Participants	03/13/2014 1:43 PM	03/13/2014 1:43 PM	orsolya.ludmany@astridbio.com	none	5 questions

Showing 1 to 8 of 8 entries

First Previous 1 Next Last

6.3.4 Listing biospecimen organizers

There is a difference between the lists of biospecimen storage levels:

- Refrigerators: Name, Location, Temperature, Creation date
- Biospecimen boxes: Name, Storing refrigerator, Temperature, Dimensions, Indexing method, Creation date

Refrigerators are listed on their individual page, while biospecimen box lists also appear on the overview page of refrigerator in which they are stored in.

6.4 Tracking user activity

Project administrators can supervise user activity and the changes in data. This information is stored in two different logs located in SmartBiobank™.

6.4.1 Event log

Click on the Event log tile to view all user activity beginning from the creation of the project. Events are displayed by time and e-mail address of its actor. The logged events are the following:

- Login: both successful and failed login attempts. The time of the login attempt and the e-mail address of the user is recorded.
- Page loads: Which pages were opened and which lists were viewed
- Creation, edits and deletion of biobank entries, profiles, questionnaires, questions and properties and storage
- Search related events: saving, loading and removing criteria

The event log list also marks when errors occurred during an event.

Event log

Time	Event level	Actor	Area	Action	Event message
03/17/2014 3:38:58.0 PM	info	orsolya.ludmany@astridbio.com	Event log	index	Event log at Mon Mar 17 15:38:58 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:38:53.0 PM	info	orsolya.ludmany@astridbio.com	Dashboard	dashboard	Dashboard at Mon Mar 17 15:38:53 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:55.0 PM	info	orsolya.ludmany@astridbio.com	Question	save	Save Question [Name] at Mon Mar 17 15:34:55 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:47.0 PM	info	orsolya.ludmany@astridbio.com	Question	create	Add new question at Mon Mar 17 15:34:47 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:34:40.0 PM	info	orsolya.ludmany@astridbio.com	Question	create	Add new question at Mon Mar 17 15:34:40 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:24:27.0 PM	info	orsolya.ludmany@astridbio.com	Questionnaire	show	Show Profile [Patient data] at Mon Mar 17 15:24:27 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:24:24.0 PM	info	orsolya.ludmany@astridbio.com	Profiles	index	Profile management at Mon Mar 17 15:24:24 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:22:42.0 PM	info	orsolya.ludmany@astridbio.com	Dashboard	dashboard	Dashboard at Mon Mar 17 15:22:42 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:12:57.0 PM	info	orsolya.ludmany@astridbio.com	Questionnaire	show	Show Profile [Patient data] at Mon Mar 17 15:12:57 CET 2014 by orsolya.ludmany@astridbio.com
03/17/2014 3:12:37.0 PM	info	orsolya.ludmany@astridbio.com	Profiles	index	Profile management at Mon Mar 17 15:12:37 CET 2014 by orsolya.ludmany@astridbio.com

Showing 1 to 10 of 738 entries

First Previous 1 2 3 4 5 Next Last

6.4.2 Change log

Changes made to data are registered and can be tracked for biobank entries and structured elements, namely Profiles, Questionnaires, Properties, Questions, Participants and Biospecimens. By carefully following the log, you can back track values to the creation of the entry.

Shown below is a Change log for a question (Diastolic blood pressure). The Change log lists details of the Actor, Date, Domain, Action, Property, Old value, New value

Diastolic blood pressure

Question text	Diastolic blood pressure	Display name	Diastolic
Description	<i>not set</i>	Hint	<i>not set</i>
Mandatory	no	Important	yes
Type	Integer	Created by	orsolya.ludmany@astridbio.com
Created	03/18/2014 4:24 PM	Last modified	03/18/2014 4:28 PM

Integer specific properties

Minimum value	<i>not set</i>
Maximum value	<i>not set</i>
Unit of measure	Hgmm

Changelog

Show 10 entries							Search:
actor	date	domain	action	property	old value	new value	
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Display name		Diastolic	
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Type	Real number	Integer	
orsolya.ludmany@astridbio.com	03/18/2014 4:26 PM	Question	modified	Important	false	true	
orsolya.ludmany@astridbio.com	03/18/2014 4:27 PM	Question	modified	Description		minimum pressure	
orsolya.ludmany@astridbio.com	03/18/2014 4:28 PM	Question	modified	Description	minimum pressure		

Showing 1 to 5 of 5 entries First Previous 1 Next Last

7 Research center

SmartBiobank™ offers many tools to make use of the data your project stores. Use these tools to control the research process using the feedbacks of the software and enhance your research results using statistical and visualization tools.

The Research center category on the Dashboard lists functions that visualize data with the help of charts. Many charts can be accessed from more than one page. Charts can be saved as an image file or can be printed directly using the upper-right drop-down menu of the charts.

7.1 Descriptive statistics

Descriptive statistics can be used for two different things:

1. It lets you monitor and improve the quality of your project by assessing the usefulness of the questions you work with.
2. It provides visualization tools to understand your results better.

Its main page collects and displays all box plots and pie charts in the project, which helps you overview your data or visually compare results.

Descriptive statistics

Adequacy of questions

Unattached questions (4/69)

- [Date of analysis](#)
- [Expected date of delivery](#)
- [Protein concentration](#)
- [Sample taken](#)

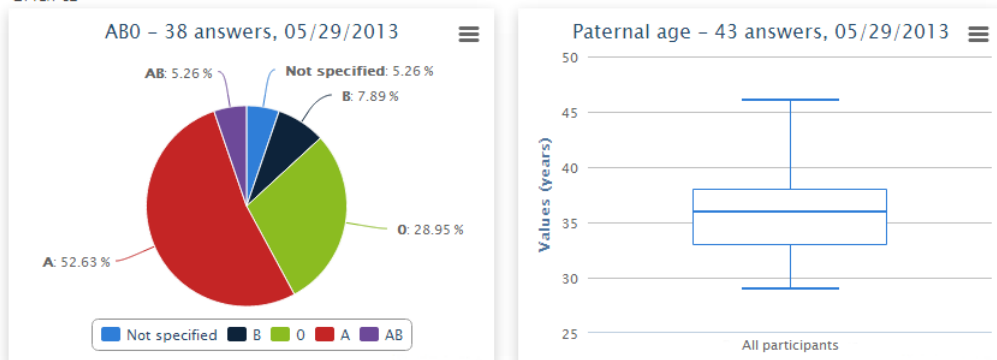
Unanswered questions (3/69)

- [Further information](#)
- [Specify disease of maternal stepsiblings\(s\)](#)
- [Specify disease of paternal stepsiblings\(s\)](#)

Constants (5/69)

- [Child with disease among maternal stepsiblings](#) (false)
- [Child with disease among paternal stepsiblings](#) (false)
- [Down syndrome in father's family](#) (false)
- [Hereditary diseases of father](#) (false)
- [Pregnancy-independent maternal diseases](#) (cardiovascular)

Charts



7.1.1 Adequacy of questions

Creating questionnaires that provide useful data can be difficult, even careful planning can result in barely usable data. The software provides assistance in assessing the adequacy of your questions based on analyzing the answers. This evaluation may highlight the need for question redesigning. The following table shows the meaning and possible implications of three evaluation categories:

Adequacy	Meaning	Implication
Unattached	The question has not been attached to any questionnaires or profiles.	Data input is not possible.
Unanswered	Either the questionnaire has not been filled yet or the answer box has been left blank every time.	Redefining the question text or changing the question type may be needed to improve adequacy.
Constant	The answer has been the same in every submission.	Answers do not provide useful information.

Charts found in the *Research center* may refine general adequacy results as highlighted by the following example.

Example: Several pie charts of Boolean type data highlight problems: around 90% of the answers were “Unknown”, and the rest were “No”. Since two different types of answer were registered, the software recognizes these questions as “Effective”, but looking at the charts makes it clear that they do not provide useful data.

7.1.2 Charts

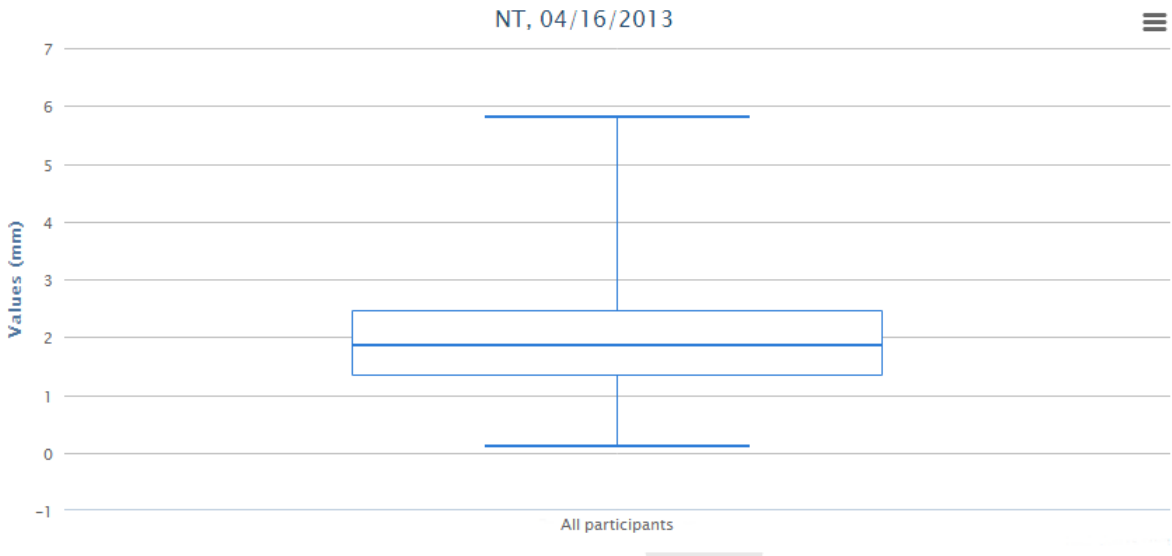
Chart visualization is available for specific types of data. Integer and Real number data are plotted as box plot, while Drop-down and Boolean types are visualized as pie charts.

Note: The charts are plotted based on all answers given to that question, even if the property or question is attached to more questionnaires or profiles. To avoid mixing data of different sources, create new questions of the above types for each profile and questionnaire.

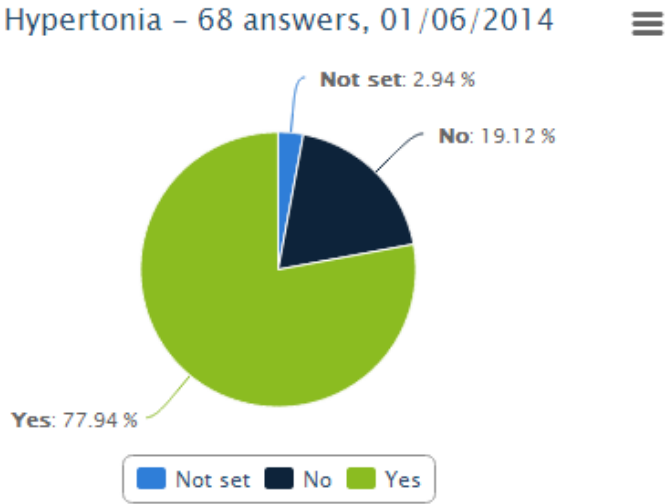
Each chart can be viewed on the overview page of the question of property of the necessary type. You can see examples of the two chart types below.

The first chart is a box plot based on a Real number type data called Nuchal Translucency, used in Down syndrome screening. The expression is usually abbreviated as NT, which was set as the Property name of the question, therefore appearing in the top of the chart. Hover your cursor over

the box plot in SmartBiobank™ to see the following statistics: minimum, lower quartile, median, higher quartile and maximum.



The second chart is from another project, and it is based on the data provided by a Boolean type question. Hovering your cursor over the pie chart sections in SmartBiobank™ opens a small window containing the section name and percentage. You can create an exploded pie chart by clicking on a section to separate one sector from the others.

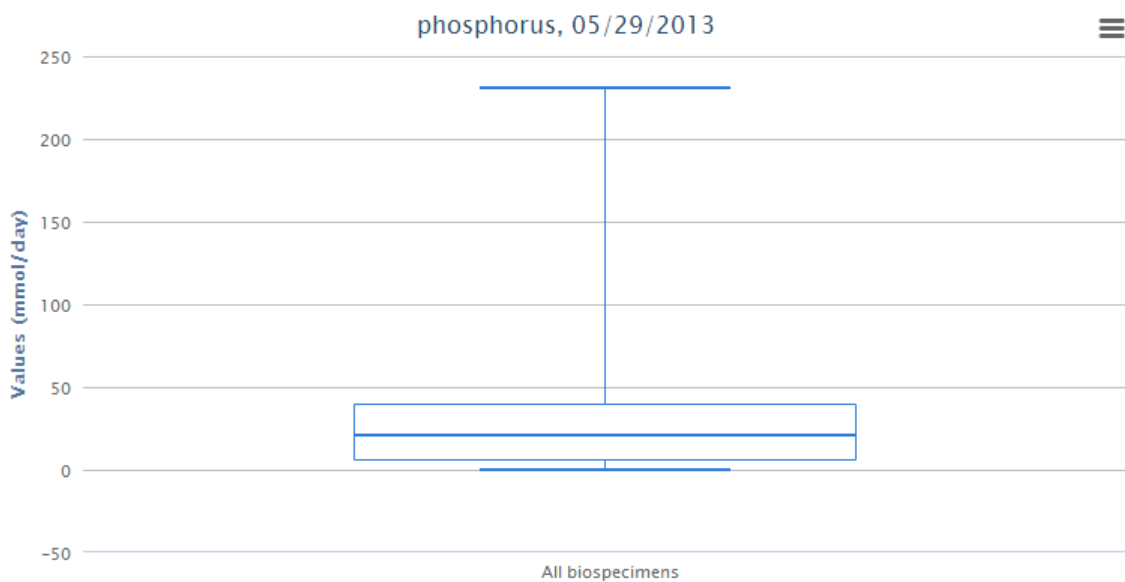


In both cases, the date in the chart title is the same date the chart was generated by the software, in other words when the page with the chart was loaded. This makes it possible to know which state the chart represents. For example, if data input is continuous, you can download a chart every day to track changes over time.

Charts can also serve as a quality control tool in your project. Below are two examples for this functionality:

Example: There might be very few charts in your project. This might be because most of your data is collected as text. If you want to improve the statistical analysis of your data, consider changing the question/property types to a drop-down selection or adding some numeric questions/properties. Please note that changing type may lead to the loss of the already entered answers.

Example: The error bars of the box plots are too big. It might mean that some records are incorrect, either because of a typo error or the wrong unit of measure had been selected.



Note: Neither the inbuilt question effectiveness indicator, nor the above stated example determines what a good project is. Different projects have different needs, and yours may not need numerical input at all.

Note: Charts of global answers include data given for other questionnaires or profiles if the question is attached to them, which may or may not influence the use of these charts in your research.

7.3 Search & Query

On the *Search & Query* page, you can search for participants or biospecimens by based on their data in related profiles or questionnaires. In this part of SmartBiobank™, the difference between questions and properties is not important and both are handled the same way, and hence we will refer to both as questions.

This sample search is based on the answers for questions in biobank entries. You first have to select the question to specify the search criteria.

Search & Query

Group name Untitled group 03/14/2014 11:58:56.594 AM [Rename](#)

Searching for Applied to

Criteria No criteria added. Click to the Add criteria button to start querying the database.



Can I help you?

- 1** Do you search Participant or Biospeciment?
Select what are you looking for
- 2** Add some criteria
With criteria you can add a restriction to a property. Several operators are available to creating criteria.
- 3** Understanding the result
The result list will contain all participant or biospeciment (depends on your choice) that correspond at least one criterion. You can combine criteria to narrow the number of the elements.
- 4** Use groups to visualize information
Rename your group by clicking the Rename link. Create several groups and go to the Data Visualization page to visualize properties in the selected groups.

Clicking on *Add criteria* opens the window that lists every question and property in the project.

Add new criteria



Filter questions

Show entries Search:

Question text	Display name	Type	Mandatory	Important
Average sleeping hours	Average sleeping hours	Integer	no	no
Blood pressure	Blood pressure	Drop Down select	no	no
Body height	Body height	Integer	no	no
Body Mass Index	BMI	Real number	no	no
Body weight	Body weight	Real number	no	no
Calorie intake per day	Calorie intake per day	Integer	no	no
Date of birth	Birth	Date	yes	yes
Date of the examination	Date of the examination	Date	no	no
Family medical history	Family medical history	Drop Down multiple select	no	no
Further information	Further information	Long text	no	no

Showing 1 to 10 of 21 entries First Previous **1** 2 3 Next Last

Select a questionnaire or profile from the *Filter questions* box to display attached questions. You can use the quick search in the upper right corner to find a specific question. (Read more about list searches in section 6.3.)

Add new criteria



Filter questions

Examination

Genomic data

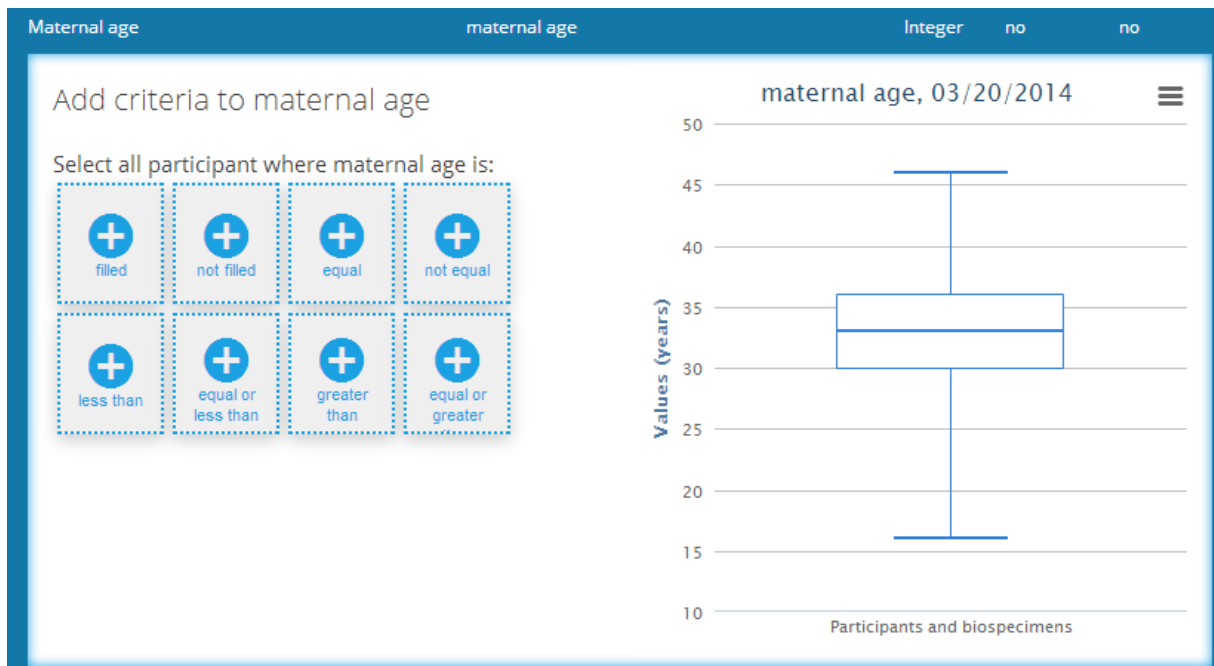
Main profile

Show entries Search:

Question text	Display name	Type	Mandatory	Important
Date of birth	Birth	Date	yes	yes
Family medical history	Family medical history	Drop Down multiple select	no	no
Further information	Further information	Long text	no	no
Gender	Gender	Drop Down select	yes	yes
Job role	Job role	Drop Down select	no	no
Physical activity level	Physical activity level	Drop Down select	no	no
Smoking	Smoking	Boolean	no	no
Work hours a week	Work hours	Integer	no	no

Showing 1 to 8 of 8 entries First Previous **1** Next Last

You also need to select an operator to further specify your search. The question and the chosen operator together define the criterion.



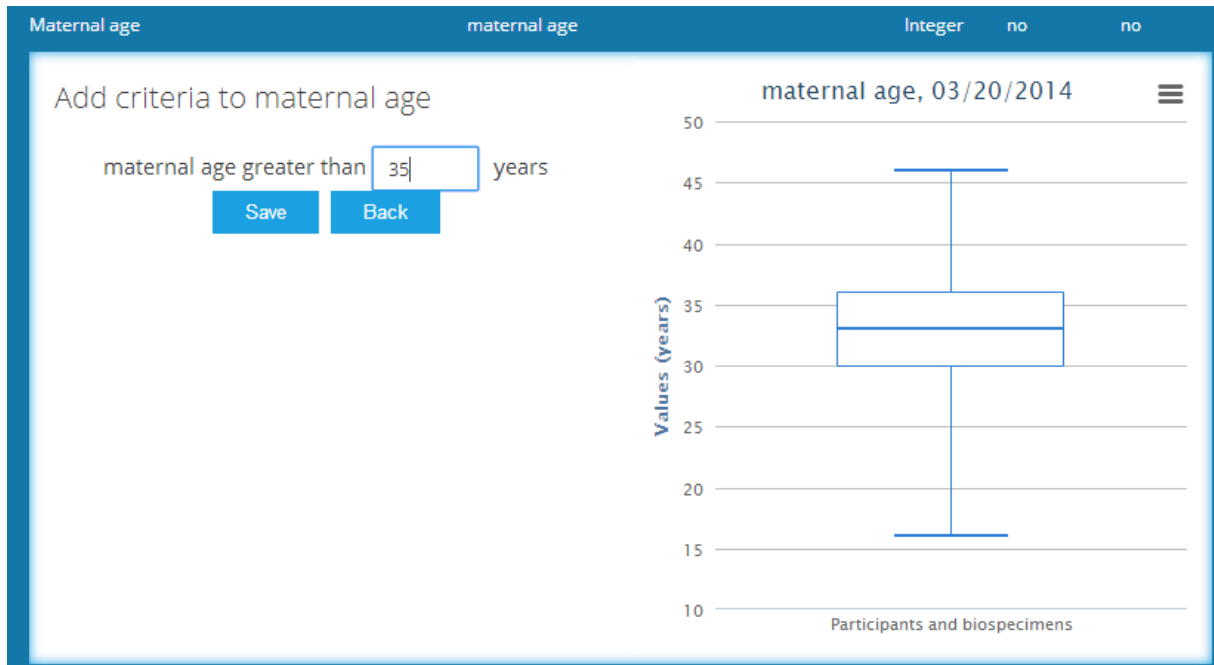
The following operators can be used:

- **Filled:** An answer is registered.
- **Not filled:** The answer box has been left blank.
- **Equal:** The answer equals the entered value.
- **Not equal:** The answer differs from the entered value.
- **Less than:** The value of the answer is lower than the entered value. Only available for numerical and date type data.
- **Equal or less than:** The value of the answer is equal or lower than the entered value. Only available for numerical and date type data.
- **Greater than:** The value of the answer is higher than the entered value. Only available for numerical and date type data.
- **Equal or greater than:** The value of the answer is equal or higher than the entered value. Only available for numerical and date type data.
- **Contains:** The answer contains the entered text fragment. Only available for text type data and participant/biospecimen links.

The chart exhibits the distribution of answers for the question (as long as the answer type is not text). Taking a look at the chart may help you decide which criteria should be used to make your search more effective. In our example above, it is immediately obvious that creating a criterion for maternal age over 46 years would be meaningless.

You can keep adding criteria. Close the window to view the list with the matching results.

After you have set the search criteria, click on *Save* and a list appears with the matching results. You can further specify your search by adding more criteria. To do this, simply click on *Add criteria* on the sidebar.



Although the number of criteria to be added is not limited, the speed of the search process is dependent on this number. In practice, a maximum of five criteria should be enough to find the required biobank entries.

Search & Query

Group name Untitled group 03/20/2014 3:32:36.826 PM [Rename](#)

Searching for Applied to

Criteria

- Maternal age greater than 35 years ✕
- Risk factors present not equal None ✕
- Does the mother take medicins regularly? equal Yes ✕

[Add criteria](#)

Group members

Show entries Search:

Participant	■ Maternal age greater than 35 years	■ Risk factors present not equal None	■ Does the mother take medicins regularly? equal Yes
1062	yes	yes	yes
1055	yes	no	yes
1034	yes	yes	yes
1030	no	no	yes
1029	yes	yes	yes
1028	yes	yes	yes

Showing 1 to 6 of 6 entries (filtered from 64 total entries) First Previous **1** Next Last

By default the drop down menus on the search results page are set to 'All' for each criterion, meaning that all entries are listed which fit at least one of the selected criteria (there is at least one 'yes' in all displayed rows). Modify this automatic selection to vary search results. If you select 'Yes' for a given criterion, the list narrows your search for those entries where this criterion is true. Similarly, if you select 'No', the list narrows your search for those entries where this criterion is not true.

Search groups are saved automatically, and unless a new search is deliberately started, the software keeps updating your current search. Click on New group to save your current search settings and start defining a new search group. Saved groups are listed on the Saved groups page that you can reach from the sidebar. Renaming your saved search can make future searches faster.

The *Search and Query* function also enables you to save your searches and use them in other SmartBiobank™ features such as *Data Visualization (7.4)*.

7.4 Data visualization

The *Data visualization* page offers statistical comparison of previously defined groups of biobank entries. First, you have to define at least two groups of biobank entries on the *Search and Query* page, see **7.1**. To make further work easier, rename your searches to know what kind of groups they contain.

After you have saved your searches, go to the *Data visualization* page and select search groups from the list in the *Groups to visualize* box. Then select questions and properties you want to compare for these groups. The ideal number of properties or questions to visualize at one time is four.

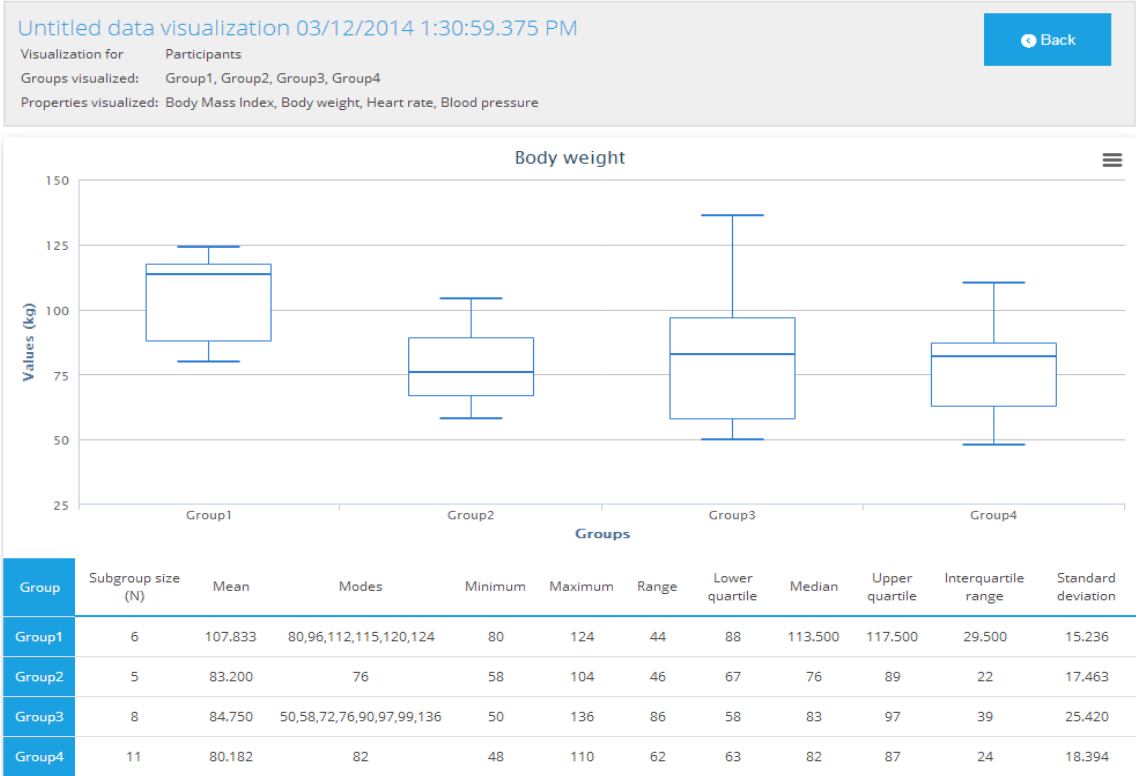
Data visualization

Data visualization name Untitled data visualization 03/12/2014 1:30:59.375 PM [Rename](#)

Visualization for	Participants	
Groups to visualize	Group1 x Group2 x Group3 x Group4 x	Please select at least one group to visualize properties. Groups can be created at Search & Query page. Save a query and it will be available as a group here!
Properties to visualize	Body Mass Index x Body weight x Heart rate x Blood pressure x Add property	Please select at least one property to visualize it!
Visualize!		

The system then generates comparable box plots and pie charts for the specified groups.

Data visualization



8 Data exchange

Data exchange allows you to save manual work by enabling you to upload questions and data from a table to SmartBiobank™ – including uploading and mapping genetic data to participants. Data exchange is only available for project owners and administrators.

8.1 Import questionnaires

Questionnaires and profiles with attached questions and properties can be imported from an Excel worksheet. For this purpose you must use our existing template. It can be downloaded from the *Import questionnaires* page, located inside the *Data exchange* page or among the hidden tiles of the Dashboard.

#	1. Question text	2. Display name	3. description	4. hint	5. type	6. required
#	Text of the question as it appears when someone fills the questionnaire. If you leave it blank, a random name will be generated.	This text will appear on charts and statistics. This column is optional. If you leave it blank, the text from "1. Question text" appears.	Description of the question. This column is optional.	This text will appear next to the answer box when someone fills the questionnaire. Use it to help users with giving appropriate answers. This column is optional.	Type of the question. It must be one of the following: shortText : Short text box with simple text input, longText : Long text box with simple text input, integer : integer number, double : real number, date : date, boolean : boolean, dropDown : selectable options, multSelDropDown : select multiple options patient : link entry to patient, sample : link entry to sample. Default value: shortText.	Whether the question is mandatory. Default value: no.
#	Participants or Biospecimens, default: Participants		Questionnaire or Profile, default: Questionnaire			
#	Applied: Participants	Type: Profile				
#	Gender	Gender			dropDown	yes
#	Date of birth	Birth			date	yes
#	Smoking				boolean	no
#	Physical activity level				dropDown	no
#	Family medical history				multSelDropDown	no
#	Job role				dropDown	no
#	Work hours a week	Work hours		average	integer	no
#	Further information				longText	no

The software recognizes one sheet as one questionnaire or profile, so it is possible to import more than one questionnaire or profile at the same time. Change the name of the sheet to your preferred questionnaire or profile name.

Note: Do not forget to indicate in the seventh row whether the sheet is attached to participants or biospecimens and whether it is a profile or questionnaire.

Technically, since the question creation and property creation differs only in the end use, we will refer to both of them as questions. In the file, options found on the question creation interface are arranged into columns while individual questions are listed in rows. Rows with a # sign in the first column contain descriptions and examples, and they will not be imported into your questionnaire. You can read about the functions of columns in **5.2.1**.

After finishing extending the template, upload the file to the *Import questionnaires* page where the software analyses its content and warns you about emerging problems such as question name duplications.

Note: You can decide to upload a revised version or proceed without the problematic questions. The software offers to rename questions with name duplications or you can choose to use the question already available in SmartBiobank™.

Import questionnaires

Your questionnaires and question must be in the appropriate format. If you do not know this format or use this function for the first time, you can download and extend our template xls file from here: [download excel file](#).

Overview

Available questionnaires	Profile Data, Regular test
Available questions	15 question processed.
Missing names	⚠ Warning: 1 name is missing!
Name duplications	<p>⚠ Warning: 6 question name already exists in the SmartBiobank. Because question names must be unique, an additional number will be added to the end of the question. You can change the name of the question in the excel file and upload again.</p> <p>⚠ Warning: 1 questionnaire name already exists in the SmartBiobank project. You can change the name of the question in the excel file and upload again.</p> <p>How do you want to resolve name duplications:</p> <div style="border: 1px solid black; padding: 2px;"><p>Rename duplicated elements</p><p>Rename duplicated elements</p><p>Do not import, use available object from SmartBiobank</p><p>Do not import, skip these objects</p></div>
Missing question type	⚠ Warning: 1 question type is missing!
Unrecognized question type	⚠ Warning: 1 question type can not be recognized in the uploaded file. Without modifications, this question will be simple short text. You can change the types in the excel file and upload again.
<div style="display: flex; justify-content: space-around;">Upload another fileProceed import</div>	

Imported questionnaires and profiles can be accessed straight after the process is finished by clicking the button with the name of the questionnaire/profile.

8.2 Copy questionnaires

If you want to use previously created questionnaires in a new project, you can simply copy them using the *Copy questionnaires* page. Besides saving you time, this option can facilitate cooperation between different projects. You can only copy from projects you are participating in. No biospecimen or participant information will be copied with this method, all your data is securely stored in its original place.

Hint: Use the copy questionnaire function only if more or less you need the same questions with the same attributes in your new projects. In case you need to change the questions substantially, you are advised to use export questionnaires function, see [0](#).

Available questionnaire

Questionnaire name	Applied to	Profile/Questionnaire	Project	Created	Last modified	Participants	Questions
Melanoma examination	Biospecimens	Questionnaire	Achat Bio	02/18/2014 3:44 PM	03/21/2014 1:56 PM	none	24 questions
Cervix samples	Biospecimens	Profile	HPV Project	03/13/2013 2:21 PM	03/21/2014 1:54 PM	none	4 questions
<div style="display: flex; align-items: center; gap: 10px;"> Copy Click the Copy button to copy the selected questionnaire/profile into the Skin Project. </div>							
Vagina swab samples	Biospecimens	Profile	HPV Project	03/19/2013 9:45 AM	03/21/2014 1:54 PM	none	4 questions
Patients under HPV examination	Participants	Profile	HPV Project	03/13/2013 2:20 PM	03/21/2014 1:53 PM	none	5 questions
Melanoma project	Participants	Profile	HPV Project	04/16/2013 11:31 AM	03/21/2014 1:53 PM	none	4 questions
Body Temperature	Participants	Questionnaire	HPV Project	04/24/2013 3:42 PM	03/21/2014 1:52 PM	none	4 questions
Regular test	Participants	Questionnaire	HPV Project	03/20/2013 11:22 AM	03/21/2014 1:51 PM	none	12 questions
Patient data	Participants	Profile	Demo Project	02/05/2014 1:02 PM	03/20/2014 1:44 PM	none	8 questions
Derivative	Biospecimens	Questionnaire	Demo Project	02/05/2014 1:02 PM	03/19/2014 6:00 PM	none	8 questions
Blood sample	Biospecimens	Profile	HPV Project	03/19/2014 5:43 PM	03/19/2014 5:44 PM	none	3 questions

Showing 1 to 10 of 246 entries First Previous 1 2 3 4 5 Next Last

8.3 Export questionnaires

You can export questionnaires and questions into the same table which is the input format of the importing function. This option can be useful when you later wish to import those questions into another project after some editing.

Export questionnaires

Please select the questionnaires you want to export.

Select questionnaires to export

Questionnaires to export	Export single questionnaire
File name	export-questionnaire-03-21-2014-14-04.xls

Export single questionnaire

Click the questionnaire you want to export.

Selected to export	Patient data Questionnaire
Change questionnaire	Change
Download	

8.4 Import Participants & Biospecimens

Participant or biospecimen data, along with their questionnaire data can be imported from an Excel file. It is a convenient tool to move your archives into SmartBiobank™. Before starting the process, you have to make sure that the profiles and questionnaires already exist in your project with all the necessary properties, questions and settings. Excel files waiting to be imported must be based on the profiles or questionnaires already present in the system.

1	A	B	C	D	E	F	G	H	I	J
2	Participants	Patient data								
3	Participant ID	Gender	Date of birth	Smoking	Physical activity level	Family medical history	Job role	Work hours a week	Further information	
4	P001	Male	03.17.1976	TRUE	Sedentary	Diabetes High blood p	Manager	60		
5	P002	Male	10.30.1982	FALSE	Light	Alzheimer's disease	Support staff	20		
6	P003	Female	01.06.1984	FALSE	Sedentary		Unemployed	0		
7	P004	Male	11.13.1963	FALSE	Moderate	Asthma	Self-employed	45		
8	P005	Female	05.18.1979	FALSE	Light		Supervisor	40		
9	P006	Male	03.19.1987	TRUE	Sedentary		Unemployed	0		
10	P007	Female	02.11.1989	FALSE	Sedentary		Unemployed	0		
11	P008	Male	07.10.1983	FALSE	Moderate	Alcoholism	Support staff	40		
12	P009	Male	03.08.1968	TRUE	Light	Gallbladder disease	Support staff	45		
13	P010	Male	06.20.1975	FALSE	Sedentary		Support staff	50		
14	P011	Male	07.29.1978	TRUE	Light		Support staff	35		
15	P012	Female	09.10.1973	FALSE	Sedentary	Heart disease Depres	Support staff	40		
16	P013	Female	12.12.1992	FALSE	Moderate	Leukemia	Support staff	30		
17	P014	Male	03.28.1979	FALSE	Light	Cancer	Self-employed	50		
18	P015	Male	12.30.1978	TRUE	Sedentary	Heart disease High blc	Supervisor	45		
19	P016	Female	07.30.1979	FALSE	Sedentary	Depression	Support staff	40		
20	P017	Male	09.04.1987	FALSE	Moderate	Alcoholism Other De	Unemployed	0		
21	P018	Female	05.11.1969	FALSE	Light	Obesity	Supervisor	55		
22	P019	Male	10.07.1983	TRUE	Sedentary	Obesity	Supervisor	60		
23	P020	Female	03.08.1968	FALSE	High		Support staff	40		
24	P021	Male	06.20.1975	FALSE	Sedentary		Unemployed	0		
25	P022	Male	08.01.1982	TRUE	Sedentary	Alcoholism Gallbladde	Manager	40		
26	P023	Male	06.06.1983	TRUE	Light		Support staff	60		
27	P024	Male	05.16.1965	FALSE	Moderate	Diabetes Obesity	Support staff	20		

Each row stands for a biobank entry and each column stands for a property or a question. Their intersections are the individual answer boxes. However, there are important rules to follow:

1. The first two rows are headings: the first row contains the exact name of the profile or questionnaire, and the second row contains the exact Question text (see 5.2.1), one name in each column.
2. The A1 cell must contain either *Participants* or *Biospecimens* to indicate what type of entries are imported.
3. A2 must state *Participant ID* or *Biospecimen ID*,
4. The first column contains the participant or biospecimen codes.

Hint: To import answers for multiple select drop-down type questions, separate answers with a vertical bar with spaces on each side. The cell will look similar to this example: [option1 | option2]

Profiles and questionnaires are organized into separate sheets. The sheet must be named to match the profile or questionnaire name. One file may contain only one sheet with the same name.

Note: *If a sheet containing profile data have the same ID twice, the second set of data will not be imported. In case of questionnaire data, multiple IDs on a sheet will result in multiple submissions of the questionnaire, all marked by the date of the import.*

However, you do not need to import questionnaire data along with profile data if the latter have already been uploaded to you project. By creating only a questionnaire sheet, you can import only questionnaire submissions for already existing biobank objects.

Note: *To import date type data, use the Excel date format.*

After the import is completed, the review message is displayed. The review consists of numerical summary of the import success, the import log, and the option to undo the action.

Right after an import is completed, information about its success is displayed. Even though some preliminary data are available before the import is performed, it is intended to check if the necessary structure is present to perform the import. Namely, it checks if the questionnaire or question names are correct, or if there are duplicated biobank objects, but it does not give information about matching the question settings or the ID format restrictions.

Note: *Make sure that the sheet containing profile data is placed before the sheet containing questionnaire data, as they are processed one after another and questionnaires need to have already uploaded profile data.*

Import Report

Import overview

	All		Imported		Error	
	Profile	Questionnaires	Profile	Questionnaires	Profile	Questionnaires
Patients	30	30	19	18	11	12
Samples	0	0	0	0	0	0
Summary	30	30	19	18	11	12

Started at	11/14/2013 9:11 AM
Started by	orsolya.ludmany@astridbio.com
Status	Finished successfully

Revoke data import

When you revoke your import ALL Patient, Sample, Filled questionnaires from your project that created within this import **will be deleted!** Be careful, once you revoke your import, there is no going back, all related data will be deleted!

By security reason, only the user created the import or the project owner can revoke an import!

[Revoke import](#)

Messages

[Download messages](#)

[Show all](#)

- Import process started at 11/14/2013 9:11:22.868 AM
- Patient AB01 from sheet Main profile imported successfully!
- Patient AB02 from sheet Main profile imported successfully!
- Patient AB03 from sheet Main profile imported successfully!
- Patient AB04 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Anemia.
- Patient AB05 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Asthma.
- Patient AB06 from sheet Main profile imported successfully!
- Patient AB07 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Diabetes | High blood pressure | Obesity | Seizures.
- Patient AB08 from sheet Main profile cannot be imported: Invalid boolean value to question "Currently smoking?": null.
- Patient AB09 from sheet Main profile imported successfully!
- Patient AB10 from sheet Main profile cannot be imported: Invalid drop down value for question "Family Medical History": Obesity.

8.4.1 Previous data imports

Imports need careful planning and large amount of preparation and it is possible that only the second or third attempt will result as desired. To make revisions easier, SmartBiobank™ offers the option to revoke previously imported datasets.

The *Previous data imports* page lists all imports performed in the project.

Previous data imports

start	executed by	patients (added/error/all)	samples (added/error/all)	all (added/error/all)	rollbacked	finish status
11/14/2013 9:37 AM	orsolya.ludmany@astridbio.com	25 / 14 / 39	0 / 0 / 0	25 / 14 / 39	no	Finished successfully
View						
Tip: Double click to open View.						
11/14/2013 9:11 AM	orsolya.ludmany@astridbio.com	19 / 11 / 30	0 / 0 / 0	19 / 11 / 30	no	Finished successfully
09/30/2013 12:29 PM	orsolya.ludmany@astridbio.com	20 / 0 / 628	0 / 0 / -2	20 / 0 / 626	no	Finished successfully
09/30/2013 10:44 AM	orsolya.ludmany@astridbio.com	20 / 1 / 628	0 / 0 / -2	20 / 1 / 626	no	Terminated by unexpected error
09/29/2013 2:45 PM	orsolya.ludmany@astridbio.com	31 / 0 / 31	0 / 0 / -2	31 / 0 / 29	no	Finished successfully
09/27/2013 9:38 AM	orsolya.ludmany@astridbio.com	31 / 1 / 163	0 / 0 / -2	31 / 1 / 161	no	Finished successfully
09/26/2013 11:24 AM	orsolya.ludmany@astridbio.com	0 / 0 / 30	30 / 0 / 30	30 / 0 / 60	no	Finished successfully
09/26/2013 11:16 AM	orsolya.ludmany@astridbio.com	0 / 0 / 30	0 / 30 / 30	0 / 30 / 60	no	Finished successfully
09/26/2013 11:16 AM	orsolya.ludmany@astridbio.com	0 / 0 / 30	0 / 30 / 30	0 / 30 / 60	no	Finished successfully
09/26/2013 11:14 AM	orsolya.ludmany@astridbio.com	30 / 0 / 30	0 / 30 / 30	30 / 30 / 60	no	Finished successfully

Showing 1 to 10 of 23 entries

First Previous **1** 2 3 Next Last

Opening a list element displays the same review that appears after the import is completed. (See 8.4.). Only the user performing the import and the project owner can revoke an import for security reasons.

8.5 Export data

Project data can be exported into Excel files so you can apply more elaborate analysis methods. You can choose to export:

- all data
- all participants
- all biospecimens
- selected profiles and questionnaires

The structure of the exported data matches the importable file structure.

Export participant & biospecimen data

Participant and biospecimen data that are stored in profiles and questionnaires can be exported here. SmartBiobank offers a generated file name, which you can change below.

What would you like to export?

You can select all patients, all samples, selected profiles, selected forms for export. Profiles or questionnaires that are not filled will not be exported.

Export mode

- Export all data:** All participant and biospecimen data will be exported from your SmartBiobank project.
- Export all participants:** All participant data will be exported from your SmartBiobank project.
- Export all biospecimens:** All biospecimen data will be exported from your SmartBiobank project.
- Export selected profiles and questionnaires:** Select profiles or questionnaires to export.

File name

Export selected profiles and questionnaires

Please select profiles and questionnaires to export. Only profiles and questionnaires that you select will be exported. At least one element must be selected!

Select to export:

Download export file

Click on the Download button to download the requested information in a single Excel file.

[Download](#)

8.6 Import genomic data

If you upload genomic data for participants, you will get a powerful tool to recognize genetic markers of diseases. The import takes four steps and there are two different files involved.

1. Create a new questionnaire applied to participants, with only one Genomic data type question attached.
2. Upload a VCF file that contains information about the genetic variants.

Genomic data import

Participants need to have Genomic data type question in a questionnaire to be able to map them. Genomic data must be in VCF format. You can download our biospecimen file here: [download VCF file](#).

Upload VCF

VCF sample number	30
VCF sample IDs	Show
Number of variants	6085

[Upload VCF again](#) [Next](#)

3. Genetic mapping: you need to attach the VCF biospecimen IDs to participants in your project. The mapping is performed with a CSV file that you have to upload in the second step. The CSV file is automatically generated based on the VCF file you have uploaded, and you can download it in the second step.

Genomic data import

Participants need to have Genomic data type question in a questionnaire to be able to map them. Genomic data must be in VCF format. You can download our biospecimen file here: [download VCF file](#).

Upload mapping file

You have to upload a mapping of the VCF biospecimen IDs and the participant IDs. This must be a CSV file, with the first column containing the VCF biospecimen IDs. You can download and extend the CSV file generated from the VCF file you have uploaded in the previous step: [vcf_ids.csv](#). Click the button or drag the mapping file from your file manager to this page to start the upload.

[Upload a file](#)

This file has two columns:

- a. The first column contains the VCF biospecimen IDs (with 'VCF_ID' written in the first cell)
- b. The second column contains the Participant IDs ('PARTICIPANT_ID' is written in the first cell). As the structure suggests, the genomic information marked by the first column will be linked to the participant indicated in the second column.

Microsoft Excel tends to insert commas to replace column borders, when the CSV file is saved. To edit the file effectively, we recommend Notepad++ that offers some advanced edit functions.

- The third step is a feedback about the success or issues about the files or the mapping. The table shows the invalid participant IDs, invalid VCF IDs, missing mappings, and valid mappings. Click on Show to see what IDs fall under each category. If you are not content with these results, you can restart the process with corrected files: click on Upload mapping again.

Upload mapping

Click the "Finalize Import" below to complete the import. Note, that only the mappings visible in the "Valid mappings" section of the above table will be loaded to SmartBiobank. The missing mapping section shows those VCF IDs which are present in the VCF file, but no mapping was provided. The "Invalid participant IDs" and "Invalid VCF IDs" sections show IDs not recognized, the mappings related to these IDs are ignored. If the results in the above table are not as you expected, use the "Upload mapping again" button to upload a fixed mapping file again.

Number of valid mappings	1 Show
Number of invalid patient IDs	29 Show
Number of invalid VCF IDs	0 Show
Number of missing mappings	29 Show

[Upload mapping again](#) [Finalize import](#)

- If the results are satisfactory you can proceed with the last step: finalize your import.

Note: You need to attach Genomic data type question to participants to be able to do the mapping. The most convenient way to do this is to create a questionnaire with this question only. It is possible to enter genomic data manually for every participant individually, but it is only feasible if the vcf file to be imported contains data for one participant only.



Development office
Astridbio Technologies Kft.

4029 Debrecen
Lőrinc pap u. 9.
Hungary

E-mail: info@astridbio.com