





# **Phenotype Database**

Jildau Bouwman









### Why a Phenotype Database?

- Multiple analysis methods (clinical chemistry, transcriptomics etc.) used for systems biological questions
- Unpublished (negative) data
- Complex study designs in nutrition
- Lack of power for systems biological questions
- Lack of individual data
- Lack of (study) meta-data (age, BMI, analysis method etc.)







### Phenotype database goals

- Collect, harmonize and distribute data
- Already available: a database for mechanistic intervention studies
- Helpdesk function/ training for new users (NuGO)
- Example on how the Phenotype Database works: analysis of the challenge response in several studies







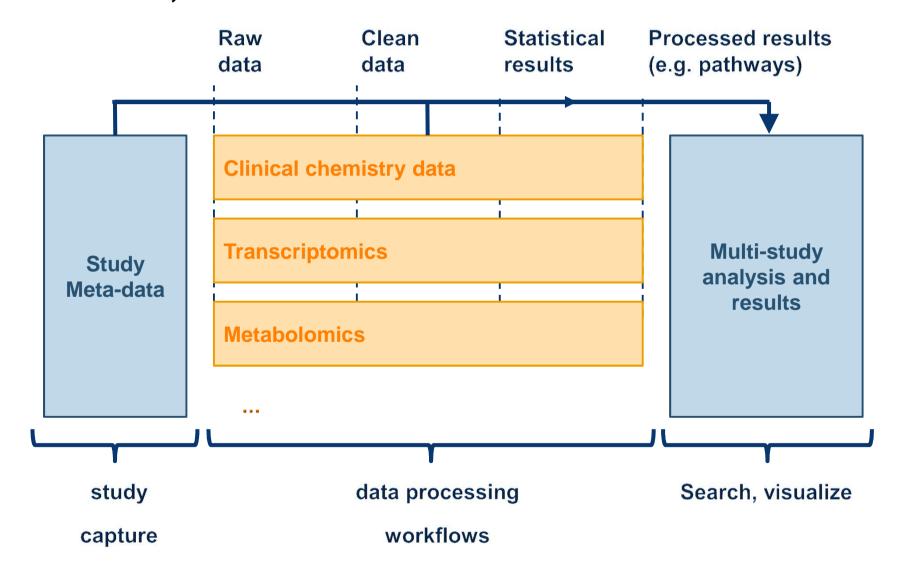
### Phenotype database goals

- Collect, harmonize and distribute data
- Already available: a database for mechanistic intervention studies
- Helpdesk function/ training for new users (NuGO)
- Example on how the Phenotype Database works: analysis of the challenge response in several studies





## Collect, harmonize and distribute data





#### Introduction

Phenotype Database is an application that can store any biological study. It contains templates which makes it possible to customize.

In order to allow flexibility to capture all information you require within a study, and to make it possible to compare studies or study data, the system uses customizable templates and ontologies. It is especially designed to store complex study designs including cross-over designs and challenges.

Phenotype Database facilitates sharing of data within a research group or consortium, as the study owner can decide who can view or access the data. In addition, Phenotype Database can stimulate collaborations by making study information publicly visible. New studies can be based on study data within the database, as standardized storage is stimulated by the system.

#### Quicksearch

Search term

more advanced searches can be performed here.

#### Quick Start

Through the studies menu you can either create, view or import studies (or study data). 'Create a new study' will guide you through several steps to include your study into the system where question marks (\*\*) will explain what information is required. You can (quick) save your study to complete it at another point in time, or use import study data to import large datasets (for example: many subjects) from an excel sheet into your study. Several data-types of different platforms (assays) can be linked to your study, like simple assays (e.g. clinical chemistry or Western blot) or metabolomics.

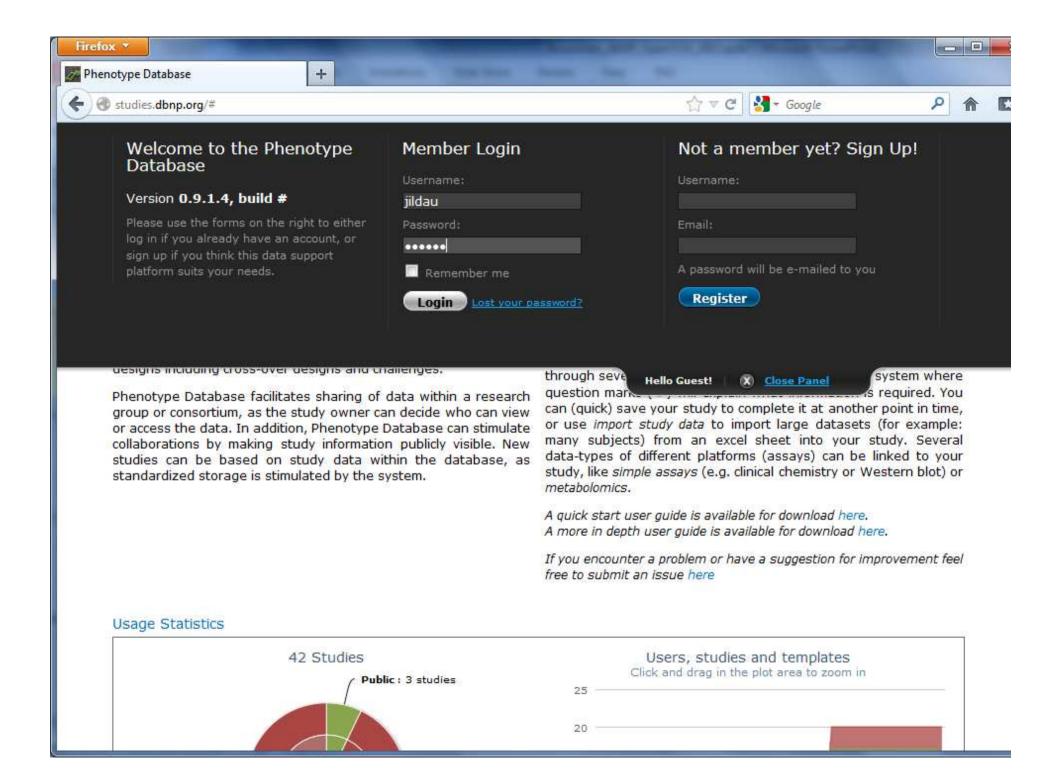
A quick start user guide is available for download here.

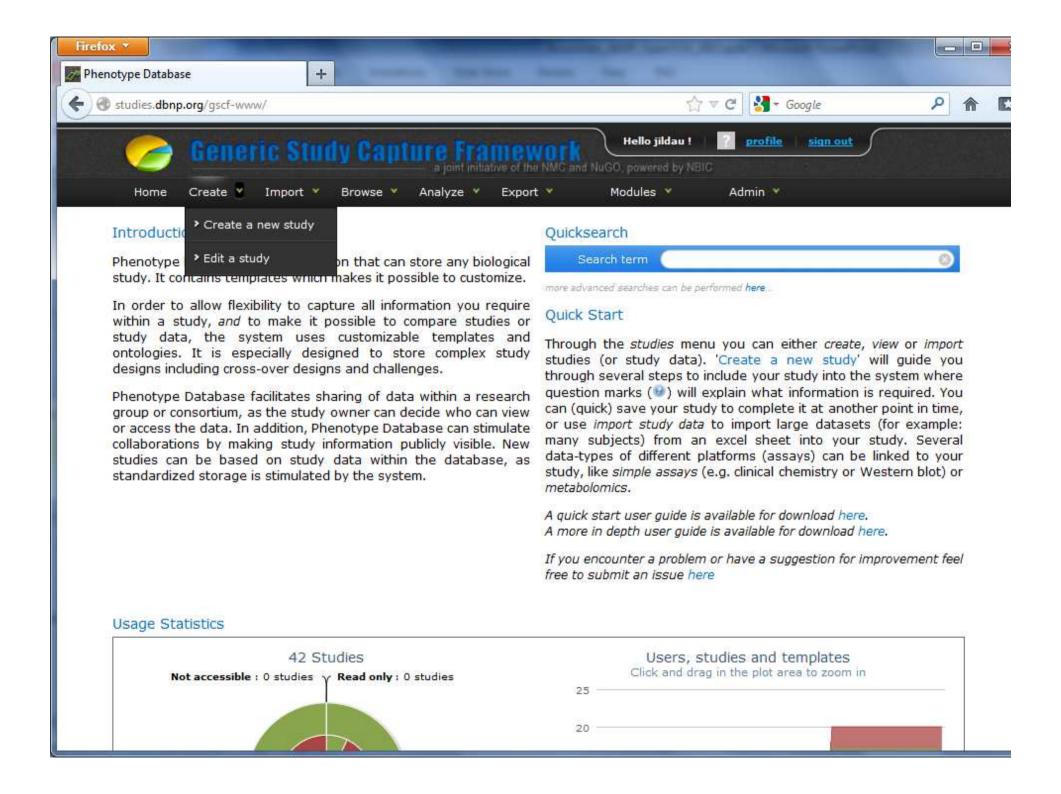
A more in depth user guide is available for download here.

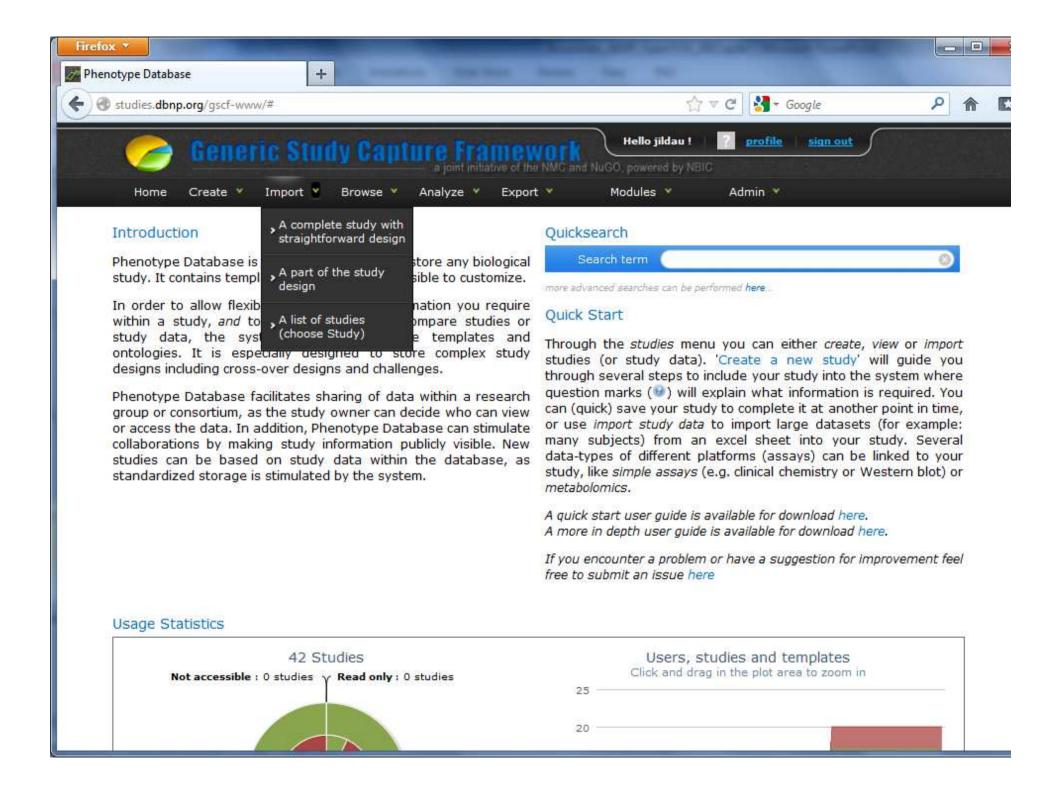
If you encounter a problem or have a suggestion for improvement feel free to submit an issue here

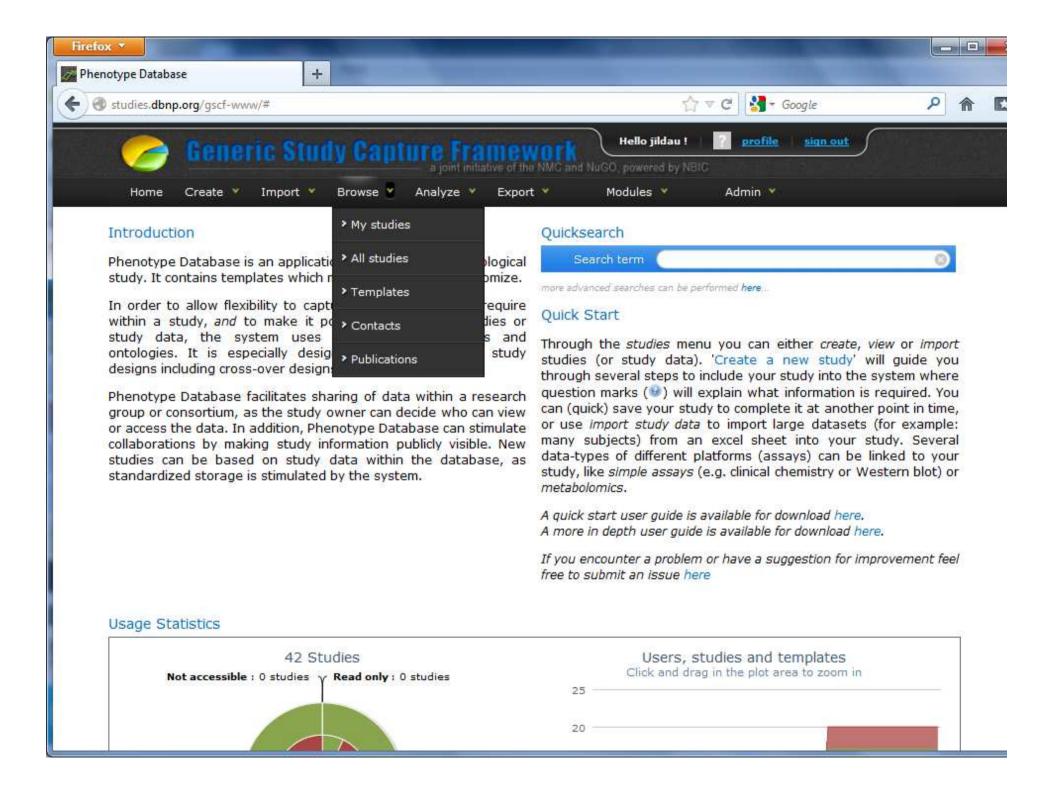
#### Usage Statistics

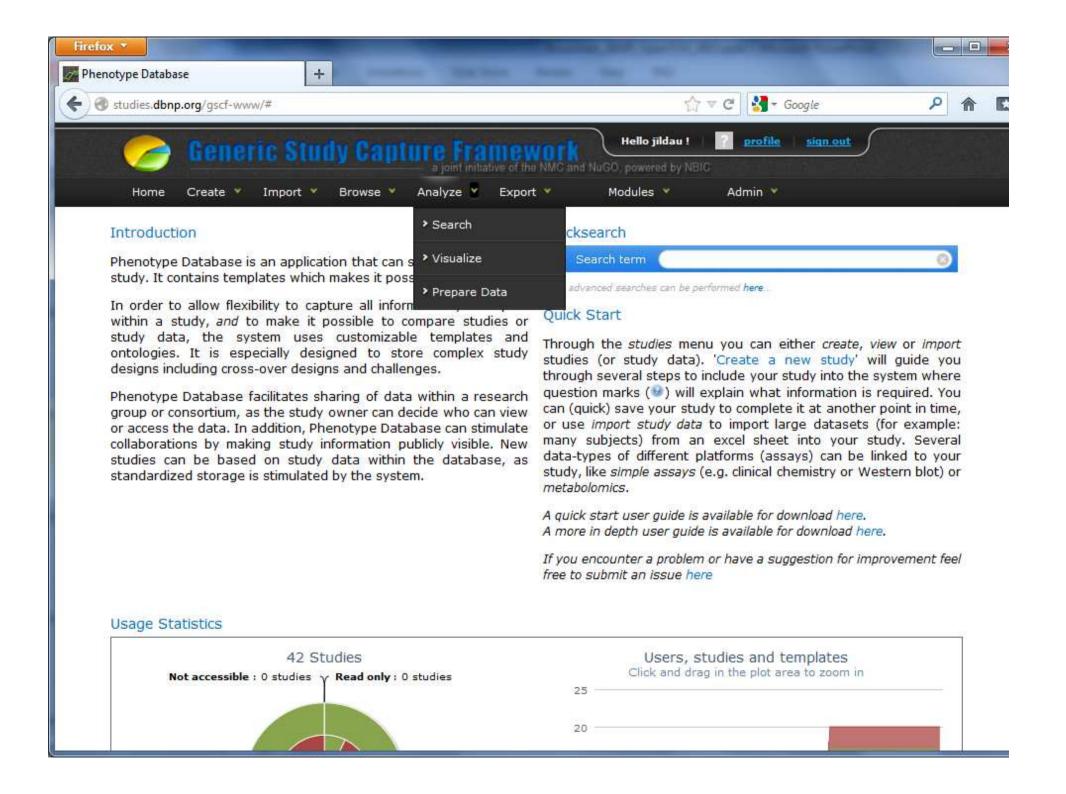


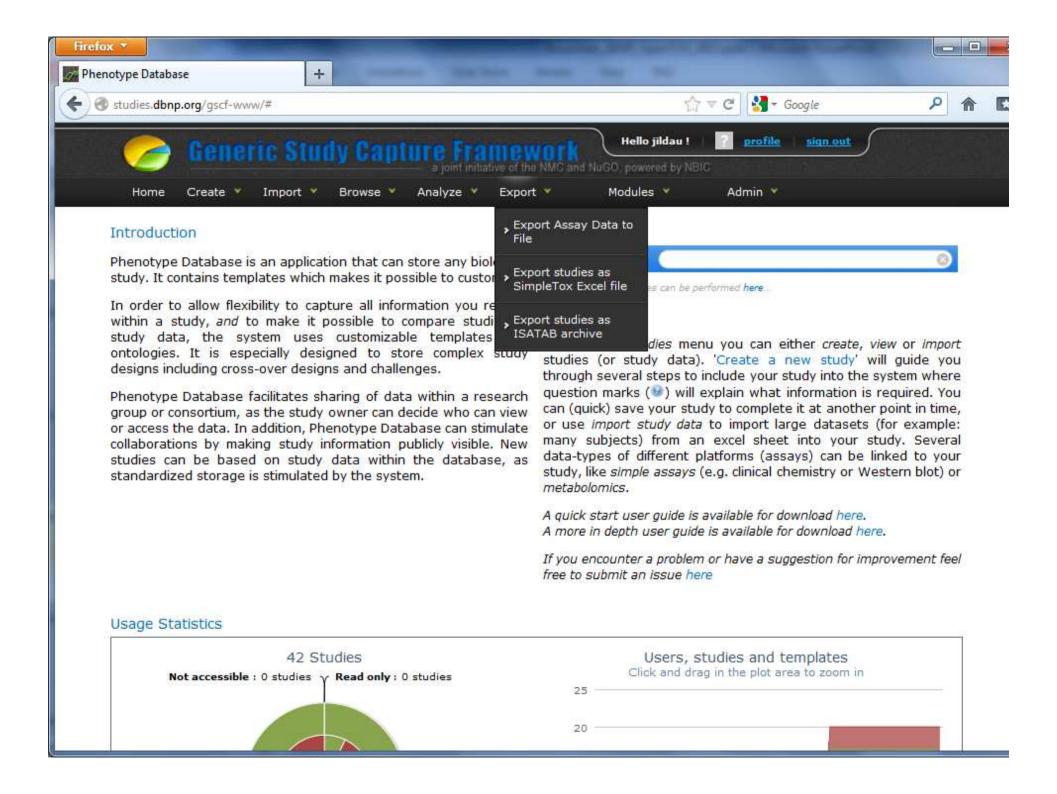


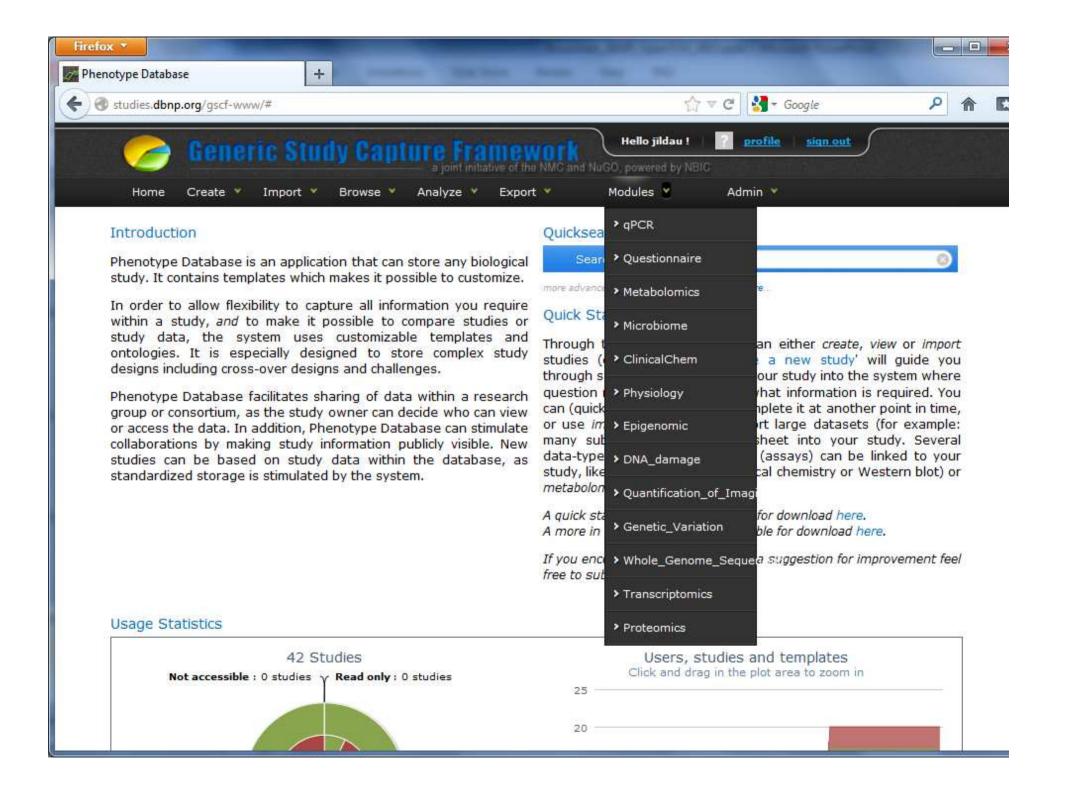


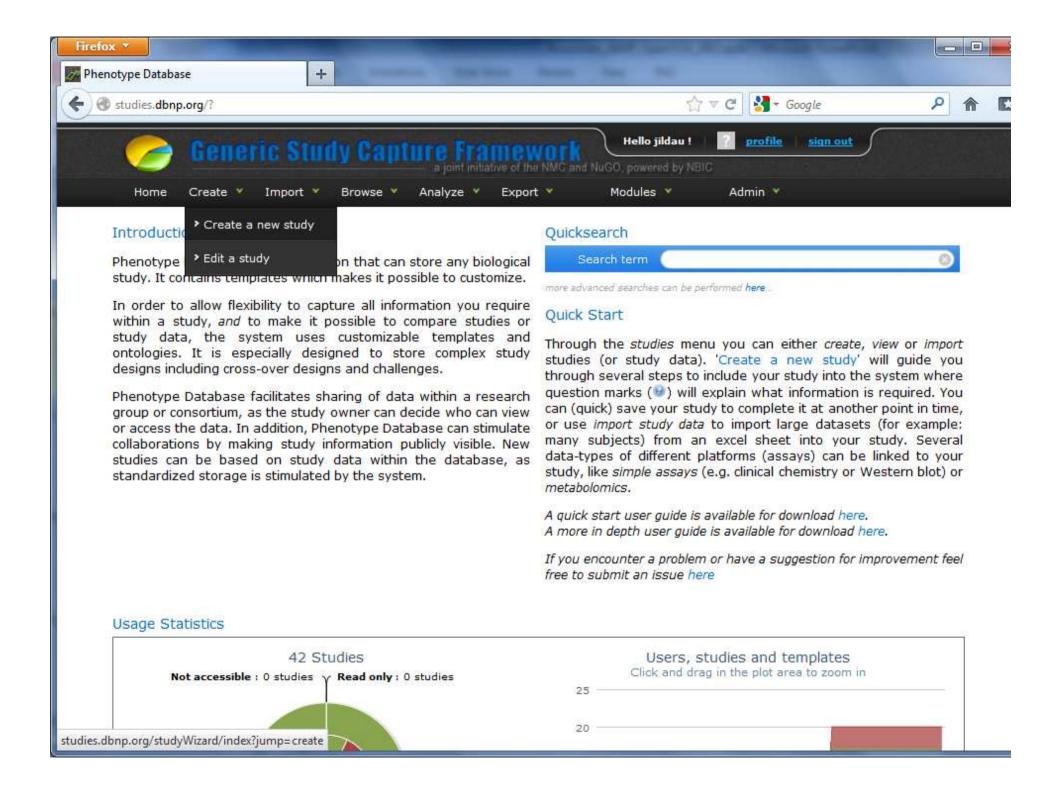


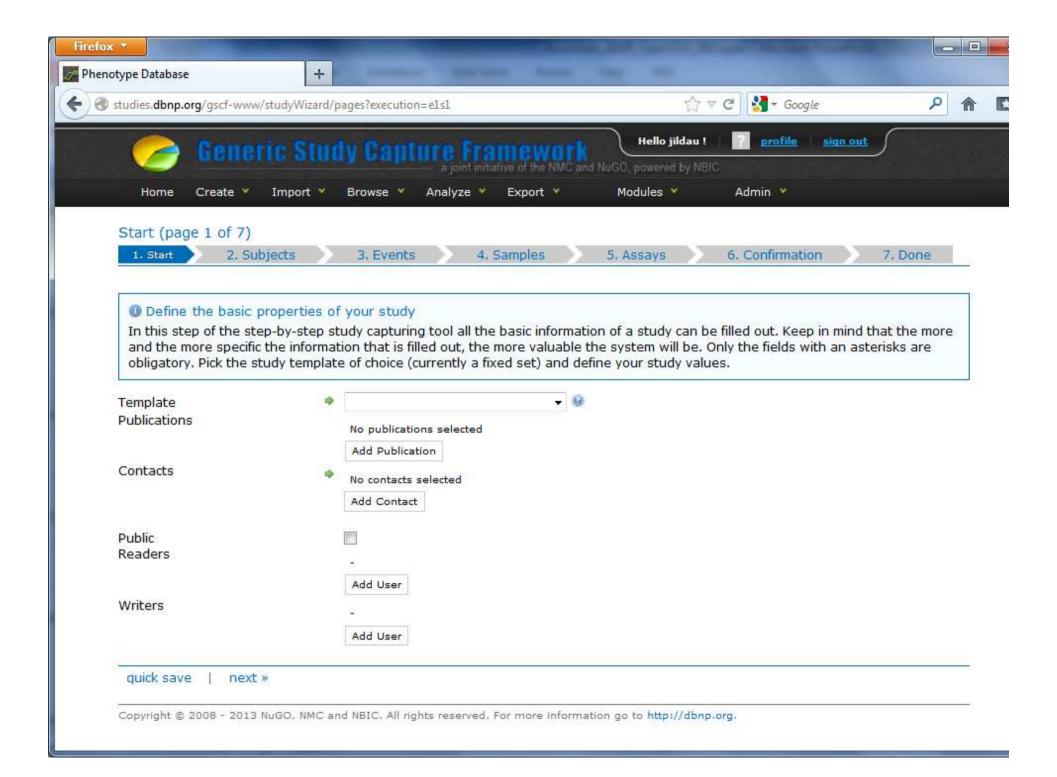


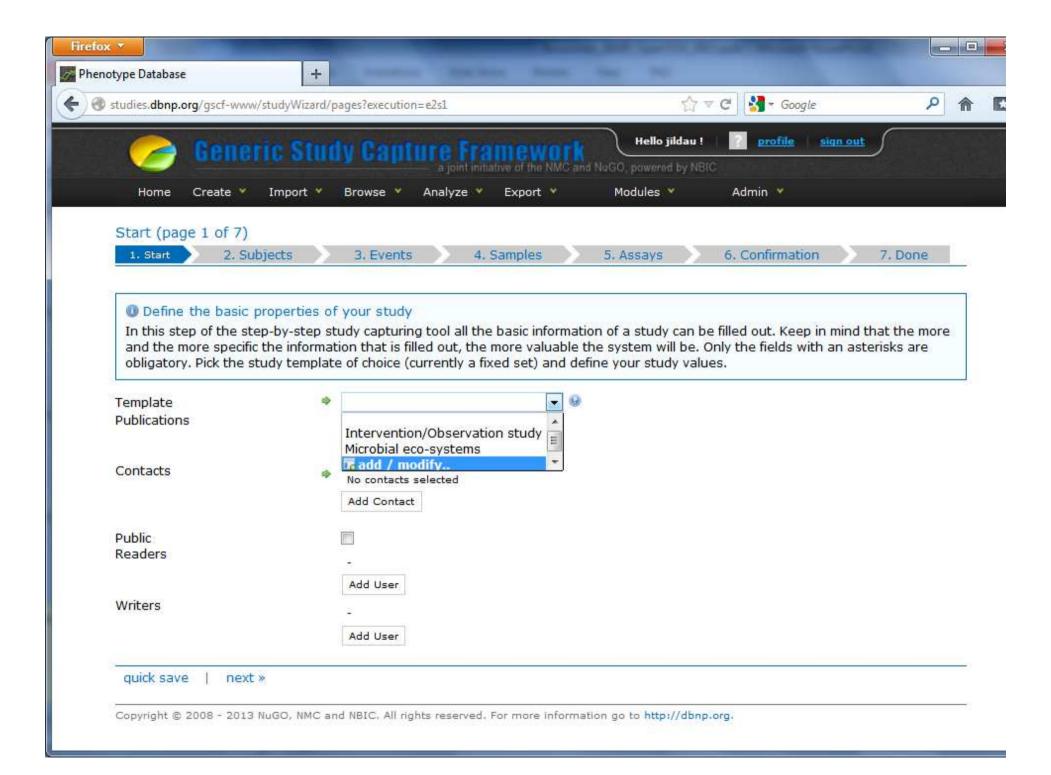


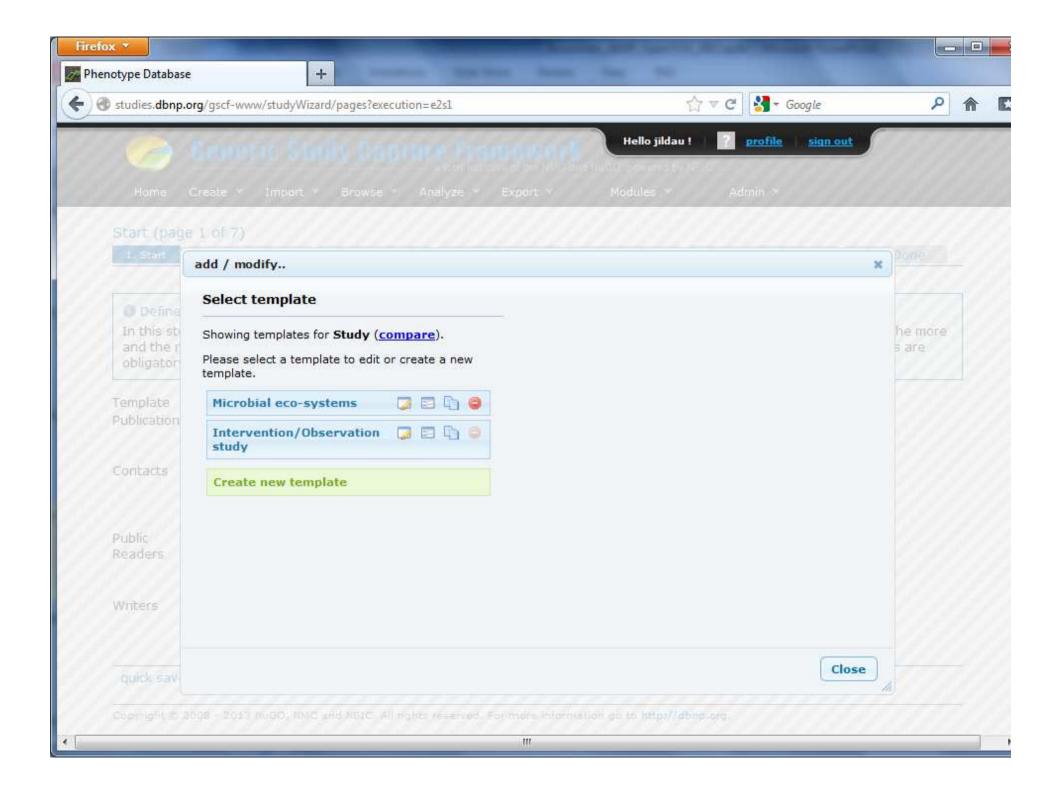


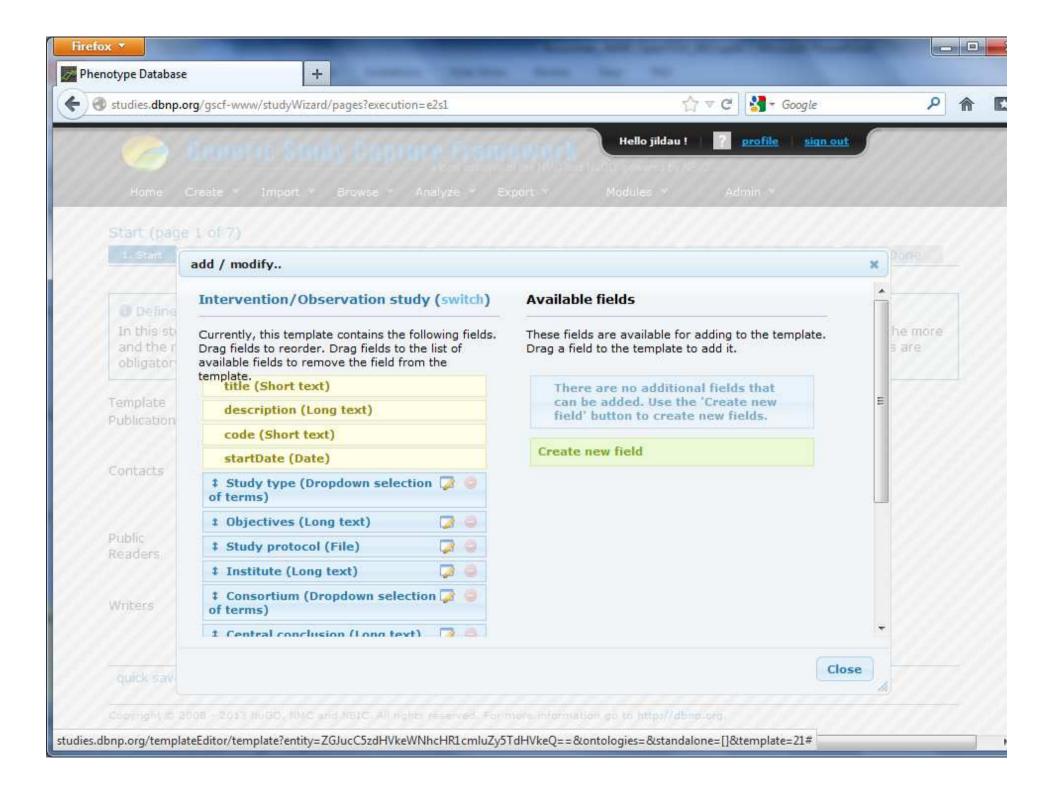


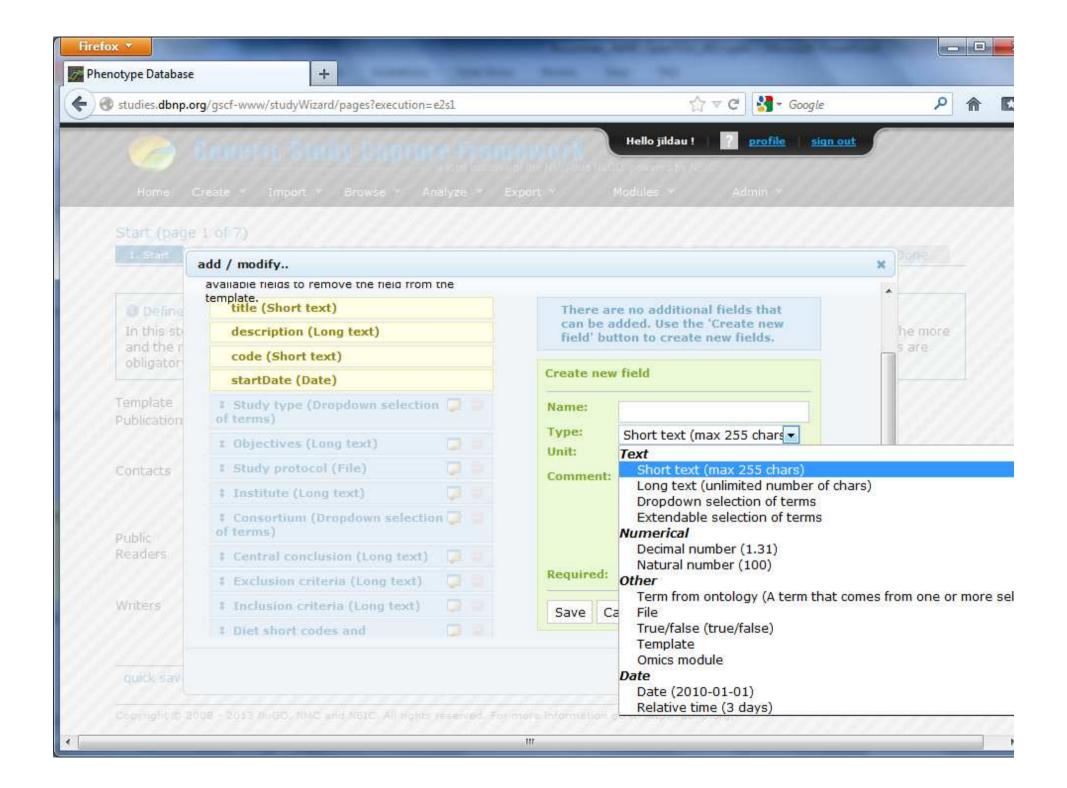


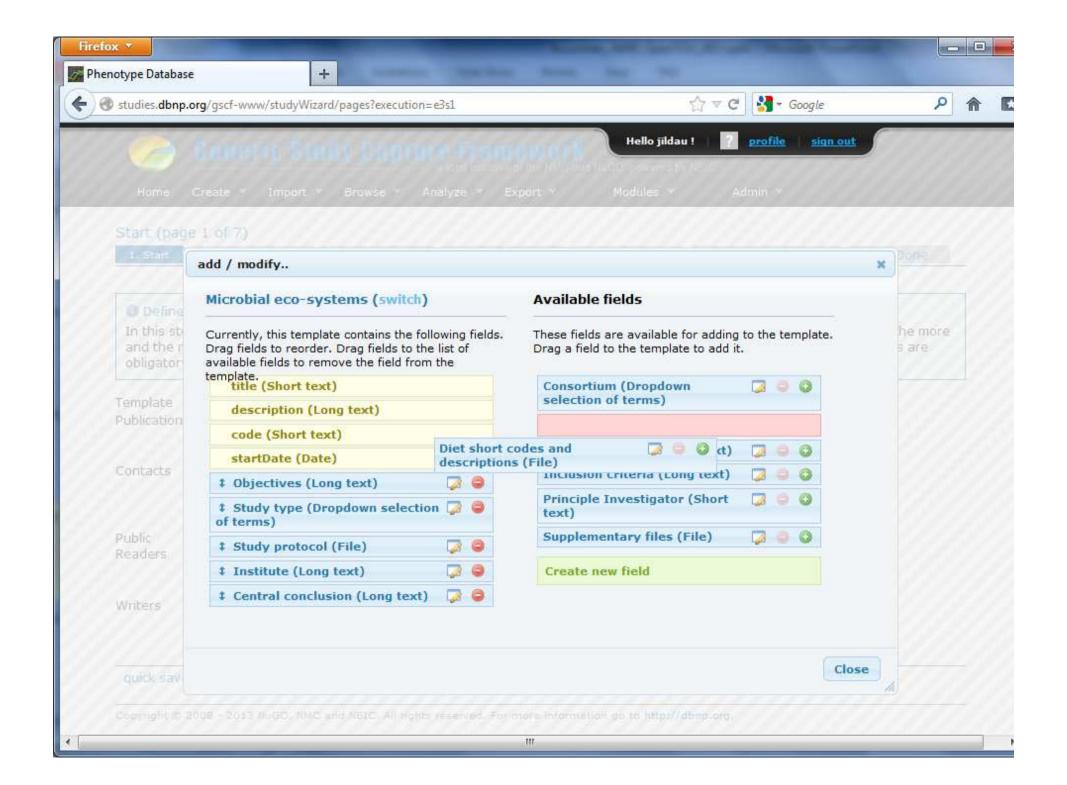


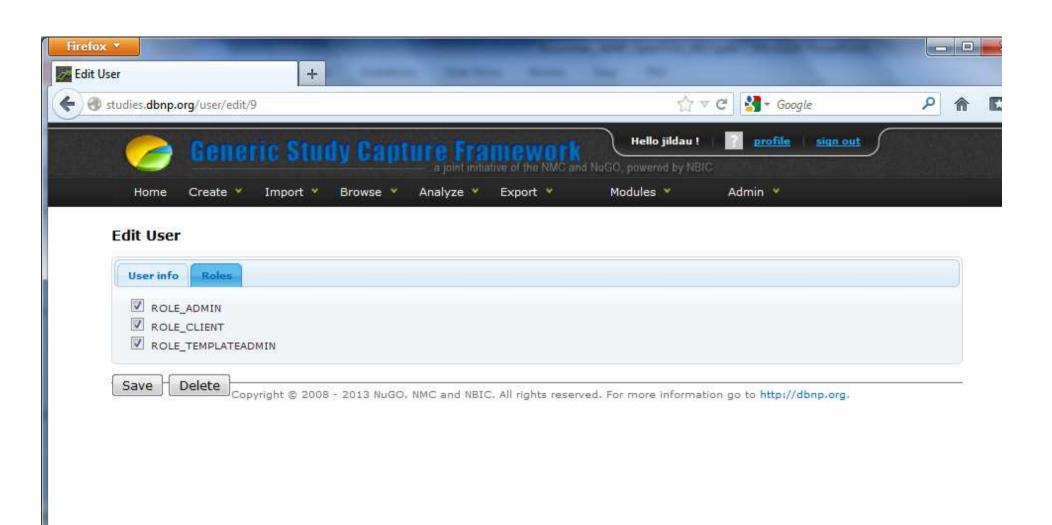


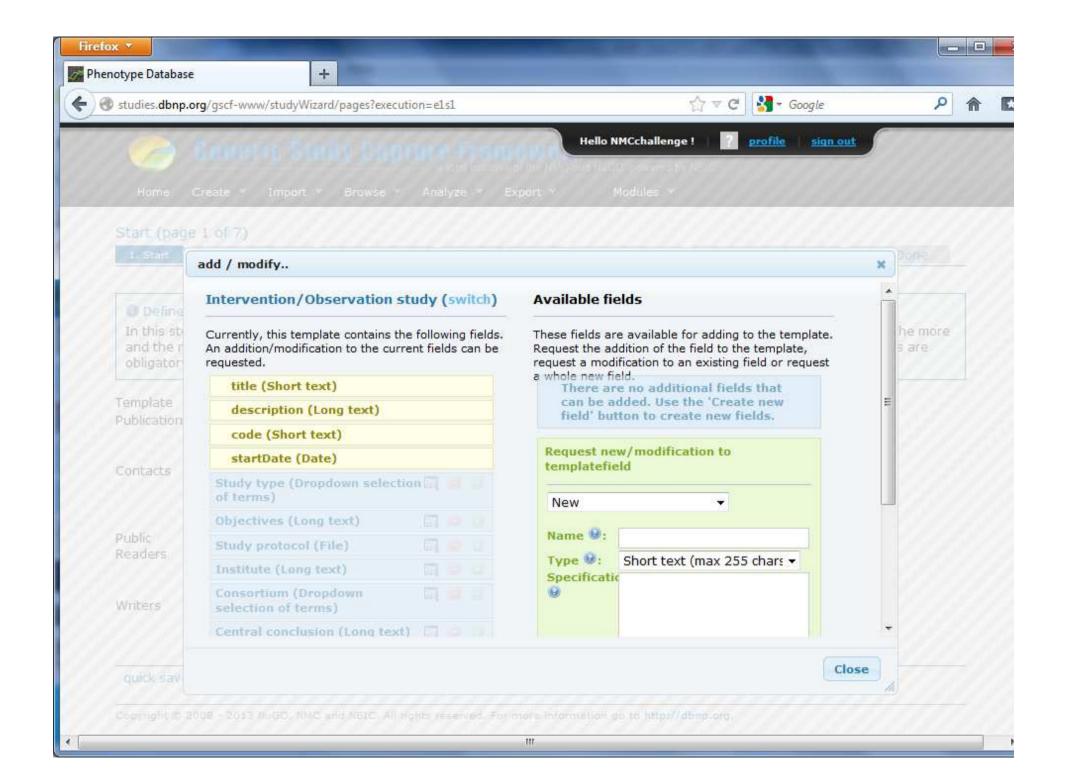


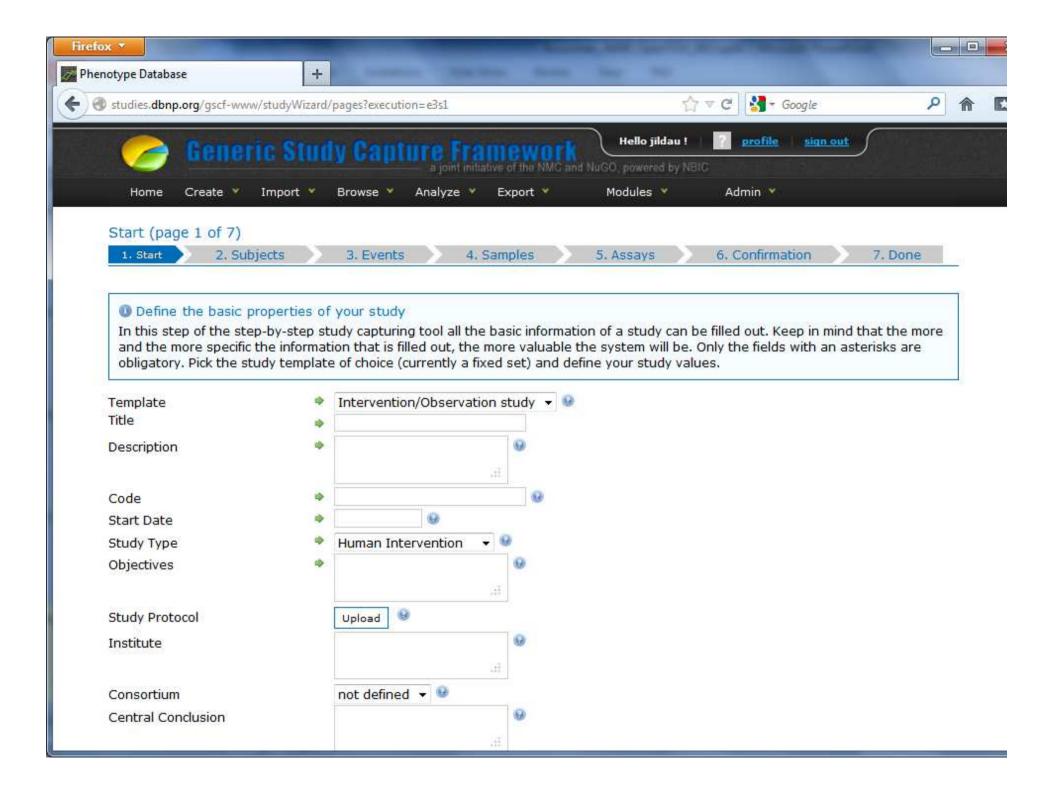


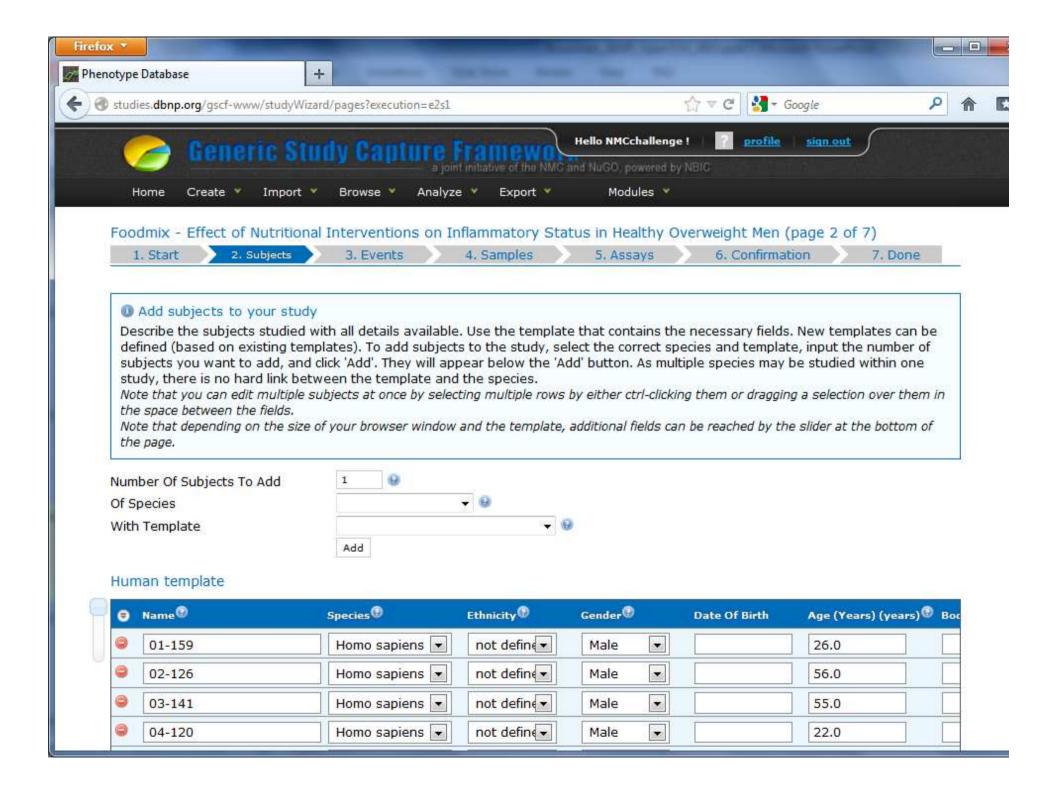










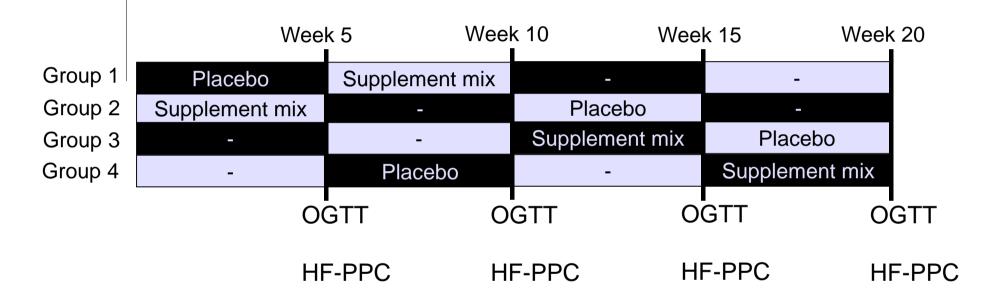


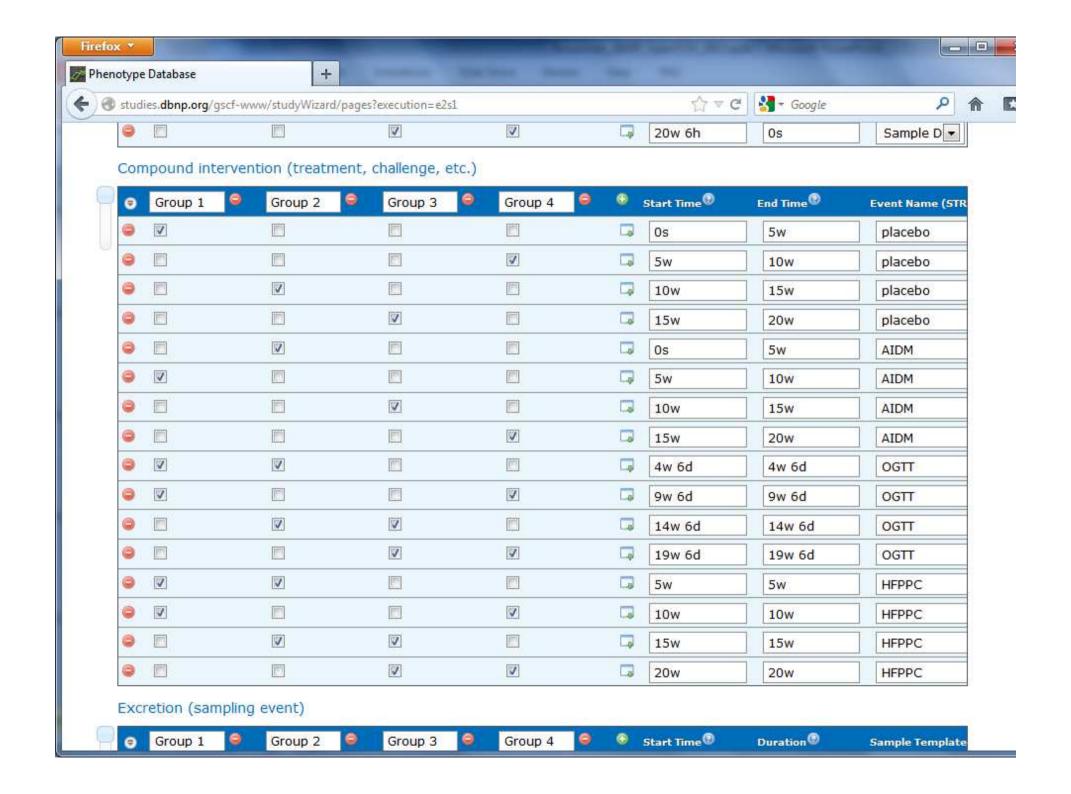






### **Example study**



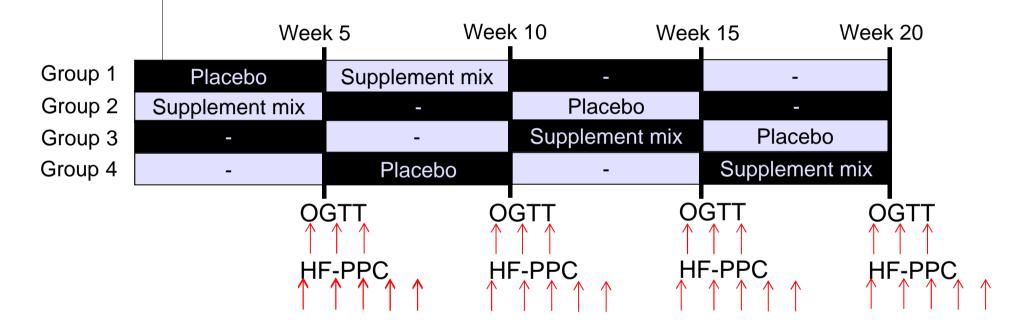


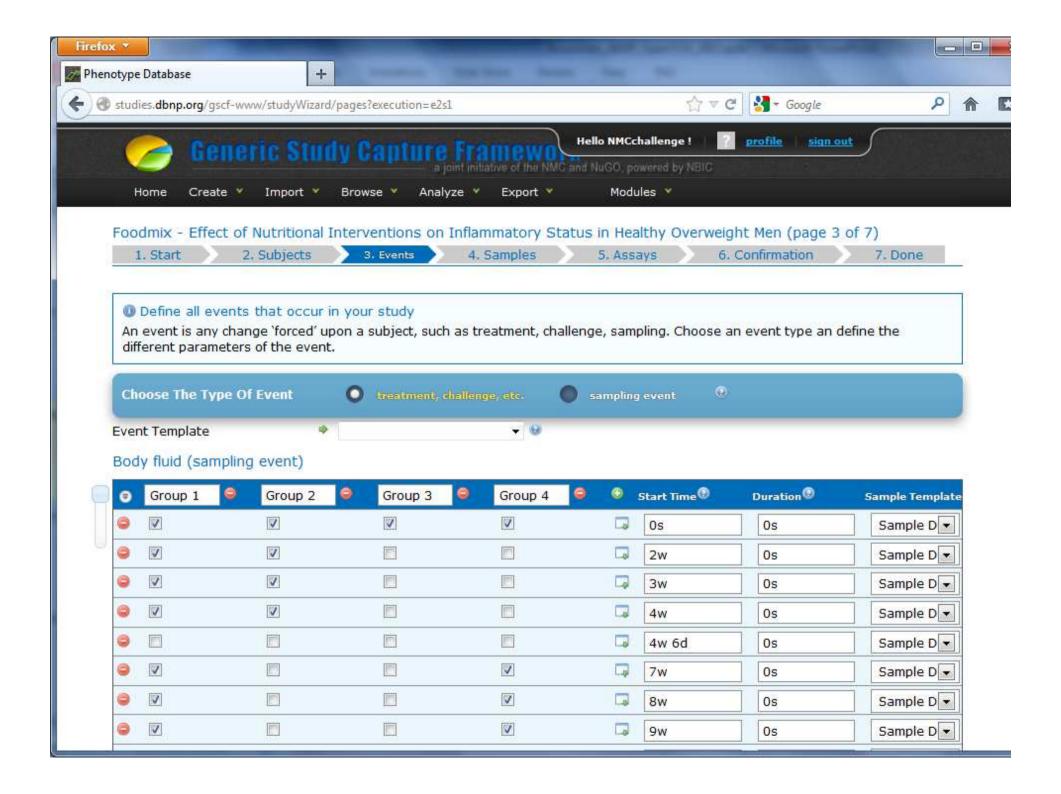






### **Example study**



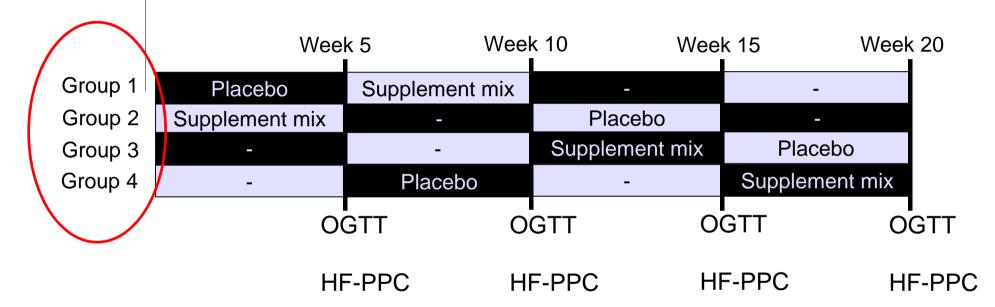


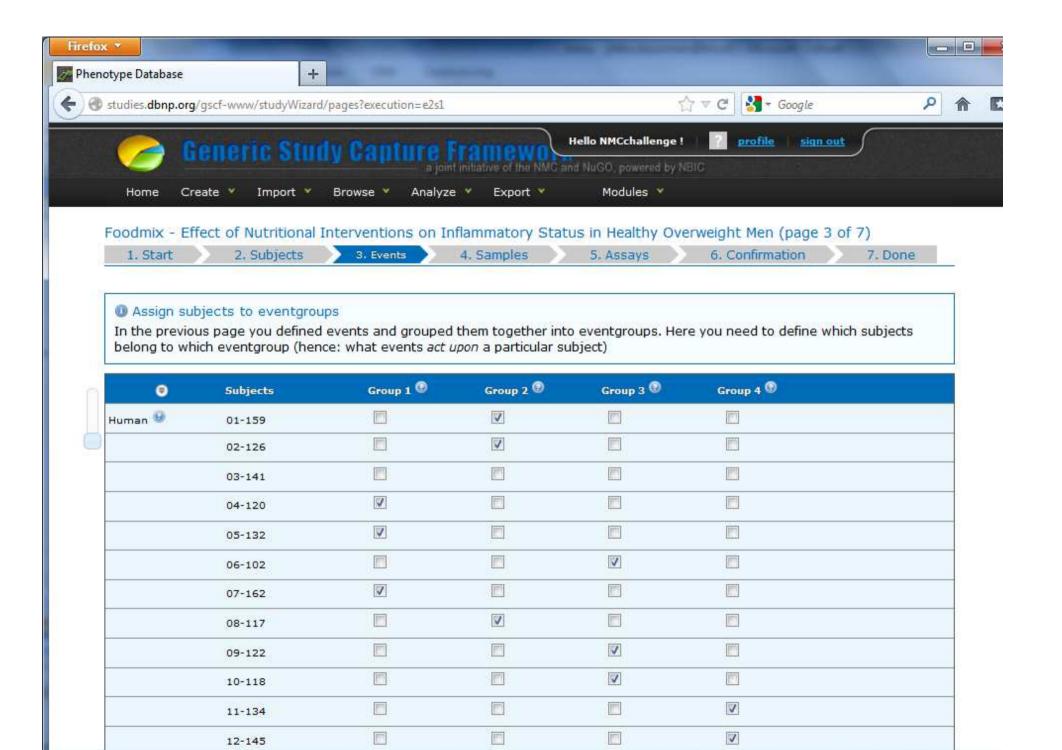






### **Example study**



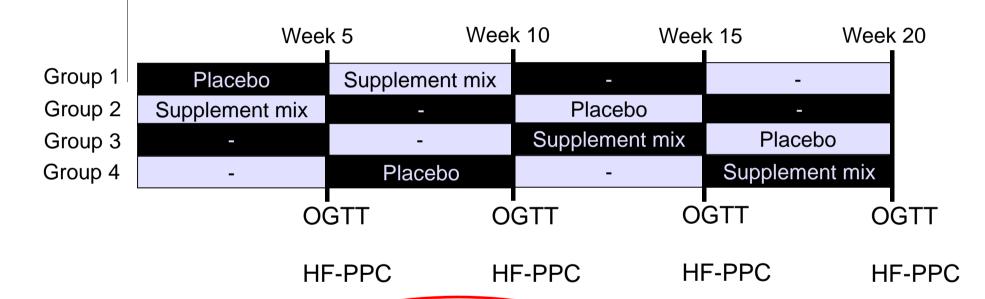






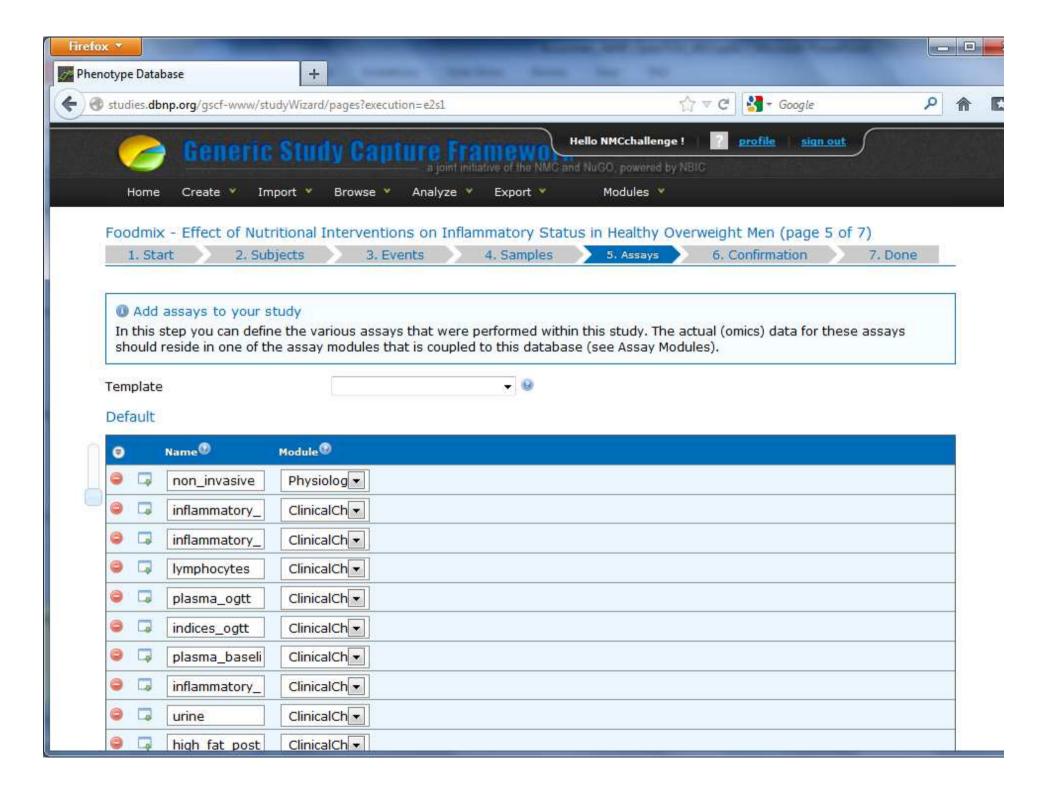


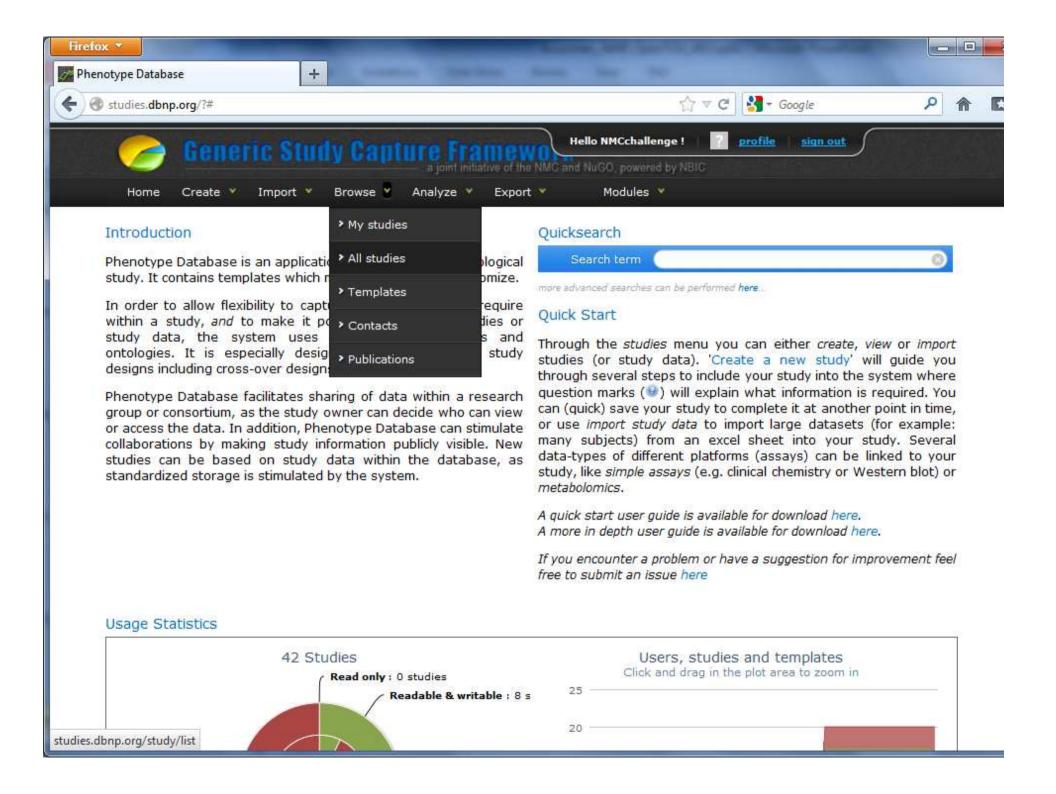
### **Example study**

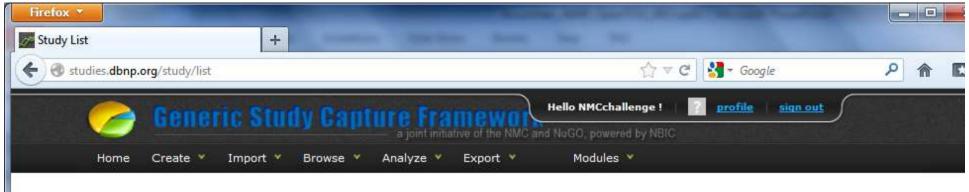


### Plasma measurements:

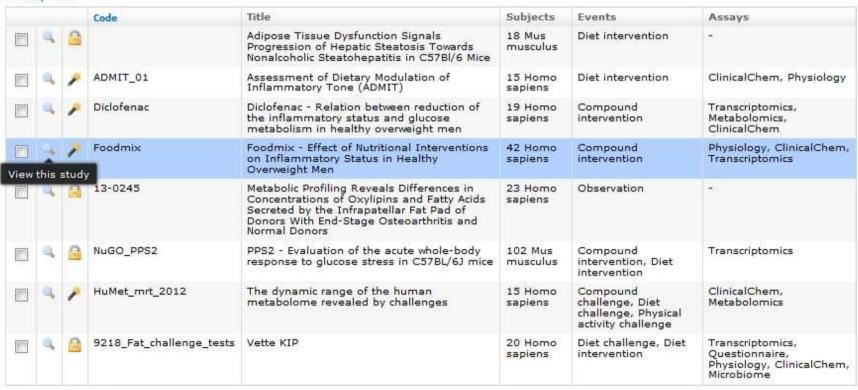
Clinical Chemistry
Lipids LC-MS (TG, SPM, LPC, PC)
Free Fatty acids LC-MS
GCMS
Multiplex protein profiling (RBM)







#### Study List

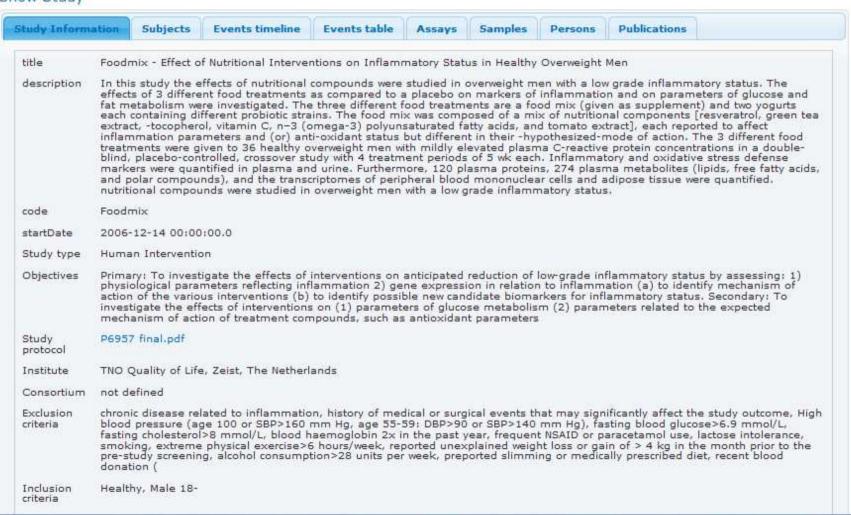


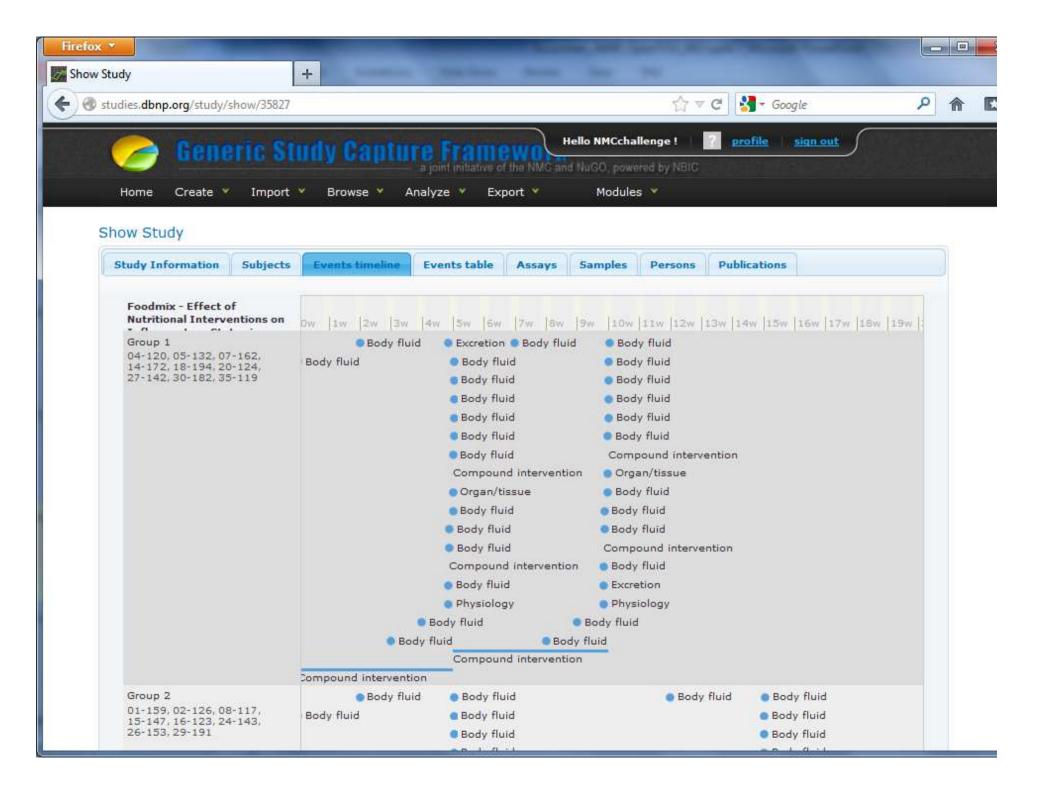
New Study Compare selected studies

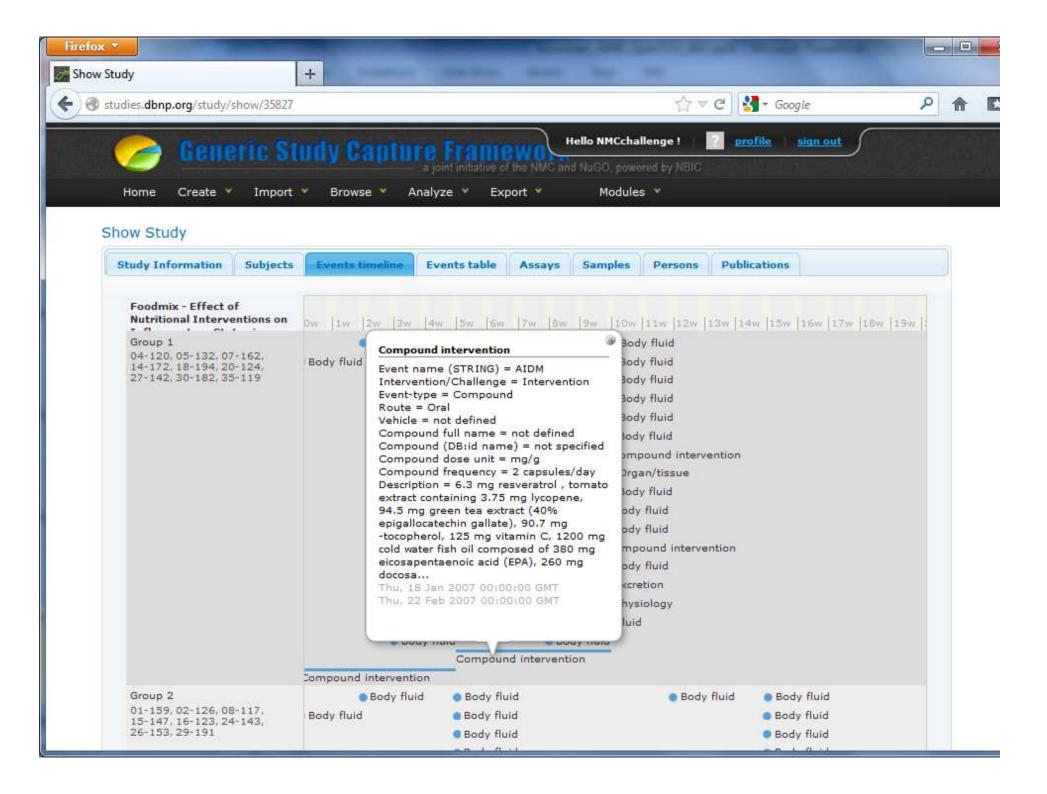
Copyright @ 2008 - 2013 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.

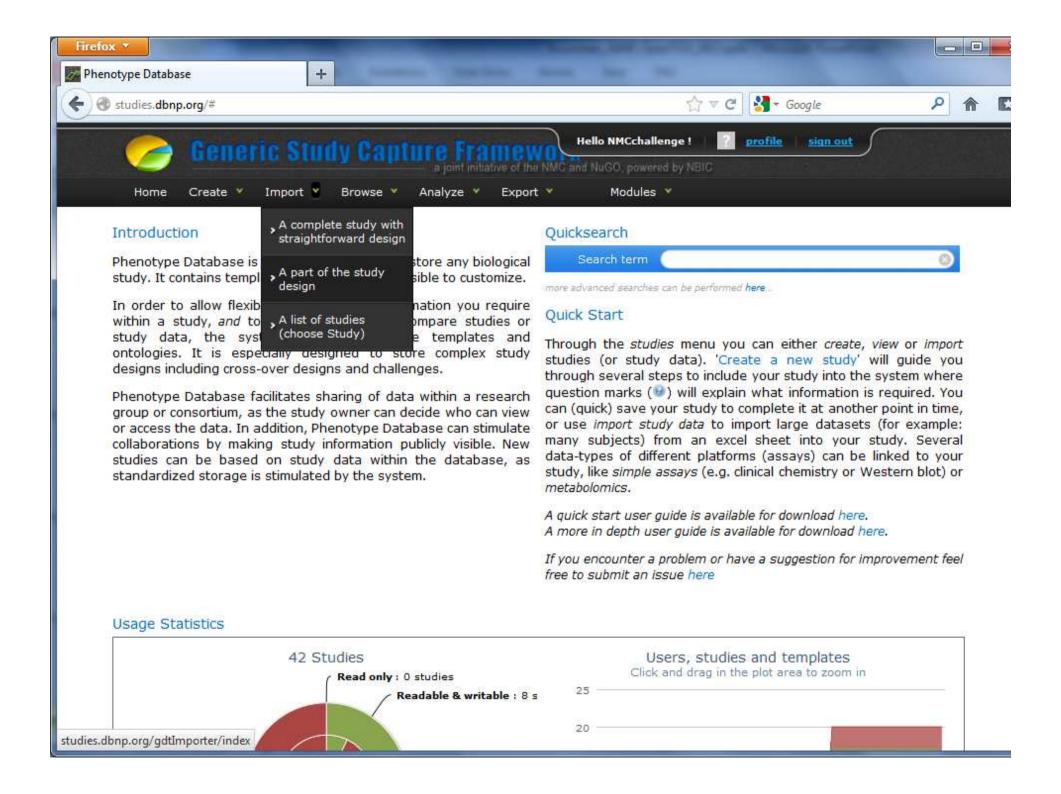


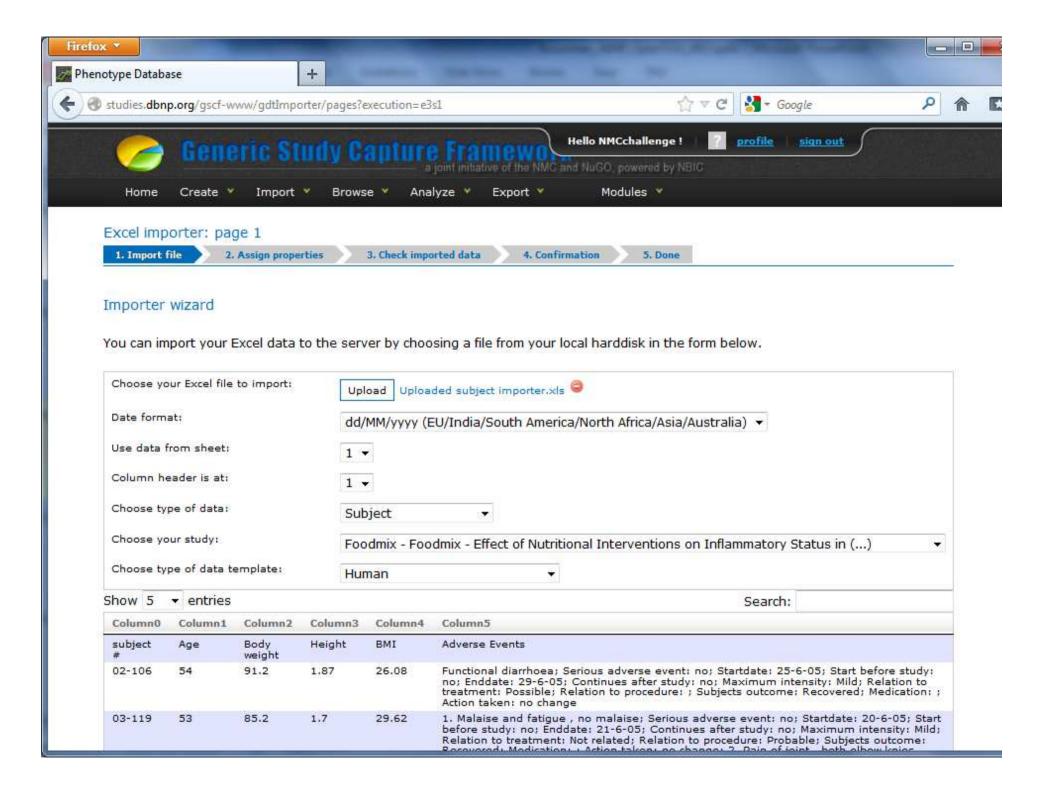
#### Show Study

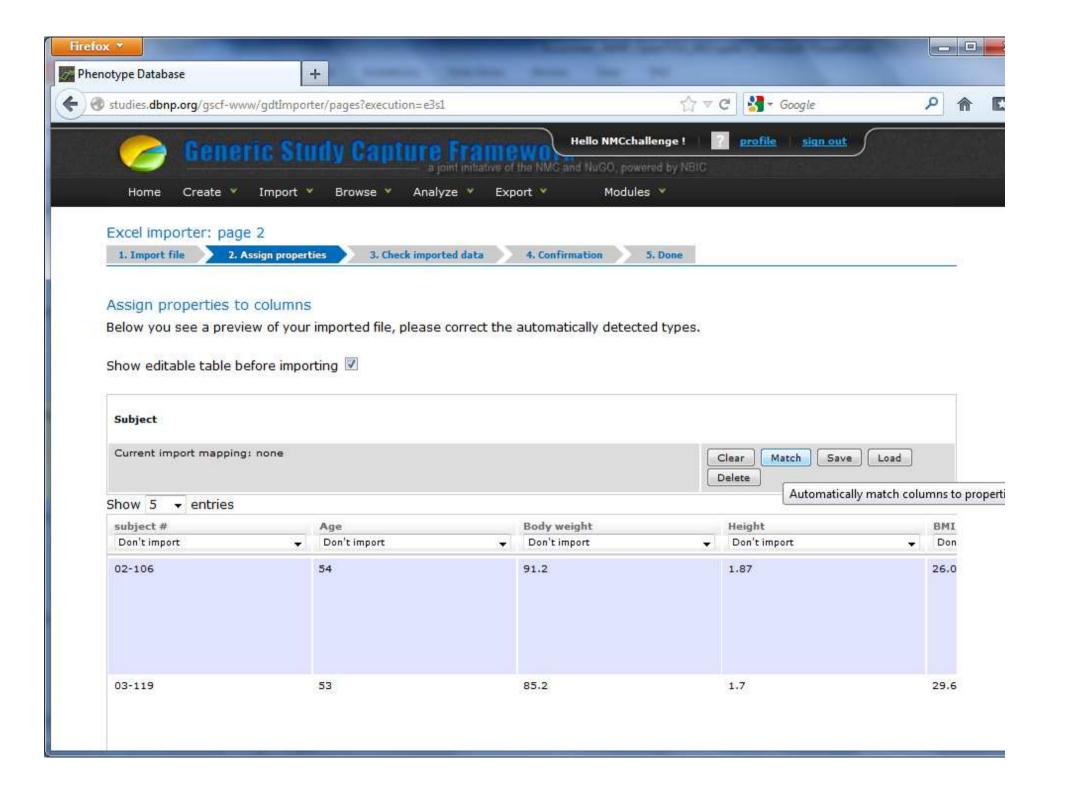


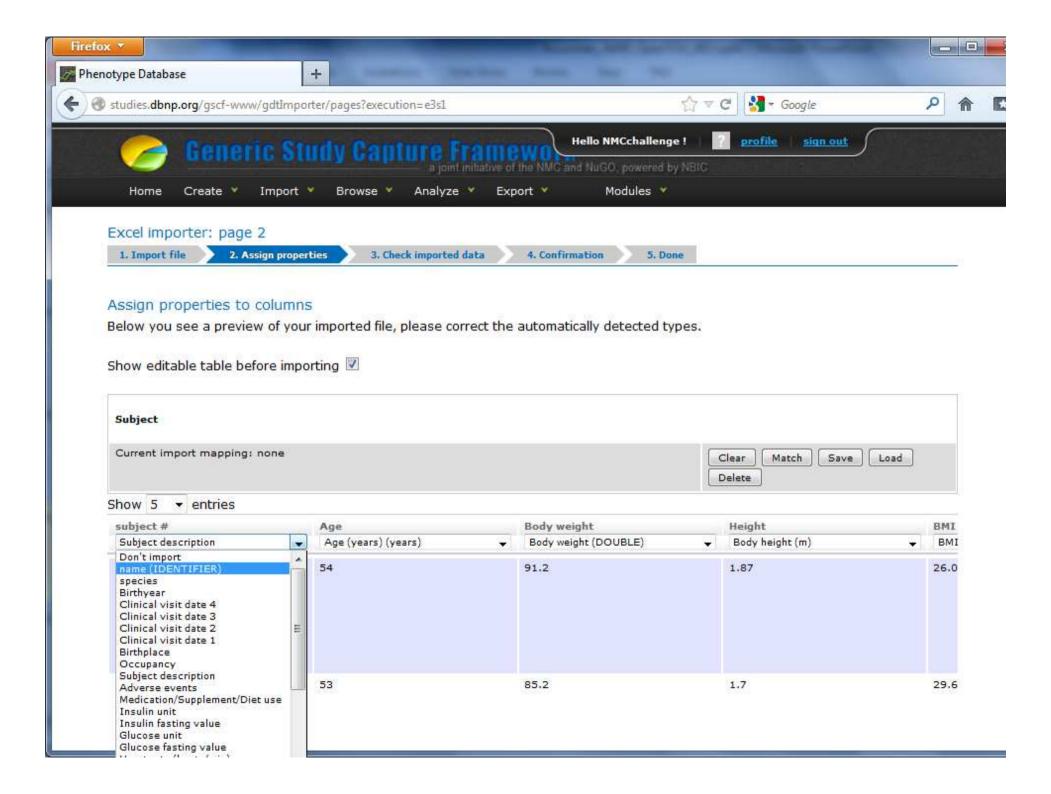


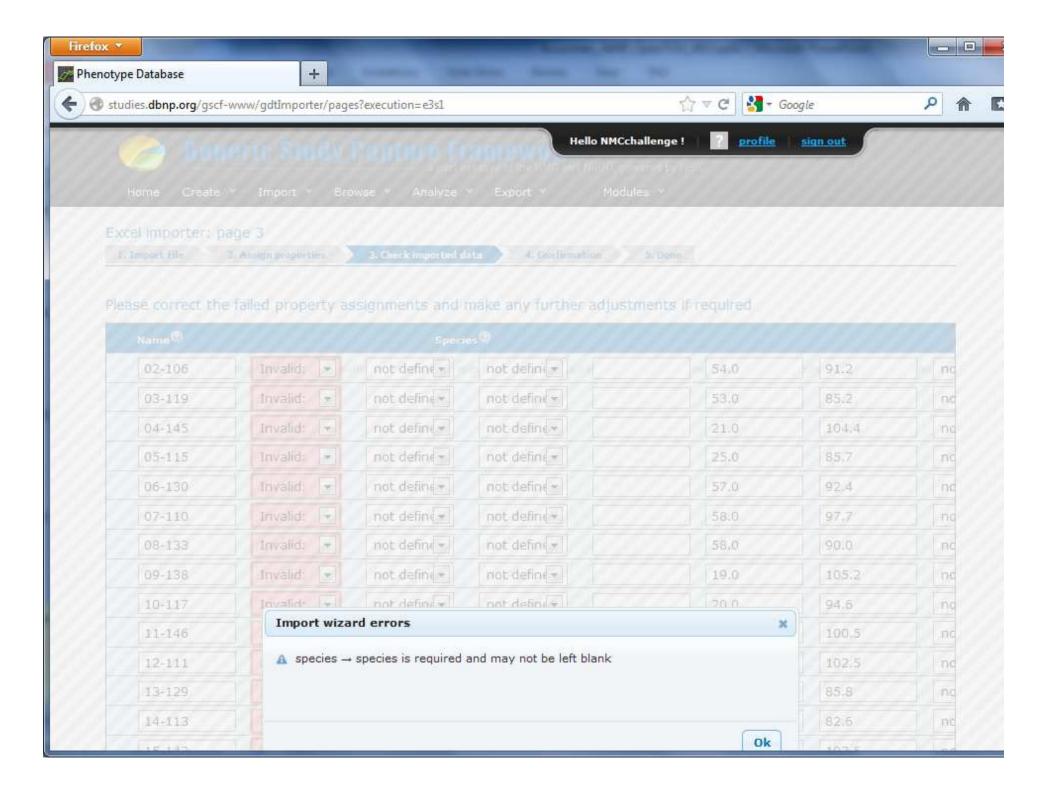


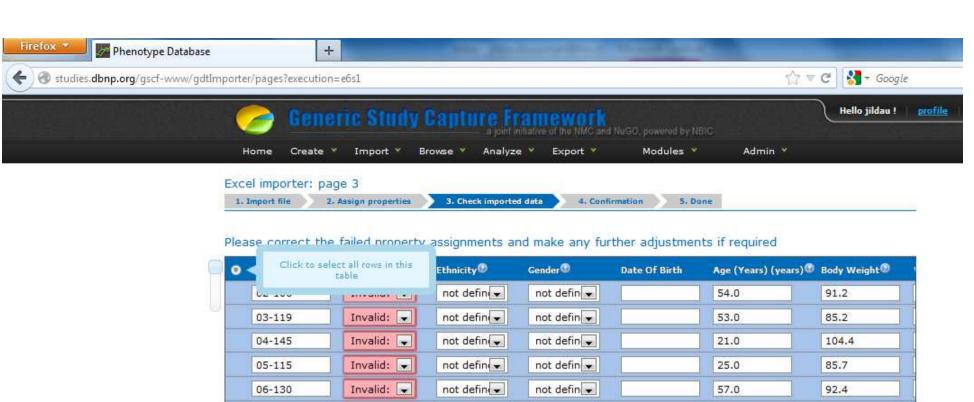




















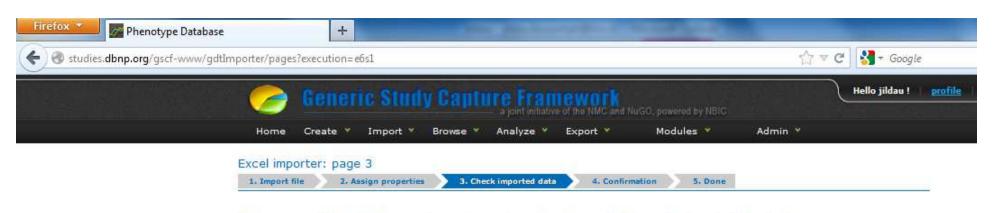












### Please correct the failed property assignments and make any further adjustments if required

Homo salvania Homo sapiens Mus musculus Rattus norvegicu Invalid: add more		not defin ▼  not defin ▼		54.0	91.2
Mus musculus Rattus norvegicu Invalid: add more	defin 🐷			53.0	85.2
Rattus norvegicu Invalid:		not defin			L Inches
add more	H.C.			21.0	104.4
. 1	defin <b>→</b>	not defin		25.0	85.7
0 Homo sa ▼	not defin	not defin		57.0	92.4
0 Homo sa ▼	not defin	not defin		58.0	97.7
3 Homo sa ▼	not defin(▼	not defin		58.0	90.0
8 Homo sa ▼	not defin	not defin		19.0	105.2
7 Homo sa ▼	not defin	not defin		20.0	94.6
6 Homo sa ▼	not defin	not defin		42.0	100.5
1 Homo sa ▼	not defin ₩	not defin		35.0	102.5
9 Homo sa ▼	not defin	not defin		53.0	85.8
3 Homo sa ₩	not defin	not defin		50.0	82.6
2 Homo sa ▼	not defin ₩	not defin		41.0	103.5
4 Homo sa ▼	not define	not defin		23.0	86.6
4 Homo sa ▼	not defin ▼	not defin		56.0	90.6
4 Homo sa ▼	not defin	not defin		60.0	81.1
0 Homo sa 🔻	not defin	not defin		40.0	92.3
2 2 4	Homo sa V	Homo sa not defin homo sa not	Homo sa not defin not defi	Homo sa not defin not defi	Homo sa not defin not defin 20.0  Homo sa not defin not defin 20.0  Homo sa not defin not defin 35.0  Homo sa not defin not defin 53.0  Homo sa not defin not defin not defin 250.0  Homo sa not defin not defin 250.0







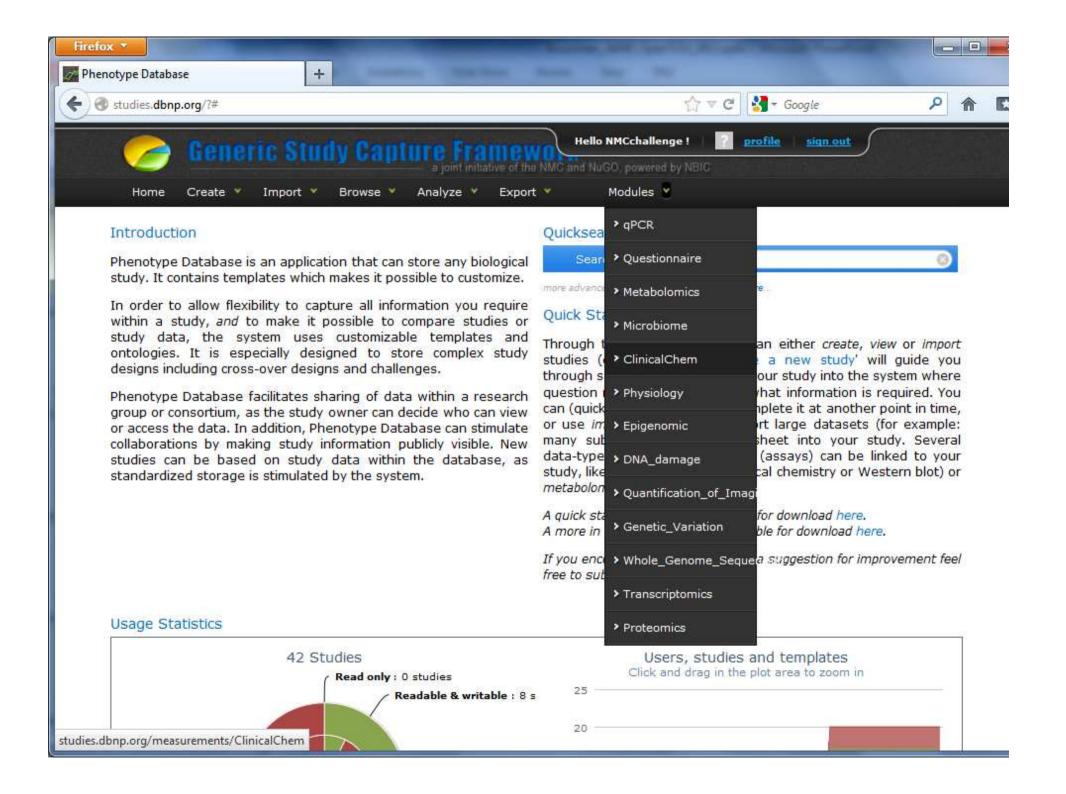














#### Introduction ClinicalChem

Here you can find the measurements for ClinicalChem.

## How do I start using this ClinicalChem module?

- 1. Browse platforms te see whether the platform of your choice already exists.
- 2. You can add features using the import function.
- 3. Then add measurements using the measurement importer.

#### What is a ClinicalChem feature?

A feature is a substance (or measurementtype / detector etc.) that is measured when a sample is taken.

For example, glucose, leptin, CD40, bodyweight, urea

### What is a ClinicalChem platform?

A platform is a technique of a certain type of data.

For example, for Transcriptomics you will have platforms such as Affymetrix & Illumina.

#### What is a measurement?

A measurement is a single value that belongs to a sample. For example, glucose level is measured for a group of samples.

## What is an assay?

An assay is a group of samples that are analyzed.

For example to obtain, clinical chemistry -, metabolomics -, or transcriptomics data

## Frequently Asked Questions

- I want to import features, give me an example.
- What kind of Excel sheets can I upload?
- Where can I edit my feature?
- Where can I edit my assay?
- I don't see my samples, what can I do?



A platform is a technique of a certain type of data.

For example, for Transcriptomics you will have platforms such as Affymetrix & Illumina.

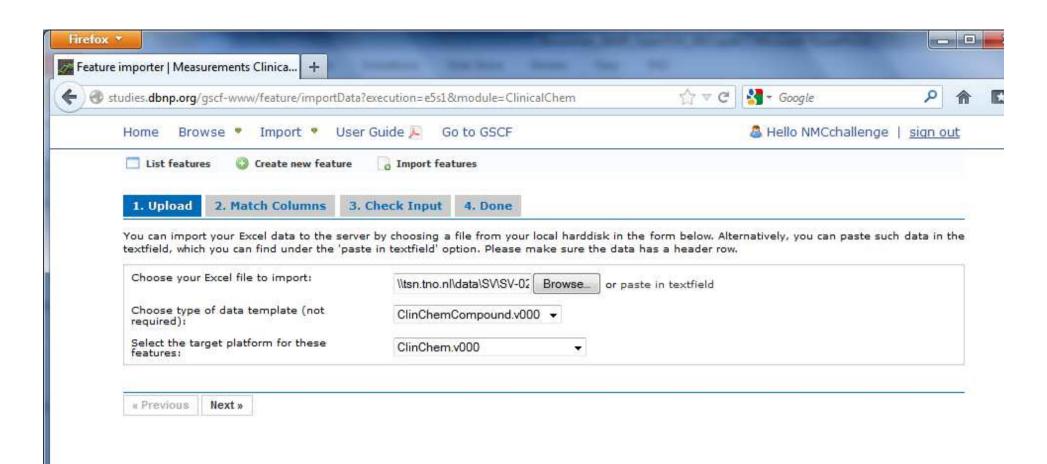
#### What is a measurement?

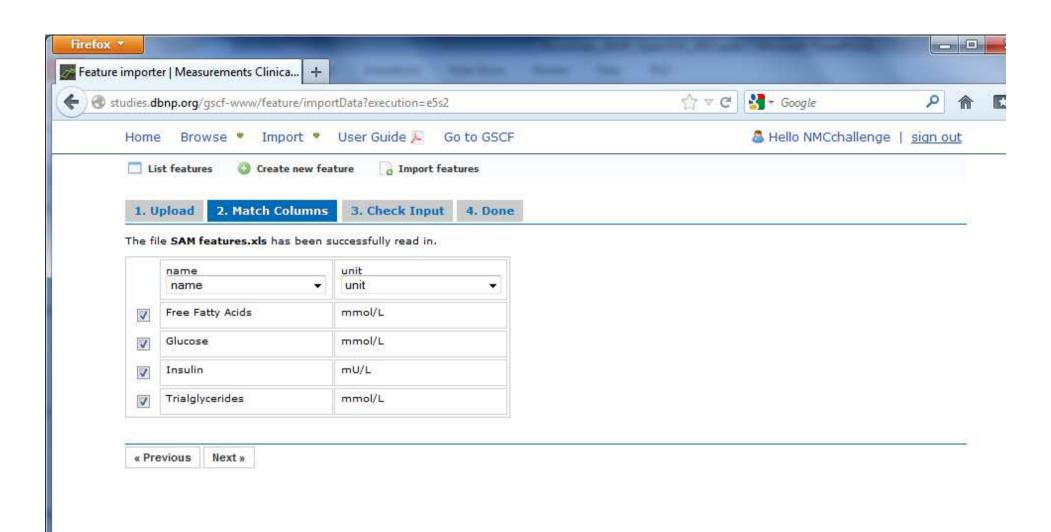
A measurement is a single value that belongs to a sample. For example, glucose level is measured for a group of samples.

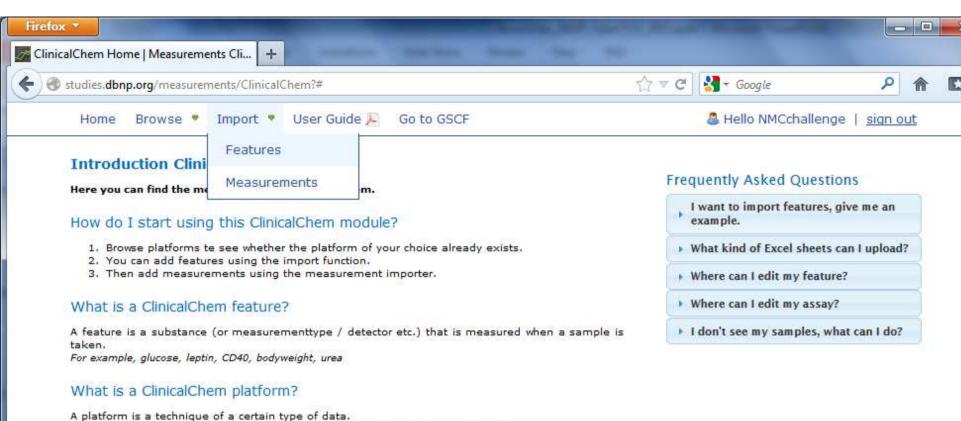
## What is an assay?

An assay is a group of samples that are analyzed.

For example to obtain, clinical chemistry -, metabolomics -, or transcriptomics data







For example, for Transcriptomics you will have platforms such as Affymetrix & Illumina.

## What is a measurement?

A measurement is a single value that belongs to a sample. For example, glucose level is measured for a group of samples.

## What is an assay?

An assay is a group of samples that are analyzed. For example to obtain, clinical chemistry -, metabolomics -, or transcriptomics data

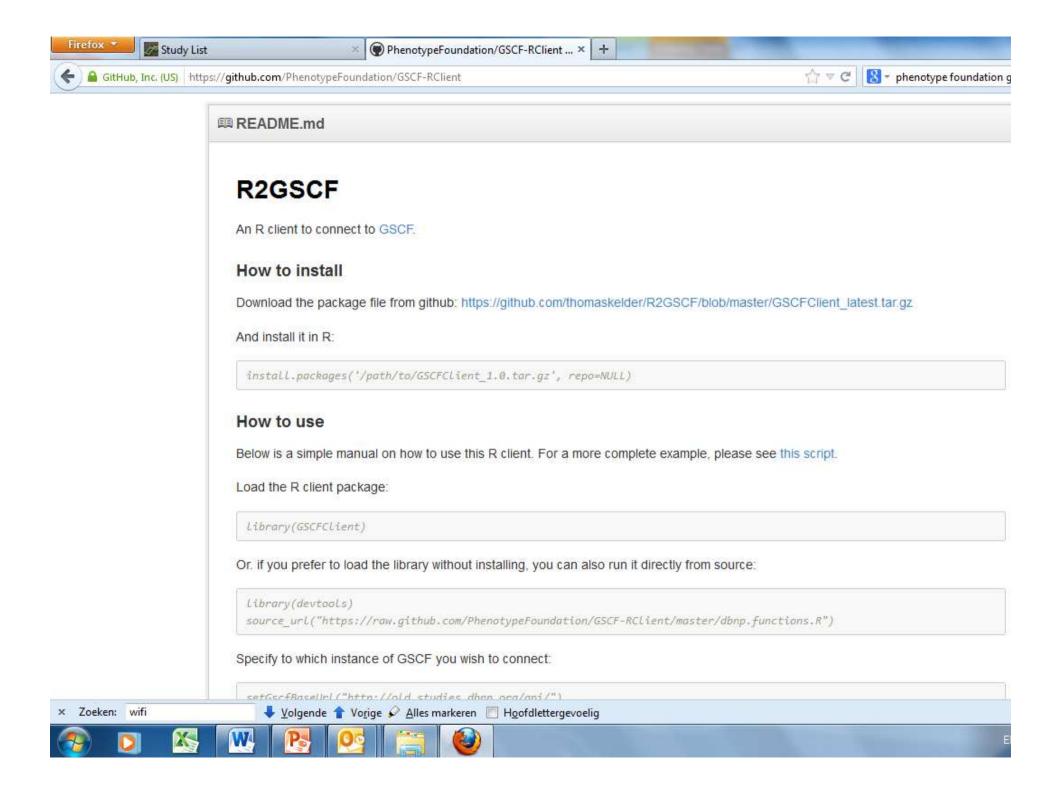






# **Output possibilities**

- For study comparison API (links to software that can analyse your data, e.g. R)
- Search option
- Excel export
- Visualisation (Pathvisio, simple graphs)







## **Technical remarks**

- Code can be found on github:
  - https://github.com/PhenotypeFoundation
- Studies.dbnp.org is the nutritional instance, but can be used by others after request (located at TNO)
- Links to other data sources possible (e.g. arrayexpress)







# Phenotype database goals

- Collect, harmonize and distribute data
- Already available: a database for mechanistic intervention studies
- Helpdesk function/ training for new users (NuGO)
- Example on how Phenotype Database works: analysis of the challenge response in several studies

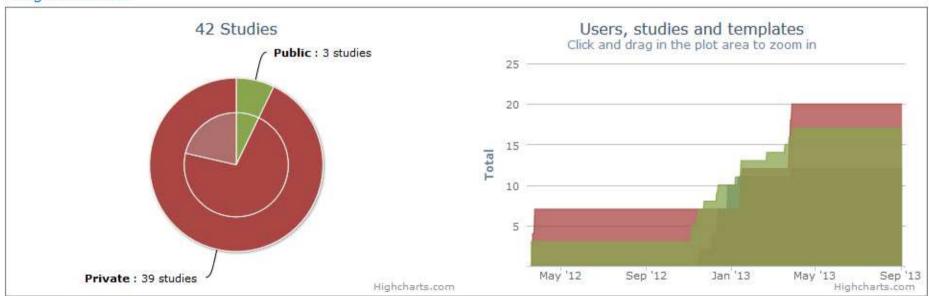






# **Current status of database**

## **Usage Statistics**









# Phenotype database goals

- Collect, harmonize and distribute data
- Already available: a database for mechanistic intervention studies
- Helpdesk function/ training for new users (NuGO)
- Example on how Phenotype Database works: analysis of the challenge response in several studies







# **Pre-training Preparations**

- Account generation creation of username/passwords
- Distribution of Phenotype Database data-uploading template (.xls file) and Quickguide to participants
- Submission of data files (even partial) at least 1 week prior to training (Participant)
- > Brief presentation of the current datasets potentially available to share, for discussion purposes during training (Participant)







# **Training**

## Format (1 ½ days)

## <u>Afternoon</u>

- Introduction to the Phenotype Database software
- Demonstration of the live site (<a href="http://studies.dbnp.org">http://studies.dbnp.org</a>)
- Discussion of the datasets to be shared in dbNP (participants)

## Next day

- Uploading of files (even partial/dummy) to assess compatibility
- Practical work and recording of potential errors/problems
- Setting of date to share full files with TNO if not already done







# Phenotype database goals

- Collect, harmonize and distribute data
- Already available: a database for mechanistic intervention studies
- Helpdesk function/ training for new users (NuGO)
- Example on how Phenotype Database works: analysis of the challenge response in several studies







## Huber et al., BMJ 2011:

## How should we define health?

The WHO definition of health as complete wellbeing is no longer fit for purpose given the rise of chronic disease. **Machteld Huber and colleagues** propose changing the emphasis towards the ability to adapt and self manage in the face of social, physical, and emotional challenges

Machteld Huber senior researcher<sup>1</sup>, J André Knottnerus president, Scientific Council for Government Policy <sup>2</sup>, Lawrence Green editor in chief, Oxford Bibliographies Online—public health<sup>3</sup>, Henriëtte van der Horst head <sup>4</sup>, Alejandro R Jadad professor<sup>5</sup>, Daan Kromhout vice president, Health Council of the Netherlands<sup>6</sup>, Brian Leonard professor<sup>7</sup>, Kate Lorig professor<sup>8</sup>, Maria Isabel Loureiro coordinator for health promotion and protection<sup>9</sup>, Jos W M van der Meer professor<sup>10</sup>, Paul Schnabel director <sup>11</sup>, Richard Smith director<sup>12</sup>, Chris van Weel head <sup>13</sup>, Henk Smid director <sup>14</sup>



The ability to adapt

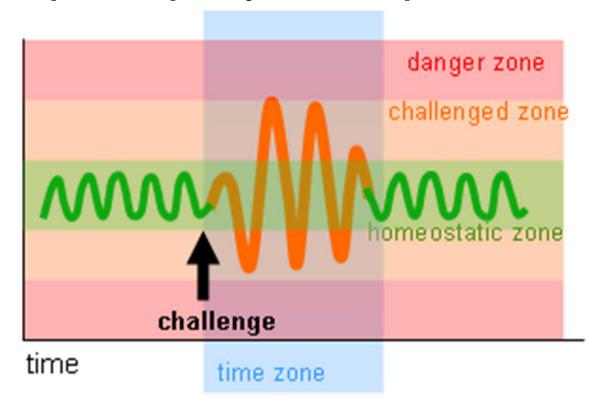






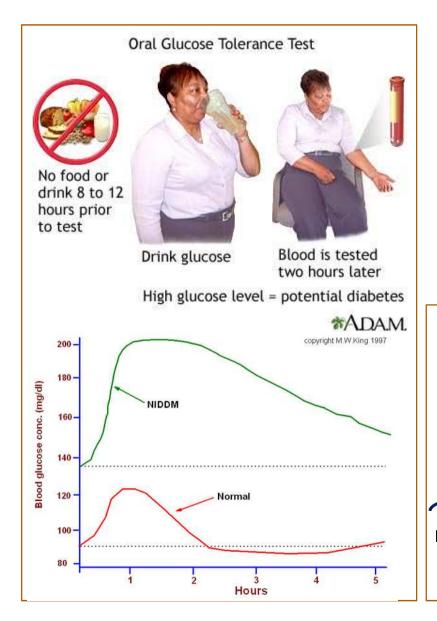


## the adaptive capacity – the response to stressors

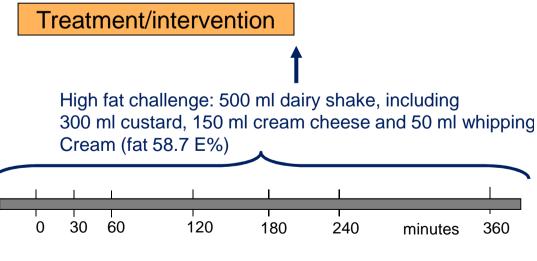


- Quantification of the effect of a treatment/intervention
- Diagnose a disease

# **Example challenges**













# Defining clinical (pre)diabetic subgroups

Description	Fasting glucose (mmol/l)	Glucose 120 min after OGTT (mmol/l)
Diabetes (T2DM)	7,0	> 11.0
IGT (Impaired glucose tolerance)	< 5.6 (normal)	7.8 – 11.0
IFG (impaired fasting glucose)	5.6 – 6.9	< 7.8
IGT&IFG	5.6 – 6.9	7.8 – 11.0
Normal	< 5.6	< 7.8







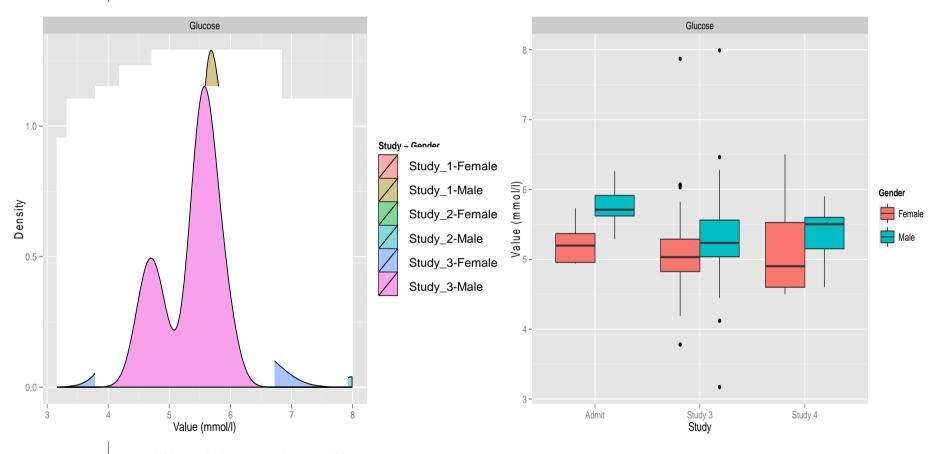
# 'The metabolic challenge test knowledge database' project goal

- > To find mechanistic explanations for differences in clinical subgroups
- Providing an integrated study evaluation platform (challenge studies uploaded)
- Demonstrate that other research questions can be answered which could not be answered by a single study





# **Explorative analysis (glucose): Male and Female**



All subjects, baseline measurements

Conclusion: An overlap in the baseline levels between the studies is observed, a further study comparison can be performed





# Complex data analysis (linear mix model and looking at covariates): Are the data suitable for virtual cohort?

	Base model: value = Study_1 * time0 + healthy + BMI : Normal						
		Value	Std.Error	DF	t-value	p-value	
Ctudy 1	study_2	-0.06	0.42	156	-0.14	0.89	
Study_1	study_3	-0.77	0.57	156	-1.35	0.18	
Time: 0	Time: 30	2.14	0.60	470	3.57	0.00	**
	Time: 60	1.98	0.54	470	3.67	0.00	**
	Time: 120	-0.19	0.44	470	-0.44	0.66	
Healthy	Diabetes	3.04	0.60	156	5.09	0.00	**
	Prediabetes	1.21	0.21	156	5.71	0.00	**
	BMI: Morbid obese 1.02 0.49	0.49	156	2.07	0.04	*	
DMI: Normal	BMI: Obese	0.78	0.29	156	2.71	0.01	00 ** 00 ** 04 * 01 *
BMI: Normal	BMI: Overweight	0.22	0.20	156	1.12	0.26	
	BMI: Underweight	-0.23	0.35	156	-0.67	0.51	
Study1:Time30	Study_2:time30	0.15	0.62	470	0.24	0.81	
	Study_3:time30	1.28	0.72	470	1.77	0.08	#
Ctudy4.TimeCO	Study_2:time60	-1.08	0.56	470	-1.94	0.05	#
Study1:Time60	Studv 3:time60	1.64	0.67	470	2.44	0.02	*

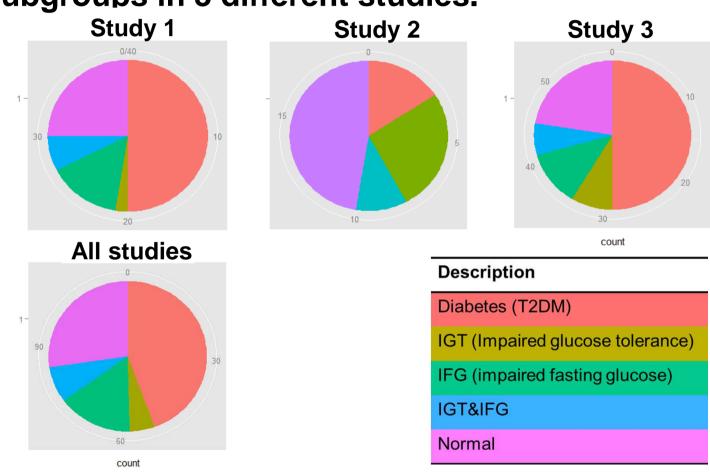
The complex data analysis shows that we can find health related differences in response to OGTT (based on the single parameter glucose) by combining studies







# Analysis of the distribution of (pre)diabetic subgroups in 3 different studies.



The 3 studies represent all 5 (pre)diabetic subgroups and therefore can be used to answer our biological question





# **Upcoming work**

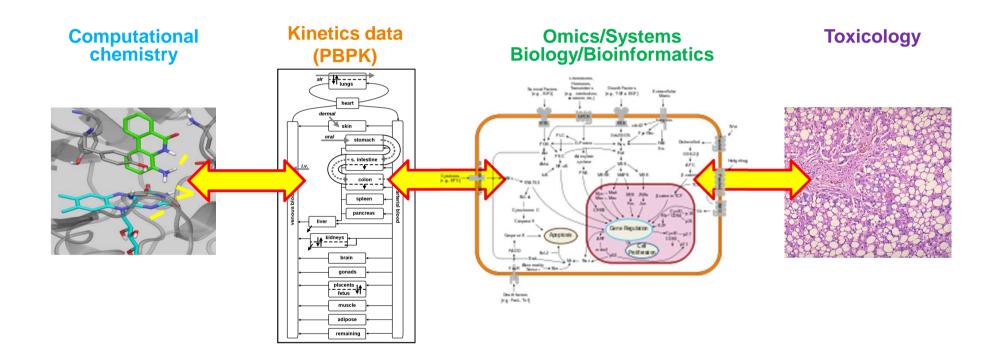
- Meta-analysis of challenges
  - On metabolomics data
  - Inclusion of more studies



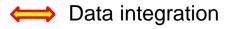




# **DIAMONDS** approach



Data Infrastructure for Applying Models ON Design and Safety



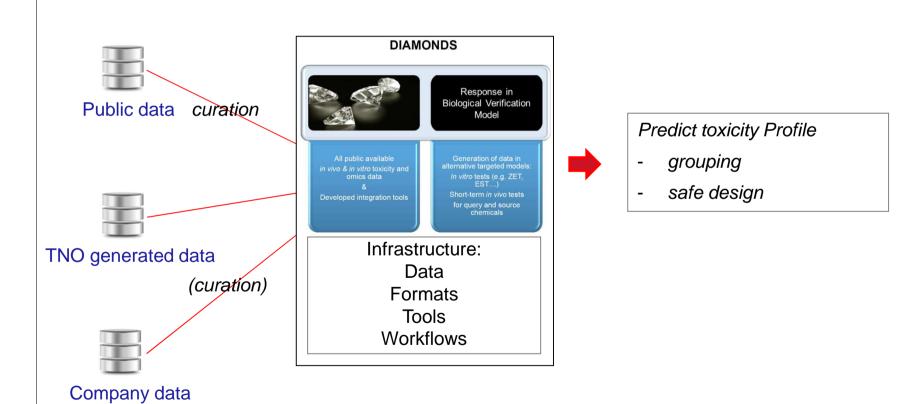






## AIM:

# Better predict toxicity profile by integrating structural and/or omics information









# **Embedding and continuation**

- Cosmos (standardization of metabolomics infrastructure): NMC involved
- > EURODISH (aiming at establishing the European Nutrition Research Institute)
- > NU-AGE
- Nutritech
- > Bioclaims
- Usage of studies and data in the Dutch-Irish collaboration on nutritional research
- JPI "Healthy Diet for Healthy Life"







# **Acknowledgements**

Jeroen Wesbeek Michael van Vliet Kees van Bochove

Tjeerd Abma

Jahn Saito

Adem Bilican

Ferry Jagers

Robert Ernst

Robert Horlings

Carina de Jong-Rubingh

Ruud Boessen

Margriet Hendriks

Theo Reijmers

Lars Eijssen

Chris Evelo

Ben van Ommen

Suzan Wopereis

Eugene van Someren

Rob Stierum

Miriam Ryan

Lorraine Brennan

...and many others









netherlands bioinformatics

## See also:

www.dbnp.org

test.dbnp.org (user=user password=useR123!)







# **Nutritional Researchers Cohort Registration**

When using the Nutritional Researchers Cohort, you can help nutritional research by sharing your data. The data can be used anonymously in different studies.

You may also choose to use the Nutritional Researchers Cohort personally, without sharing your data. That way, no researchers will have access to your data.

I understand, and want to share data

I only want to use my data personally

http://nrc.dbnp.org

#### **GET INFORMED**

- Welcome
- Introduction
- · Researcher Terms of Use
- Grant Rights
- Watch Video
- Checkpoint
- Acknowledge Understanding

#### REGISTRATION

- Consent Form
- User Profile
- Confirm registration

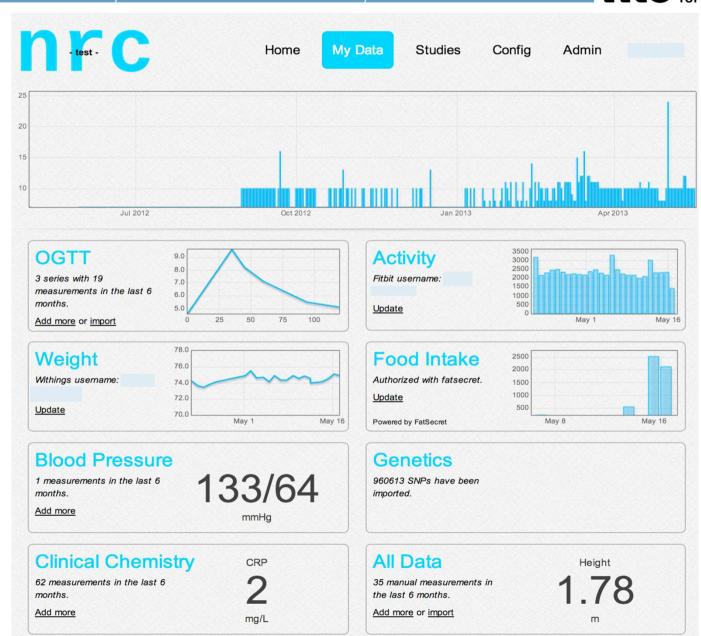
## ADD YOUR DATA

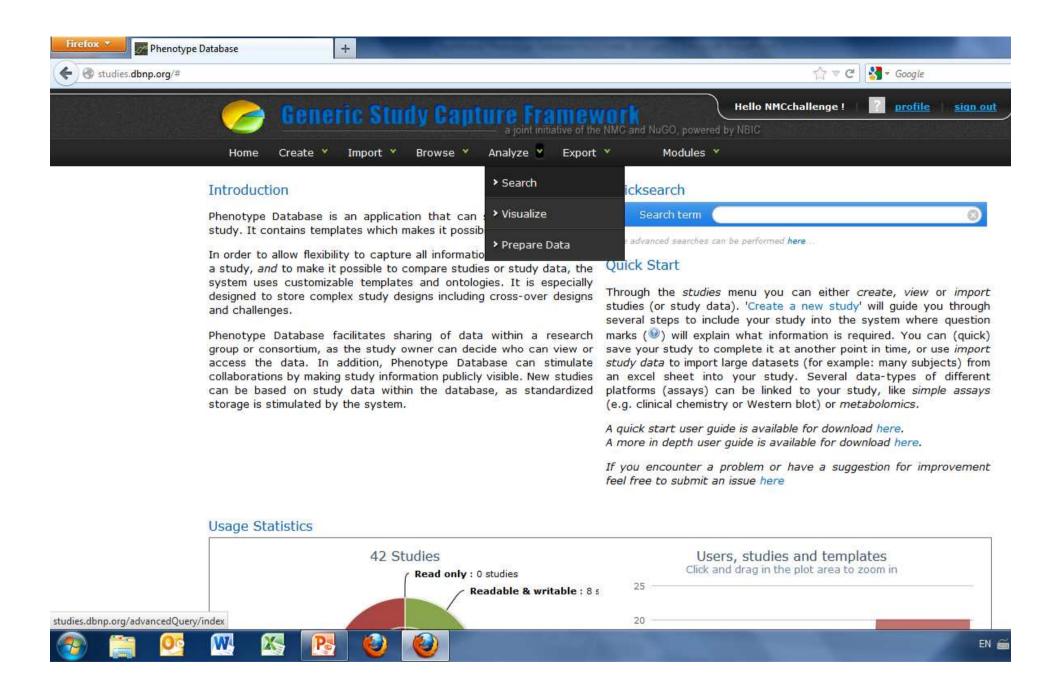
- · Sign In
- Upload Data

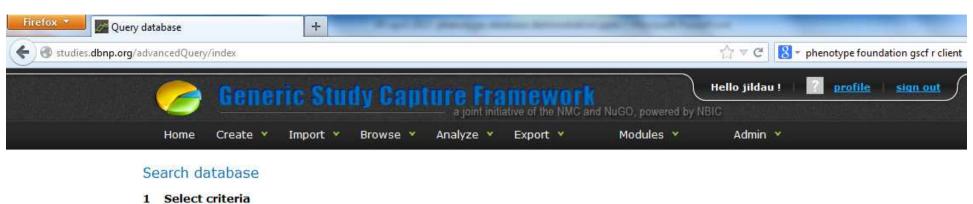












N.B. Comparing numerical values is done without taking into account the units. E.g. a weight of 1 kg equals 1 grams.



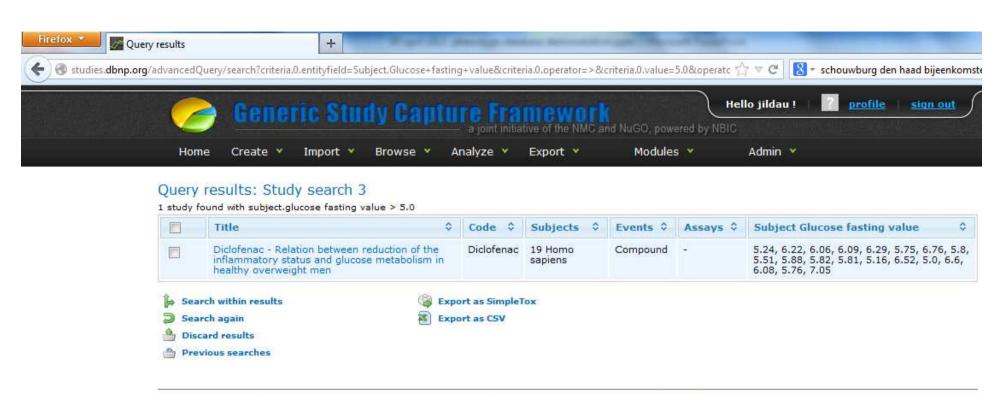
Run query

Search

Previous searches

Copyright @ 2008 - 2013 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.





Copyright © 2008 - 2013 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.





### Select the assay you want to export data from

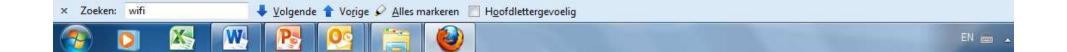
With this exporter you can export (meta) data about samples from an assay to a file. First, select a study from the first list and then select an assay from that study from the second list.

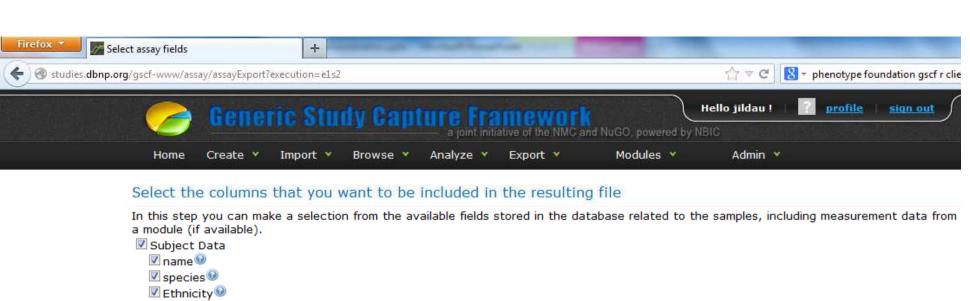
Foodmix - Effect of Nutritional Interventions on Inflammatory Status in Healthy Overweight Men

plasma\_ogtt 

Submit

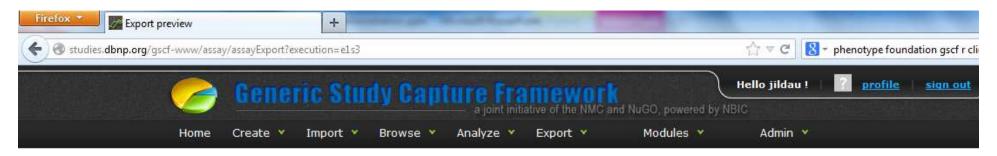
Copyright @ 2008 - 2013 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.











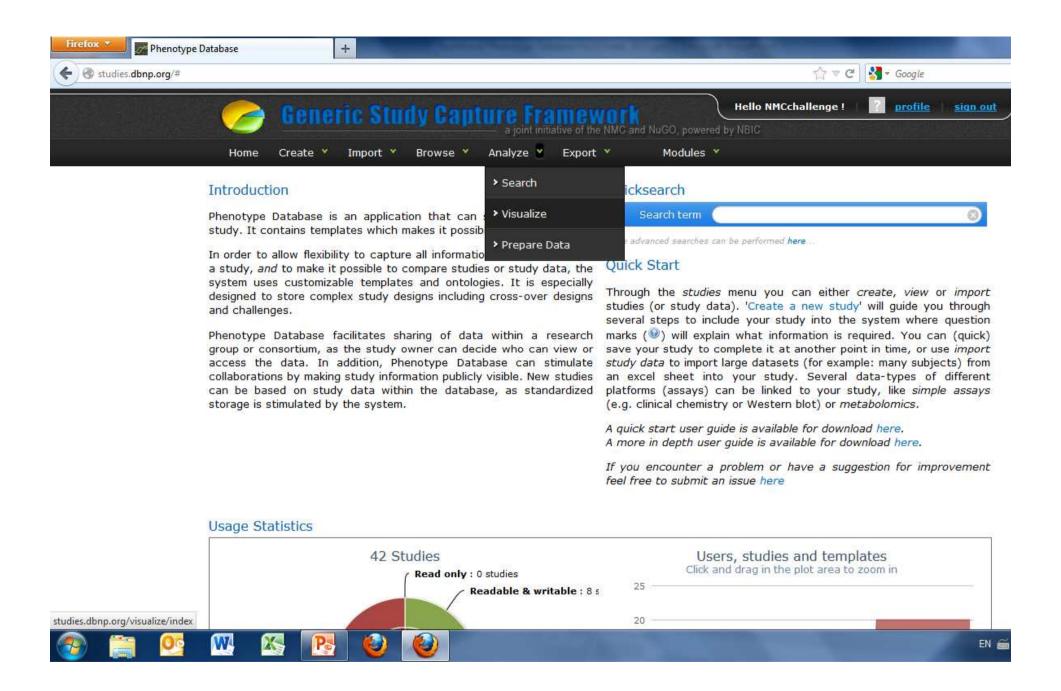
## Below you see a preview of the resulting file, click OK to download

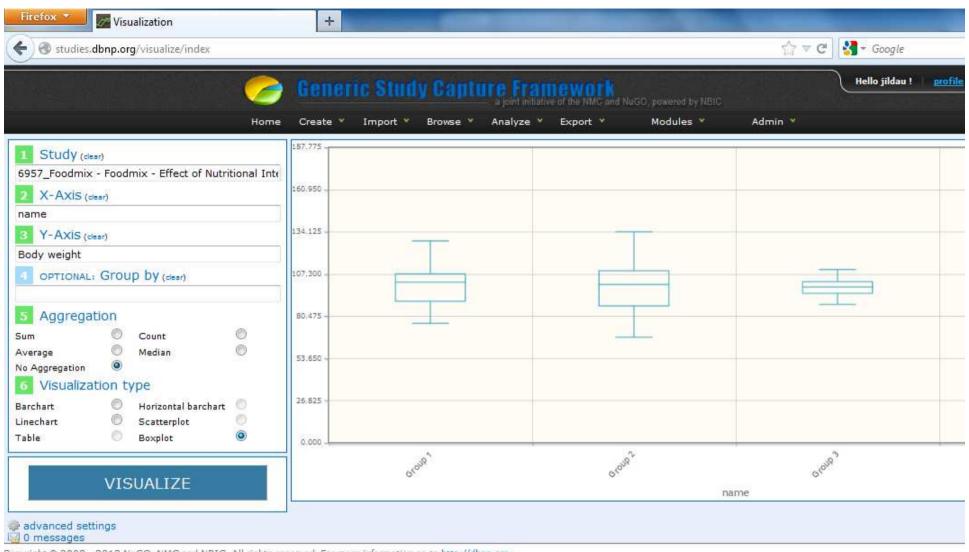
Subject Data					***
name	species	Ethnicity	Gender	Age (years) (years)	
01-159	Homo sapiens	not defined	Male	26.0	
01-159	Homo sapiens	not defined	Male	26.0	
01-159	Homo sapiens	not defined	Male	26.0	
***			***		***

#### OK Cancel

Copyright @ 2008 - 2013 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.







Copyright © 2008 - 2012 NuGO, NMC and NBIC. All rights reserved. For more information go to http://dbnp.org.















