<OpenClinica Import & Export Guide>

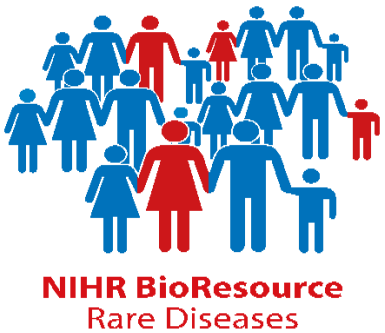
Purpose of this document: This document is a step-by-step guide on how to import data from a comma separated values (csv) file into OpenClinica, and export data from OpenClinica. Knowledge of how to set up a study in OpenClinica and basic CRF design is assumed.

Version: 2.3

Date: June 2016

Authors: Vera Matser, Csaba Halmagyi, Ping Yu

NIHR Rare Diseases Translational Research Collaboration
Department of Haematology, University of Cambridge
National Health Service Blood and Transplant Cambridge
Long Road, Cambridge
CB2 0PT



Document control

Version	Date	Who	Changes
1.0	21/08/2014	Catherine Titterton	First draft
1.1	26/01/2015	Catherine Titterton	Inclusion of section 4 – Advanced Data Import; section 5 – CIT Data Importer
2.0	27/07/2015	Ping Yu	Guidelines to prepare Data for import. New version Data import tool: ODIN. Rewrite Data export from OpenClinica
2.1	18/12/2015	Ping Yu	Add information about coding data for importing.
2.2	01/03/2016	Ping Yu	Add information about RDCIT Secure Data Service importing and a new feature about ODIN
2.3	15/06/2016	Ping Yu	Add a section of ODIN synchronise

Check if there is a later version of this document on the RDCIT website.
(http://rdcit.org/openclinica/oc-training/user-documentation/)

Contents

1 How to use this guide 5

2 Preparing your data for import 6

2.1 ODIN template 6

2.2 Repeating items 6

2.3 Converting Date format to Text 9

2.4 Preparing coding data for importing into OpenClinica 12

3 Importing data 13

3.1 Last preparation before importing the data 13

3.2 ODIN 13

3.3 Mapping Files 23

3.4 Data import to the RDCIT Secure Data Service 24

4 Synchronise studies 24

4.1 Synchronise module 25

4.2 Synchronise Study using ODIN 27

5 Data Export from OpenClinica 33

5.1 Create a Dataset 33

5.1.1 Define the dataset 33

5.1.2 Generate and download Dataset 37

6 Reference: 40

1 How to use this guide

OpenClinica uses an international file format called CDISC ODM XML. In order to import data directly into OpenClinica the data file needs to adhere to this format. This guide explains how to import data from a comma separated values (csv) file using a mapping programme such as ODIN.

There are a number of tools and approaches available for importing data into OpenClinica. Each can be used singly or in parallel, depending on the state of your data and level of expertise or familiarity with the tools. All tools are accessible to the end user and do not require specialist IT technical skills. Tools available include:

- ODIN (developed by the Clinical Infrastructure team, you can download source files from Here.)
- OC Data Importer (developed by https://community.openclinica.com/extension/ocdataimporter)

Which tool should I use?

There is no hard and fast rule which tool will be most appropriate for you use, they can all do the job with the same end results. The key differences between the tools are as follows:

- CIT Data Importer includes an option to skip populating events and fields for which no data has been collected against a particular subject.

- OC DataImporter has good error checking prior to data load into OpenClinica. You can rerun the generation of the XML file over and over again until the report files show no errors. CIT data importer runs the import immediately and only shows errors after import which can then be corrected with a second import (OpenClinica keeps an audit log of changes).

- Both CIT and OC data importers have a manual mapping process. The advanced data mapping process outlined in section **Error! Reference source not found.** can be used with both tools, a lthough the excel formulas are currently set up to create the OC data import mapping file, not the mapping file for CIT.

- TDS Importer is useful for those comfortable working in Access and primarily used for loading subjects and scheduling events, not for loading study data.

To be able to import data into OpenClinica you will need to have setup a study to receive your data import. It is expected that you will already have gone through the process of creating your CRFs and that these have been tested and loaded into OpenClinica and assigned to an event.

2 Preparing your data for import

2.1 ODIN template

ODIN will expect to receive your data file in a csv comma separated values file

	A	B	C	D	E	F	G	H	I
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	Data Item 1	Data Item 2	Data Item 3
2	ADD123		2015-05-01	PID_ADD123	F	1984-11-14			
3	ADD124		2015-05-01	PID_ADD124	M	1980-09-05			
4	ADD125		2015-05-01	PID_ADD125	F	1975-10-23			
5	ADD126		2015-05-01	PID_ADD126	F	1979-03-23			
6	ADD127		2015-05-01	PID_ADD127	M	1980-09-24			
7	ADD128		2015-05-01	PID_ADD128	M	1979-11-11			
8	ADD129		2015-05-01	PID_ADD129	F	1976-07-22			
9	ADD130		2015-05-01	PID_ADD130	F	1985-09-16			
10	ADD131		2015-05-01	PID_ADD131	M	1980-04-22			
11	ADD132		2015-05-01	PID_ADD132	F	1976-06-11			
12									

2.2 Repeating items

To be able to import data using ODIN (and most other importing tools) your data file needs to be formatted in such a way that each row (here 2 to 11) is a subject and the columns contain data for those subjects (A to I). In this case we have a repeating item for HPO terms, the columns G to I contain the HPO terms. Depending on the subject there is one or up to three HPO terms. To import this data as a repeating item we need to have the repeating measurements in one cell separated by 2 colons.

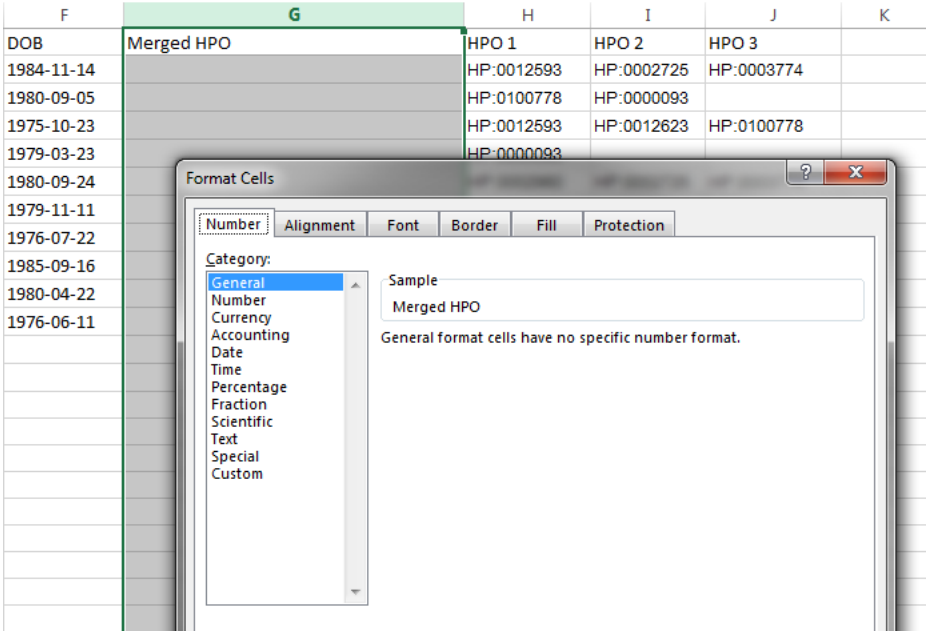
The instructions below will explain how to use a formula in Excel to correctly format your data.

Note: We would always recommend copying your data file instead of manipulating the data in your primary data file.

R18

	A	B	C	D	E	F	G	H	I
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	HPO 1	HPO 2	HPO 3
2	TRN_001			TRN_001	F	1984-11-14	HP:0012593	HP:0002725	HP:0003774
3	TRN_002			TRN_002	M	1980-09-05	HP:0100778	HP:0000093	
4	TRN_003			TRN_003	M	1975-10-23	HP:0012593	HP:0012623	HP:0100778
5	TRN_004			TRN_004	M	1979-03-23	HP:0000093		
6	TRN_005			TRN_005	F	1980-09-24	HP:0002960	HP:0002725	HP:0003774
7	TRN_006			TRN_006	F	1979-11-11	HP:0012593	HP:0000093	
8	TRN_007			TRN_007	M	1976-07-22	HP:0012623	HP:0000093	HP:0003453
9	TRN_008			TRN_008	F	1985-09-16	HP:0000093		
10	TRN_009			TRN_009	M	1980-04-22	HP:0006562	HP:0002725	HP:0002960
11	TRN_010			TRN_010	F	1976-06-11	HP:0003453	HP:0100778	HP:0003774

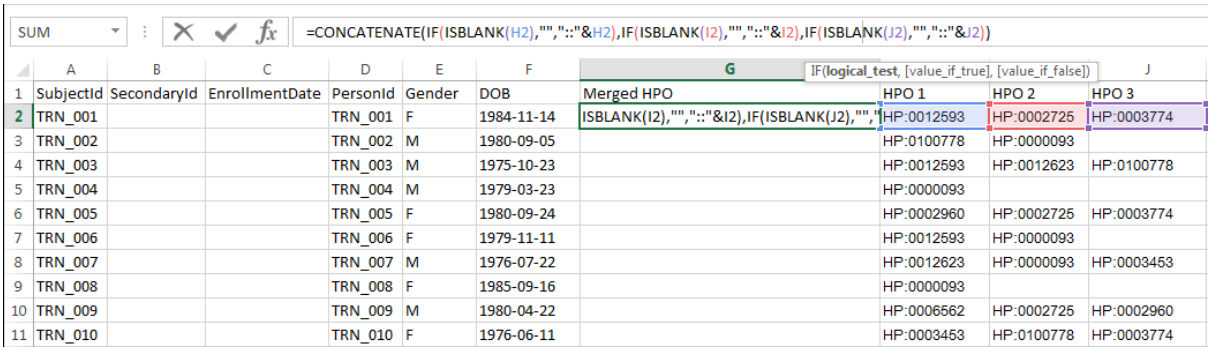
1. Add a new column to your excel sheet, *Merged HPO.*
2. Check that the format of the column is set to “General”



3. Add the following formula to the first cell in your *Merged HPO* column. You can either select the data columns (here H2, I2 and J2) by selecting the first cell in that column or by writing the cell coordinates (H2, I2 and J2) you want to combine.

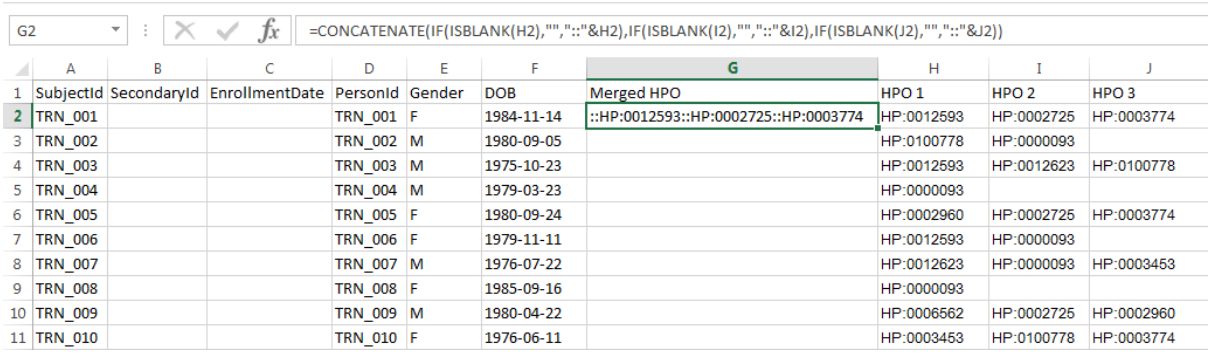
```
=CONCATENATE(IF(ISBLANK(H2),"","::"&H2),IF(ISBLANK(I2),"","::"&I2),IF(ISBLANK(J2),"","::"&J2))
```

The formula has 3 parts, the ISBLANK formula checks if the cell is blank and the IF Statement will insert nothing if the cell is blank and ":: + cell’s value". The CONCATENATE formula allows the 3 IF statements (one for each data column) to be combined.



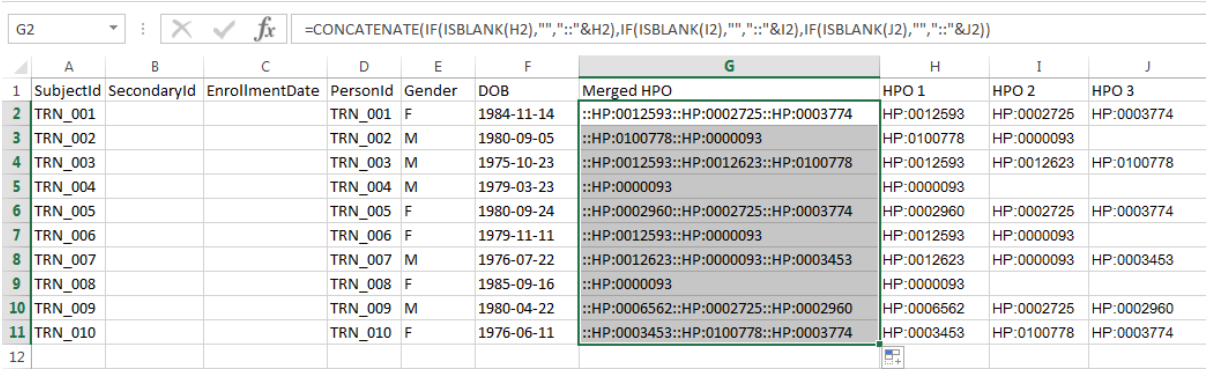
	A	B	C	D	E	F	G	H	I	J
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	Merged HPO	HPO 1	HPO 2	HPO 3
2	TRN_001			TRN_001	F	1984-11-14	ISBLANK(I2),"","::"&I2),IF(ISBLANK(J2),"","'	HP:0012593	HP:0002725	HP:0003774
3	TRN_002			TRN_002	M	1980-09-05		HP:0100778	HP:0000093	
4	TRN_003			TRN_003	M	1975-10-23		HP:0012593	HP:0012623	HP:0100778
5	TRN_004			TRN_004	M	1979-03-23		HP:0000093		
6	TRN_005			TRN_005	F	1980-09-24		HP:0002960	HP:0002725	HP:0003774
7	TRN_006			TRN_006	F	1979-11-11		HP:0012593	HP:0000093	
8	TRN_007			TRN_007	M	1976-07-22		HP:0012623	HP:0000093	HP:0003453
9	TRN_008			TRN_008	F	1985-09-16		HP:0000093		
10	TRN_009			TRN_009	M	1980-04-22		HP:0006562	HP:0002725	HP:0002960
11	TRN_010			TRN_010	F	1976-06-11		HP:0003453	HP:0100778	HP:0003774

4. Note that you can drag the formula to the right to make Excel automatically apply the same formula to any other column you might want to combine (not applicable in this example).



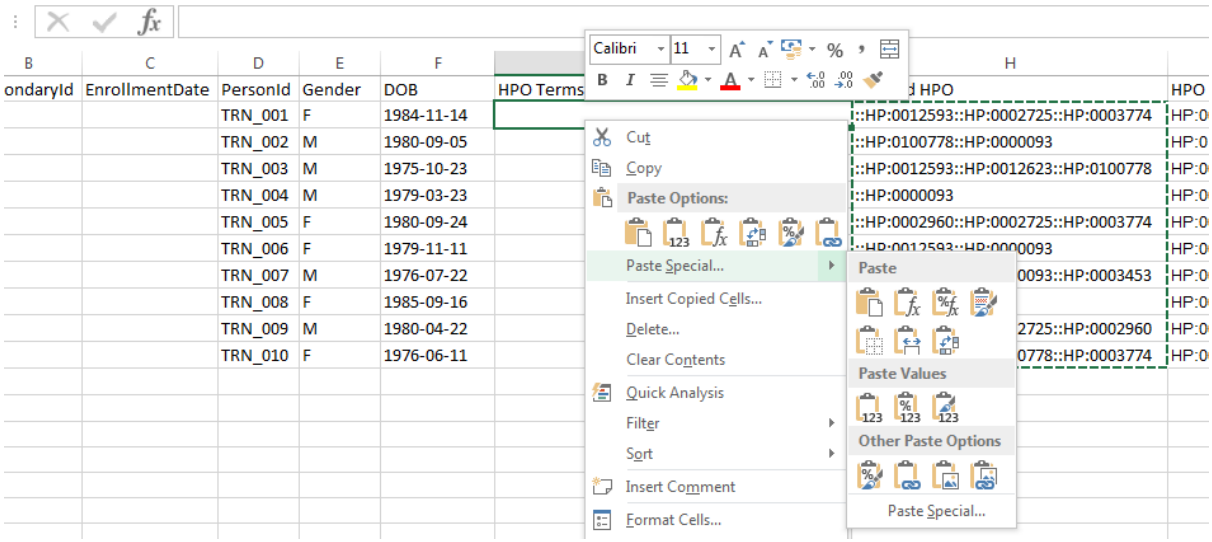
	A	B	C	D	E	F	G	H	I	J
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	Merged HPO	HPO 1	HPO 2	HPO 3
2	TRN_001			TRN_001	F	1984-11-14	::HP:0012593::HP:0002725::HP:0003774	HP:0012593	HP:0002725	HP:0003774
3	TRN_002			TRN_002	M	1980-09-05		HP:0100778	HP:0000093	
4	TRN_003			TRN_003	M	1975-10-23		HP:0012593	HP:0012623	HP:0100778
5	TRN_004			TRN_004	M	1979-03-23		HP:0000093		
6	TRN_005			TRN_005	F	1980-09-24		HP:0002960	HP:0002725	HP:0003774
7	TRN_006			TRN_006	F	1979-11-11		HP:0012593	HP:0000093	
8	TRN_007			TRN_007	M	1976-07-22		HP:0012623	HP:0000093	HP:0003453
9	TRN_008			TRN_008	F	1985-09-16		HP:0000093		
10	TRN_009			TRN_009	M	1980-04-22		HP:0006562	HP:0002725	HP:0002960
11	TRN_010			TRN_010	F	1976-06-11		HP:0003453	HP:0100778	HP:0003774

5. Drag the cell with the formula (G) down your subject rows to apply the formulas to the other subjects.

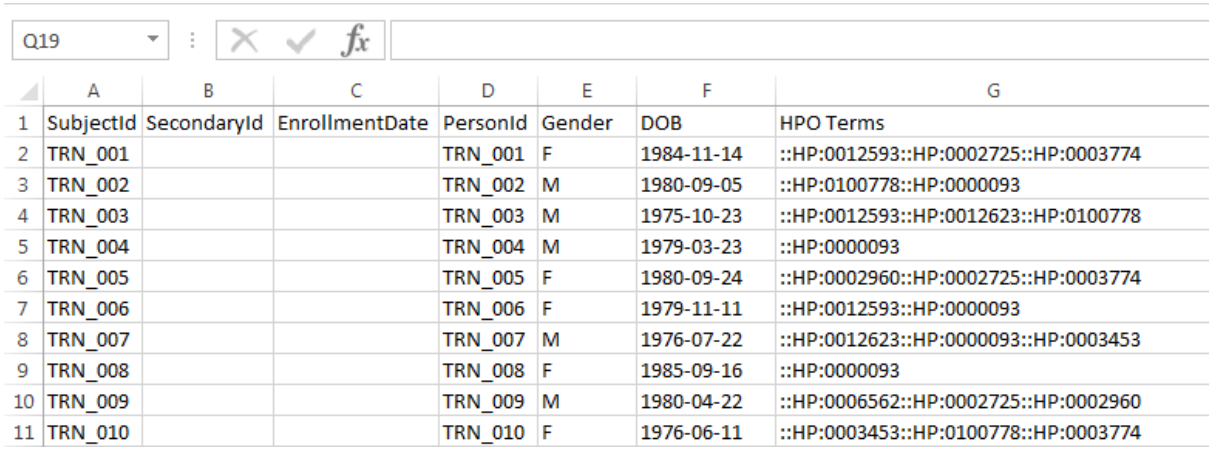


	A	B	C	D	E	F	G	H	I	J
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	Merged HPO	HPO 1	HPO 2	HPO 3
2	TRN_001			TRN_001	F	1984-11-14	::HP:0012593::HP:0002725::HP:0003774	HP:0012593	HP:0002725	HP:0003774
3	TRN_002			TRN_002	M	1980-09-05	::HP:0100778::HP:0000093	HP:0100778	HP:0000093	
4	TRN_003			TRN_003	M	1975-10-23	::HP:0012593::HP:0012623::HP:0100778	HP:0012593	HP:0012623	HP:0100778
5	TRN_004			TRN_004	M	1979-03-23	::HP:0000093	HP:0000093		
6	TRN_005			TRN_005	F	1980-09-24	::HP:0002960::HP:0002725::HP:0003774	HP:0002960	HP:0002725	HP:0003774
7	TRN_006			TRN_006	F	1979-11-11	::HP:0012593::HP:0000093	HP:0012593	HP:0000093	
8	TRN_007			TRN_007	M	1976-07-22	::HP:0012623::HP:0000093::HP:0003453	HP:0012623	HP:0000093	HP:0003453
9	TRN_008			TRN_008	F	1985-09-16	::HP:0000093	HP:0000093		
10	TRN_009			TRN_009	M	1980-04-22	::HP:0006562::HP:0002725::HP:0002960	HP:0006562	HP:0002725	HP:0002960
11	TRN_010			TRN_010	F	1976-06-11	::HP:0003453::HP:0100778::HP:0003774	HP:0003453	HP:0100778	HP:0003774
12										

6. Add a new columns to your excel sheet (*HPO Terms*) and Copy & Paste Value the data from *Merged HPO* across to the new columns.



7. After quickly checking that there were no errors you can now delete the extra columns (H to K) from your excel sheet



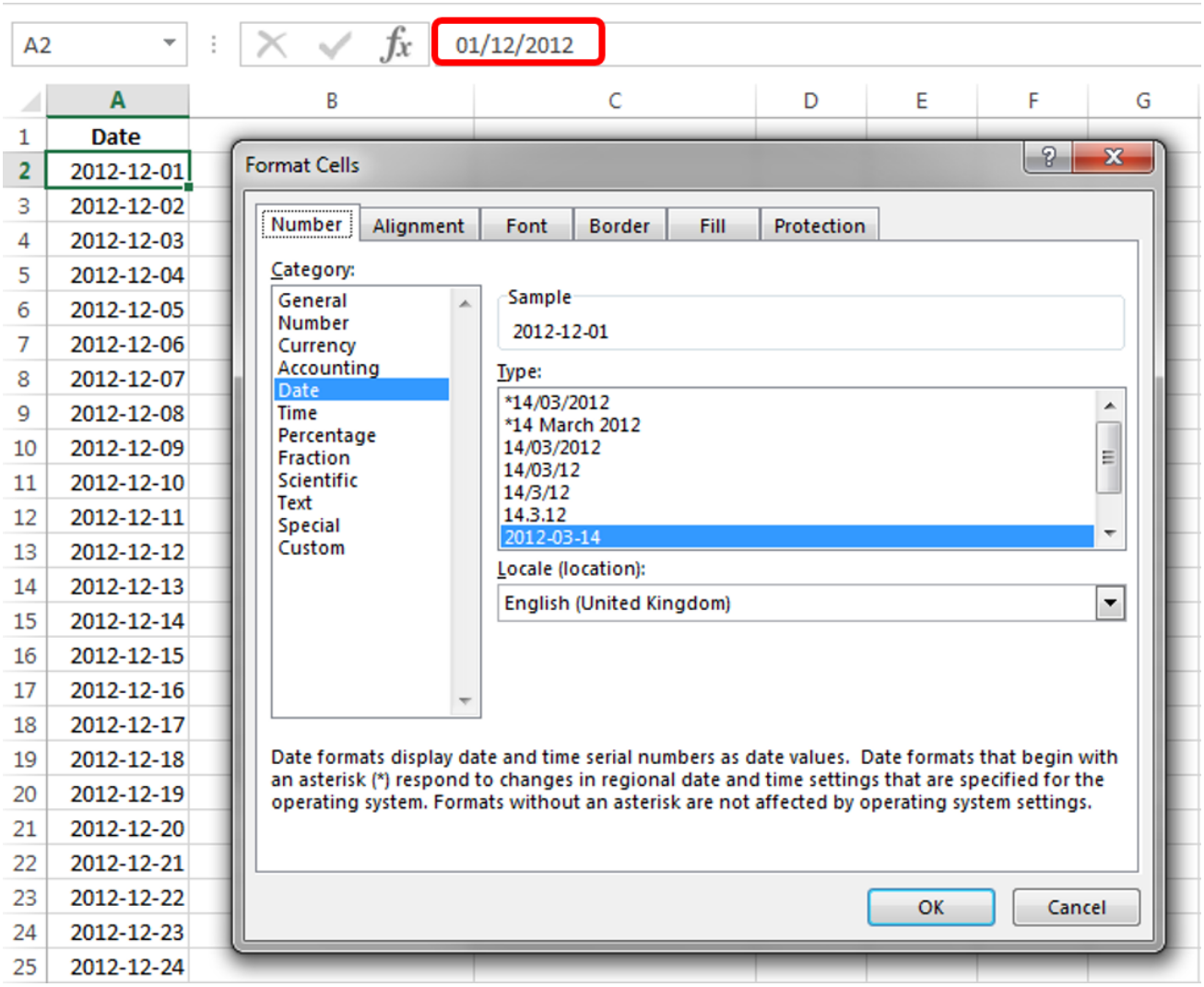
	A	B	C	D	E	F	G
1	SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB	HPO Terms
2	TRN_001			TRN_001	F	1984-11-14	::HP:0012593::HP:0002725::HP:0003774
3	TRN_002			TRN_002	M	1980-09-05	::HP:0100778::HP:0000093
4	TRN_003			TRN_003	M	1975-10-23	::HP:0012593::HP:0012623::HP:0100778
5	TRN_004			TRN_004	M	1979-03-23	::HP:0000093
6	TRN_005			TRN_005	F	1980-09-24	::HP:0002960::HP:0002725::HP:0003774
7	TRN_006			TRN_006	F	1979-11-11	::HP:0012593::HP:0000093
8	TRN_007			TRN_007	M	1976-07-22	::HP:0012623::HP:0000093::HP:0003453
9	TRN_008			TRN_008	F	1985-09-16	::HP:0000093
10	TRN_009			TRN_009	M	1980-04-22	::HP:0006562::HP:0002725::HP:0002960
11	TRN_010			TRN_010	F	1976-06-11	::HP:0003453::HP:0100778::HP:0003774

You now have the merged repeating items for HPO codes in one cell ready for import.

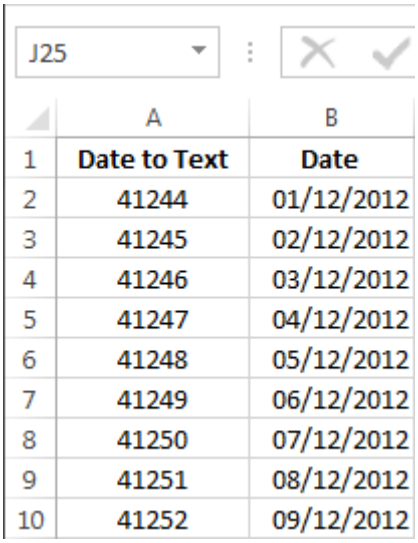
2.3 Converting Date format to Text

The internationally accepted date format is ISO 8601 (yyyy-mm-dd) and it is the format that OpenClinica uses to process any date information. All dates in the ODIN data file will have to be in ISO 8601 format. ODIN will check the correct formatting of the Enrollment Date and the Date of Birth since they are part of the mandatory 6 columns in the ODIN data file but if there are dates in the subsequent data columns the problem of wrong date formatting will not be apparent until the user uploads imports the xml file using OpenClinica.

If you change the date format in excel to match ISO 8601 this is a visual change only, the underlying date format is still dd-mm-yyyy. Additionally there is the problem that Excel reformats the csv file according to its own rules.

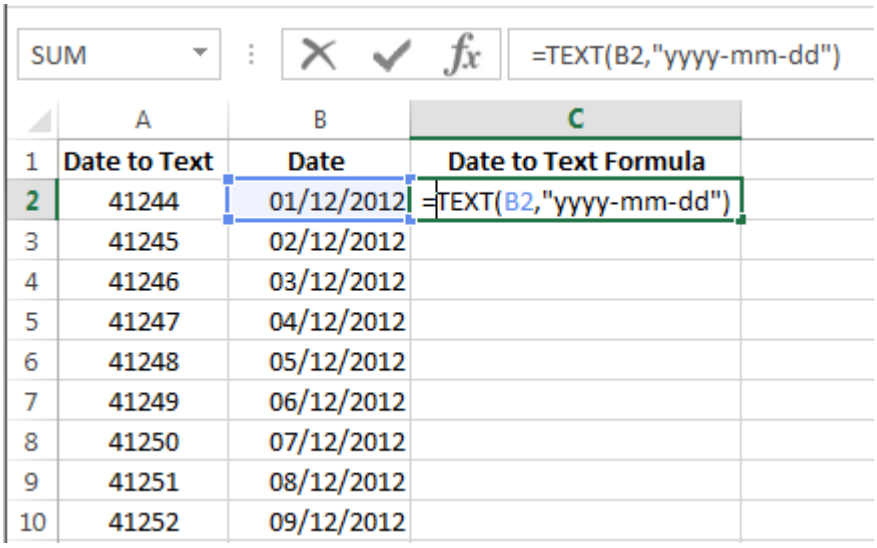


Converting a date column to Text format in Excel results in a value that counts the number of days since 1900.



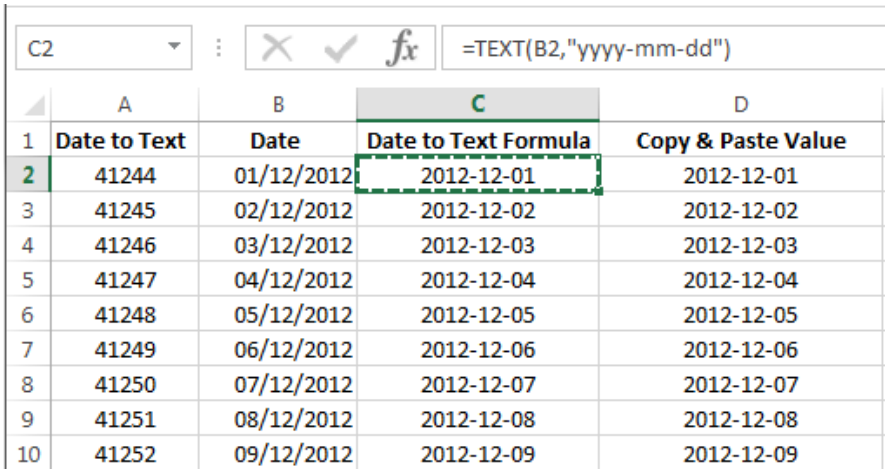
	A	B
1	**Date to Text**	**Date**
2	41244	01/12/2012
3	41245	02/12/2012
4	41246	03/12/2012
5	41247	04/12/2012
6	41248	05/12/2012
7	41249	06/12/2012
8	41250	07/12/2012
9	41251	08/12/2012
10	41252	09/12/2012

Using the Excel TEXT formula it is possible to convert dates to a text value in the yyyy-mm-dd format (or date format).



	A	B	C
1	**Date to Text**	**Date**	**Date to Text Formula**
2	41244	01/12/2012	=TEXT(B2,"yyyy-mm-dd")
3	41245	02/12/2012	
4	41246	03/12/2012	
5	41247	04/12/2012	
6	41248	05/12/2012	
7	41249	06/12/2012	
8	41250	07/12/2012	
9	41251	08/12/2012	
10	41252	09/12/2012	

The *Date to Text Formula* column (C) now holds the date in the ISO 8601 format but it is still a formula rather than a value. Add a new column and Copy & Past Value across from the *Date to Text Formula* column (C).



C2 =TEXT(B2,"yyyy-mm-dd")

	A	B	C	D
1	Date to Text	Date	Date to Text Formula	Copy & Paste Value
2	41244	01/12/2012	2012-12-01	2012-12-01
3	41245	02/12/2012	2012-12-02	2012-12-02
4	41246	03/12/2012	2012-12-03	2012-12-03
5	41247	04/12/2012	2012-12-04	2012-12-04
6	41248	05/12/2012	2012-12-05	2012-12-05
7	41249	06/12/2012	2012-12-06	2012-12-06
8	41250	07/12/2012	2012-12-07	2012-12-07
9	41251	08/12/2012	2012-12-08	2012-12-08
10	41252	09/12/2012	2012-12-09	2012-12-09

You can rename the Co*py & Paste Value* column to *Date* and delete the additional columns.

2.4 Preparing coding data for importing into OpenClinica

Using DECODE function in CRFs is the best way for capturing the related ontology interpretation code for data items (Refer to Data Plan Design).

As the DECODE function in the CRF does not trigger when data is imported into OpenClinica, the relevant clinical codes must be added into the raw dataset prior to importing.

Using the “Asthma” as an example, the codes need to be added to the dataset for importing, and modify the CRF against importing correspondently.

Subject	Asthma	Asthma_HPO
1	1	*HP:0002099*
2	0	
3	1	*HP:0002099*
4	0	

Insert a new column to add the HPO codes

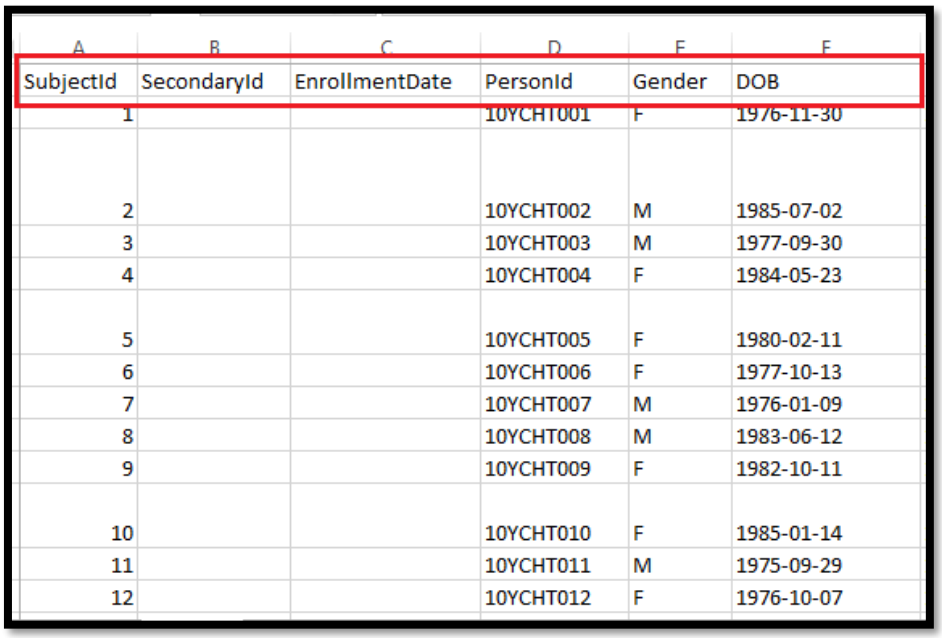
"1" means the subject has Asthma and a HPO code should be added, "0" means the subject doesn't have Asthma, so no HPO code added

3 Importing data

3.1 Last preparation before importing the data

After the process of cleaning your data, the next steps are to:

1. Columns A to F must contain the column headings **SubjectId**, **SecondaryId, EnrollmentDate, PersonId, Gender** and **DOB** as shown below. The import is case sensitive so ensure column headings are exactly as show.

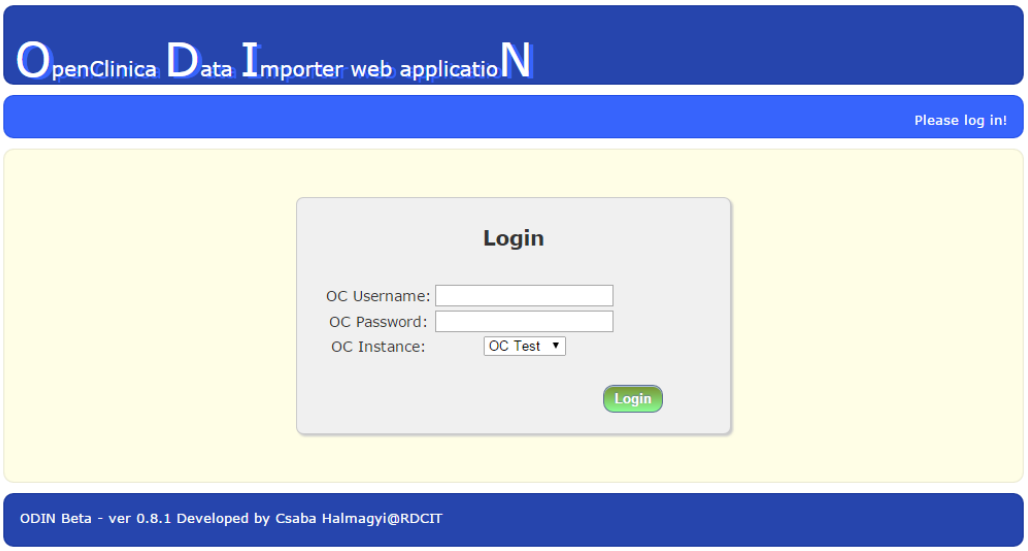


A	B	C	D	E	F
SubjectId	SecondaryId	EnrollmentDate	PersonId	Gender	DOB
1			10YCHT001	F	1976-11-30
2			10YCHT002	M	1985-07-02
3			10YCHT003	M	1977-09-30
4			10YCHT004	F	1984-05-23
5			10YCHT005	F	1980-02-11
6			10YCHT006	F	1977-10-13
7			10YCHT007	M	1976-01-09
8			10YCHT008	M	1983-06-12
9			10YCHT009	F	1982-10-11
10			10YCHT010	F	1985-01-14
11			10YCHT011	M	1975-09-29
12			10YCHT012	F	1976-10-07

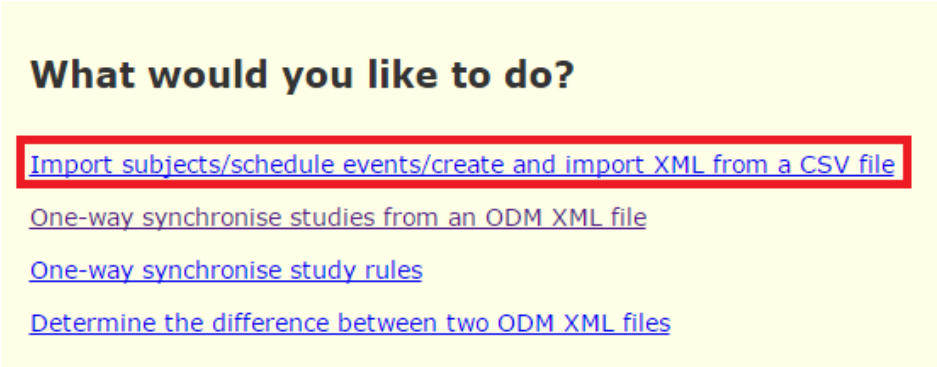
2. If your study is not collecting date of birth (**DOB**) or is not using **PersonId**, **SecondaryId, EnrollmentDate** or **Gender** the import file still expects to see the column headings. The first column is the **SubjectId** which should never be empty.
3. Save your data file as [studyname].csv.

3.2 ODIN

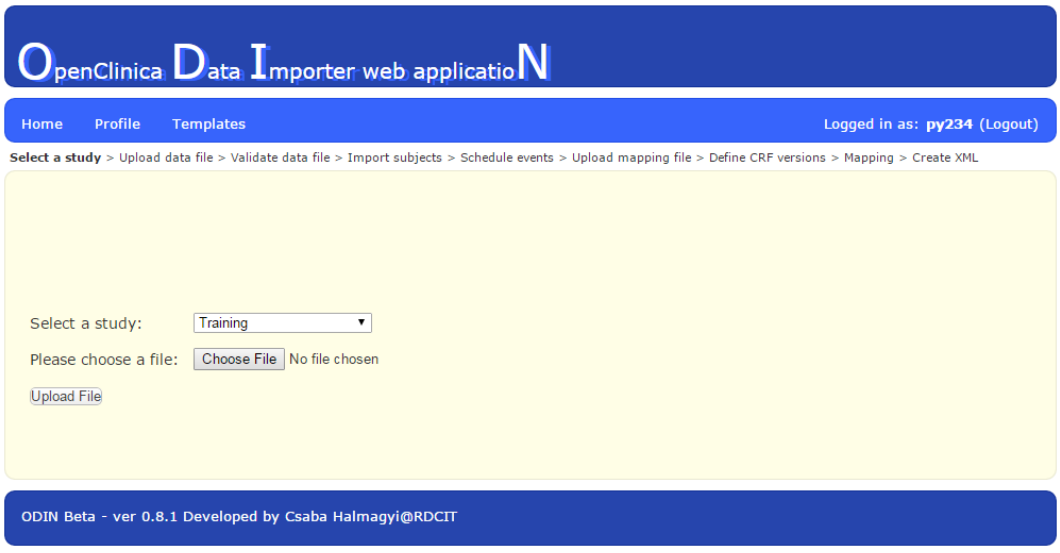
To login to ODIN visit the following URL: http://yourwebserver.com/odinbeta where you should be able to see the ODIN login screen.



To log in you have to choose an OpenClinica instance and enter your OpenClinica username and password. Your username should be authorised to use the web services. After a successful login, it leads to the task menu, choose “Import subjects/schedule events/create and import XML from a CSV file”:

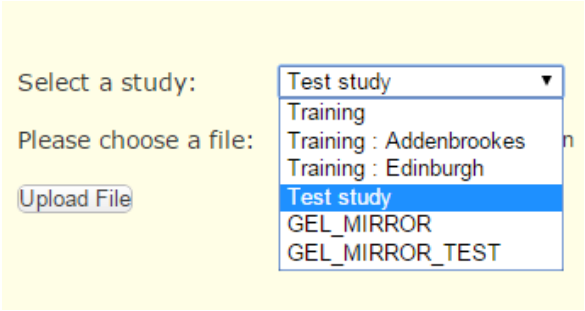


the ODIN main menu will be displayed:

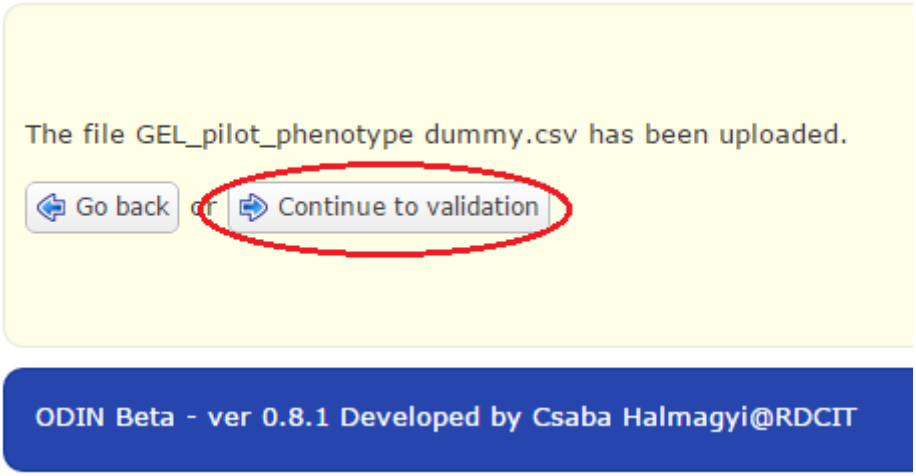


1. Select a study and upload a data file.

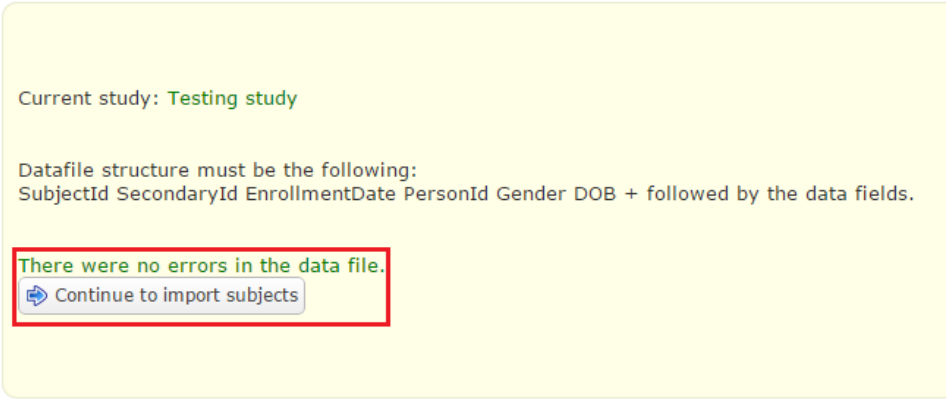
In the main menu, select a study which you would like to import against from the dropdown list, click ‘Choose File’ and navigate to the file [studyname].csv



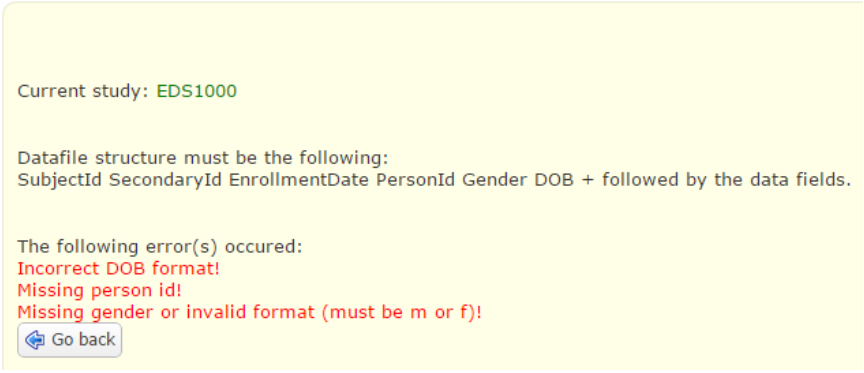
2. Click “Upload file”, the validation page opens:



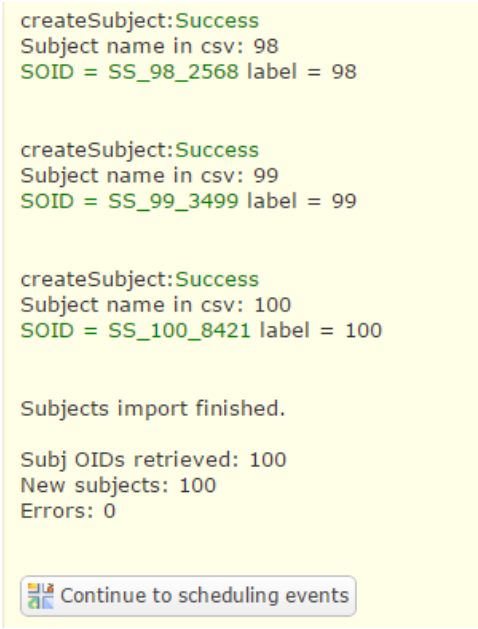
3. Click “Continue to validation” of the upload file. If there are no errors, you can click “Continue to import subjects”.



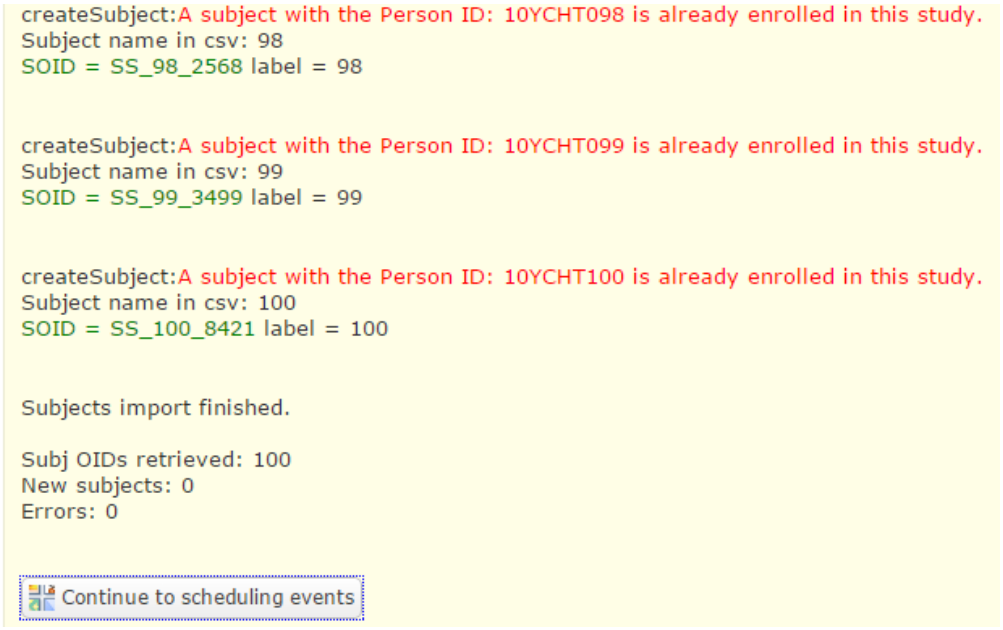
ODIN has some basic validation scripts built in which run after a data file has been uploaded. If you have any errors in the data file, ODIN will notify you about these errors. Any errors have to be fixed before continuing to the subjects import. Please note that during this step only the first six columns will be validated by the application. After creating and uploading the CDISC ODM XML file to OpenClinica other errors (such as invalid data values, etc.) could be possibly spotted.



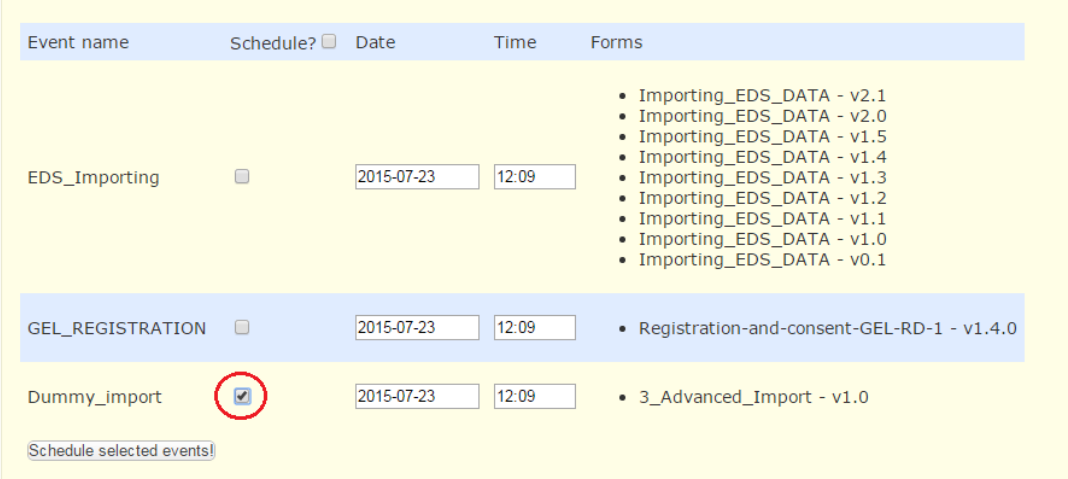
4. After Clicking “Continue to import subjects”, “Continue to scheduling events” opens if there is no error.



If the subjects have already been loaded to the study, new subject records will not be created, instead the below message will show “subject X exists for study Y”. New records will be show “successful” with a summary of new records shown at the bottom of the screen.



5. Click “Continue to scheduling events”. The next screen shows all the events in your study, select the ones you want to schedule, accept the default date or add a specific date. (Note you cannot schedule individual dates for each subject – if you need to schedule specific dates, your data import file will need to be split according to event date)



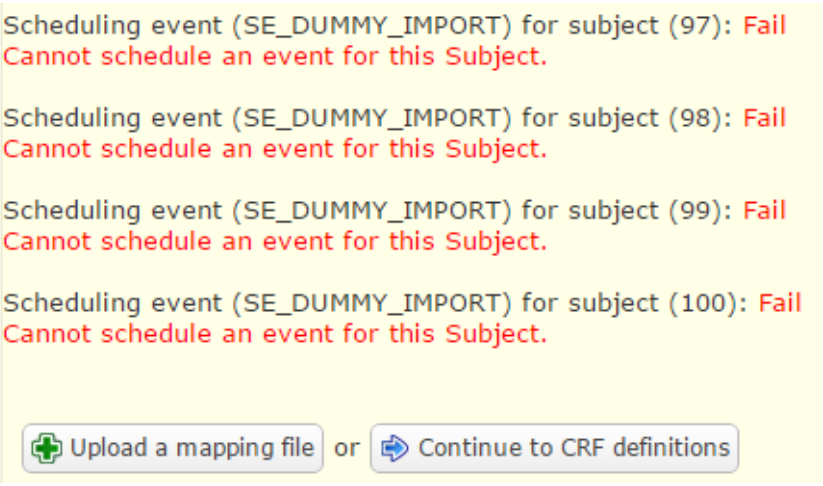
ODIN will display all events in the current study with all the CRFs assigned to them. Any repeating events in the study will be marked with an orange background.

ODIN will use the current date and time when scheduling the events unless you provide a different date and time. You can schedule selected events or if the events have already been scheduled uncheck all events and schedule none.

NOTE: Be careful when scheduling a repeating event as every time you schedule an event a new occurrence will be created. When dealing with repeating events ODIN will import data against the most recent occurrence of the event.

6. Click “Schedule selected events”. The next screen will show a line for every event per subject as scheduled with “success”.

If the event has already been scheduled for a particular subject, a warning will show.

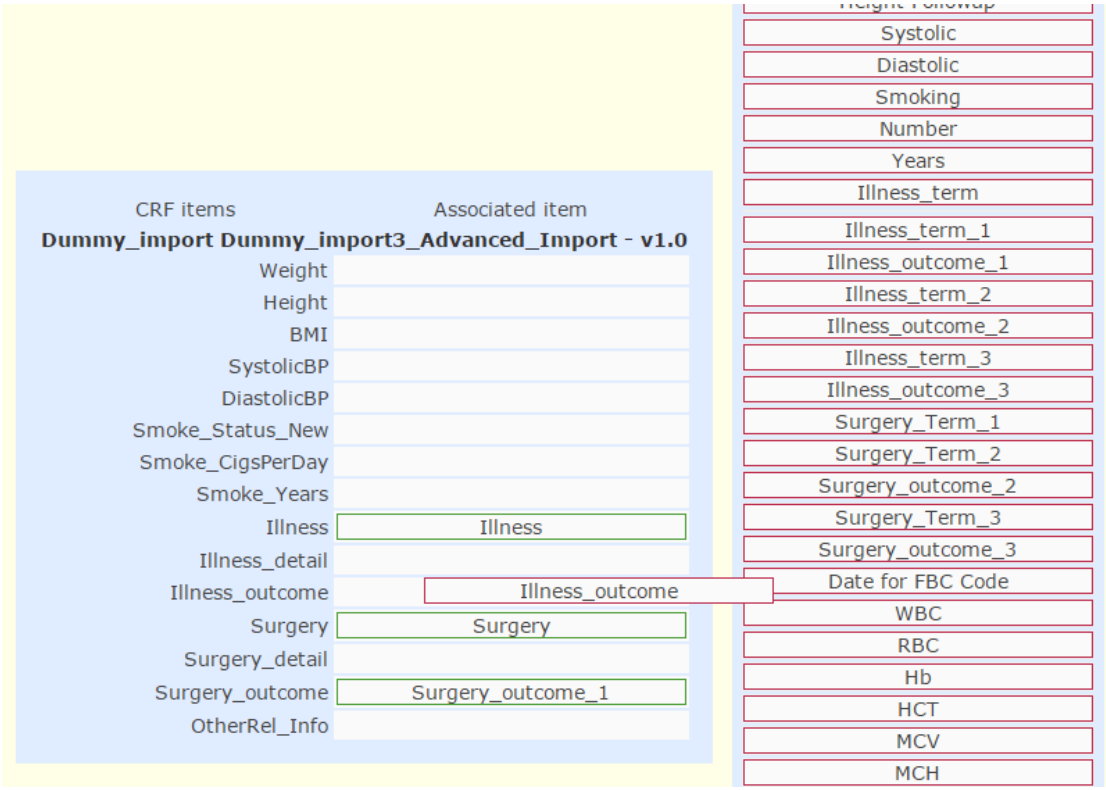


7. Upload an existing mapping file or “Continue to CRF definitions”

You can upload a previously created mapping file to skip the manual mapping process. This step is optional. If you don’t have a mapping file they can continue to define the CRF versions.

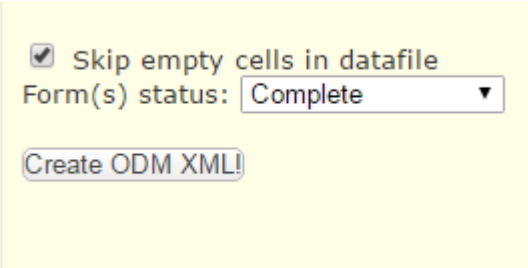
You can define the CRF versions they would like to import data against. All the events with all the assigned CRF versions will be displayed here. Unchecking forms/items first in this phase will prevent those items from being used during the mapping process, and select the event and the assigned CRF you wish to schedule.

8. Click “Continue to mapping”, the mapping module displays. Only the fields from the event ticked on the previous screen show on the left under CRF items. On the right side there are the .CSV data file headers. Drag and drop the appropriate fields from the CSV import file against the CRF fields.

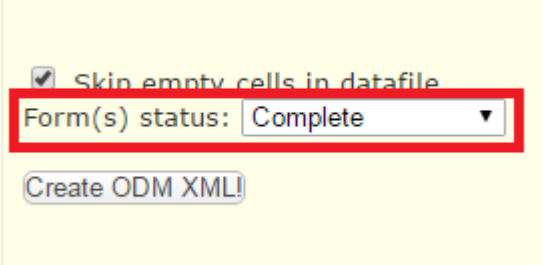


You can remove the associations by dragging the associated item and dropping it back to the right side. If a mapping file was uploaded before defining the CRF versions the manual association can be skipped, but the associations from the mapping file cannot be overwritten manually.

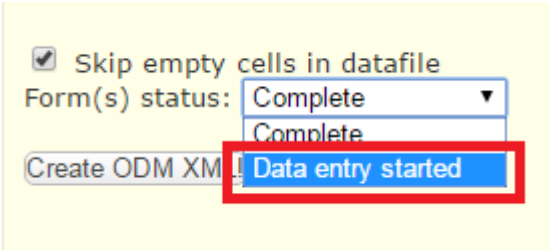
Any fields not mapped, will not have data imported against them. If you have a mix of subjects who do and don't have data for a particular event, tick the bottom box "Skip empty cells in datafile".



The default value of "Form(s) status" is "Complete". It indicates "Green tick" in the Subject Matrix once the import is completed.



You might change the status to "Data entry started", and it will indicate "Orange" in the Subject Matrix once the import is completed. In this case, the clinicians will verify the data of all the Subjects and manually mark the CRF as complete after reviewing the data.

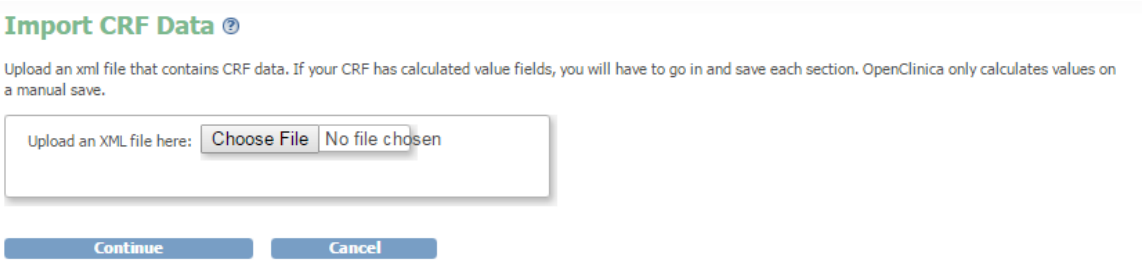


After finishing the associations the ODM XML file can be created. There must be at least one header assigned to an item in order to create the XML file.

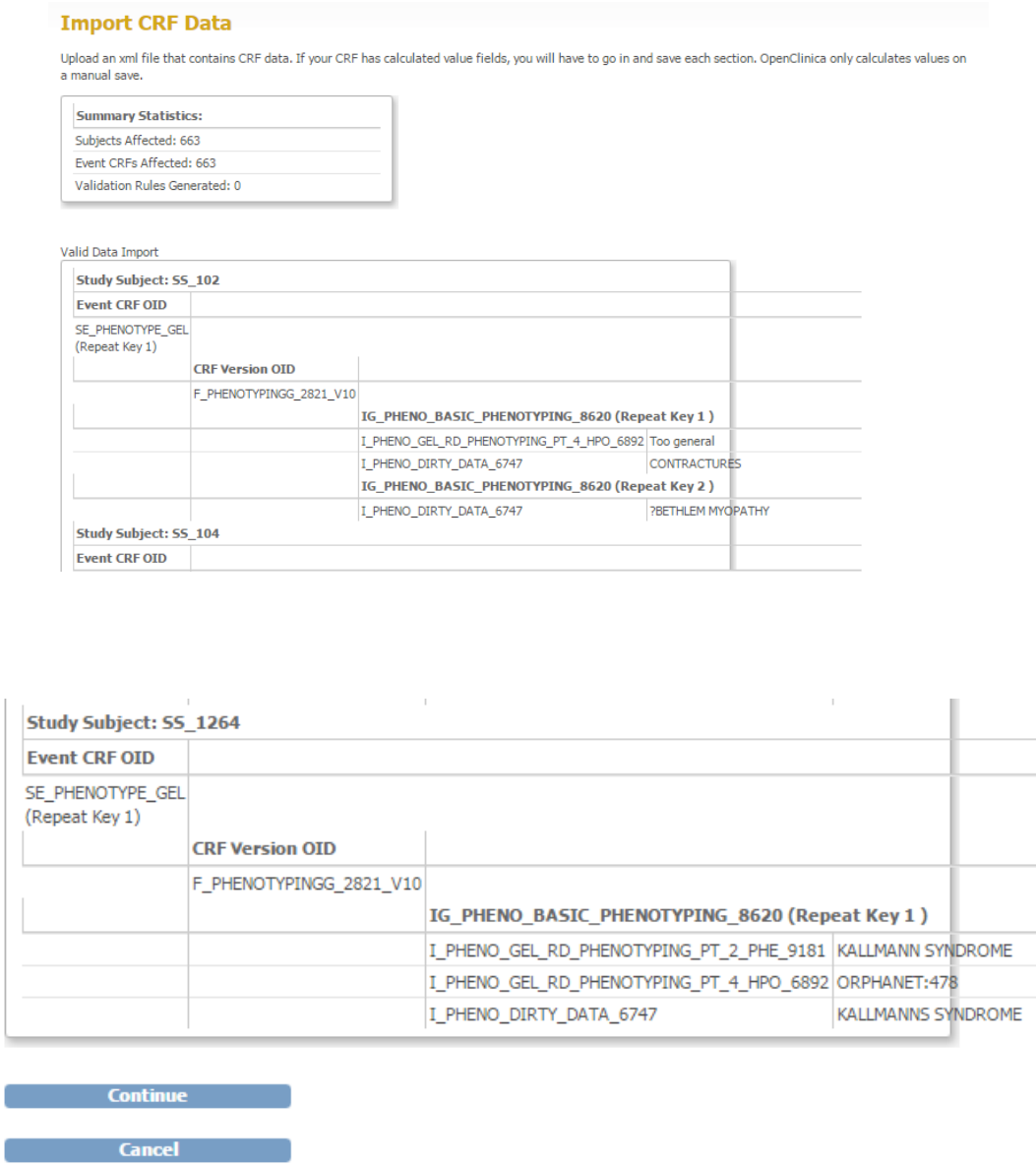
9. ODIN will create a CDISC ODM XML file based on the mapping rules created previously. This file can be download by clicking on the button as shown below.

If you are allowed to import directly to this OpenClinica instance the “Import data from this XML” button will be displayed, but we strongly recommend to disable this option. Instead of importing directly to OpenClinica, download the XML file and use the OpenClinica “Import Data”

10. In OpenClinica, under Task >Import Data. The Import Data page opens:

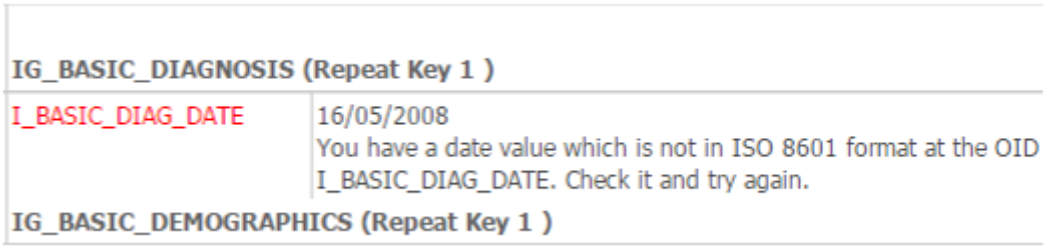


Choose file which you would like to import, and click “Continue” to start the importing process. OpenClinica import task will validate the XML data prior to the import and provide error messages if any errors were found. If there is no error, click “Continue”. And the data importing is completed.

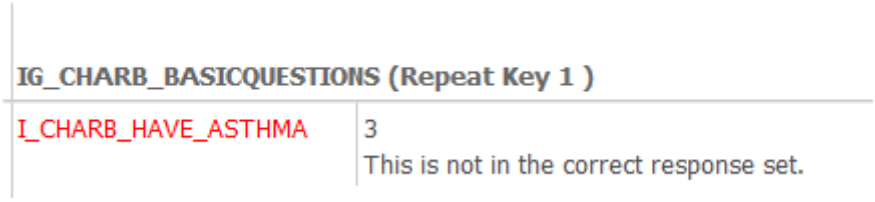


There are some common errors and problems which are not filtered by ODIN:

1. The data in your CSV data file is stored in a wrong format;



2. The data loaded from the .CSV file violates OpenClinica response set (e.g. the input is expected to be 1 or 2 and you want to import 3)



3. Using special character encode which is not supported by OpenClinica (the ODM XML will be invalid)
4. Creating an invalid XML due to mismapping the OpenClinica items and the CSV headers.

If an error was found, you are prompted to upload a new XML file at the bottom.

By clicking on the "Start a new import", ODIN will reset the current import session, and redirect you to the main menu and generate a new session Id for a new import. If you return to the main menu without resetting the import session and start a new import, all XML and mapping rules will be overwritten. Please don't use the web browsers 'go back' arrow back to the previous page as ODIN may lose some critical information.

3.3 Mapping Files

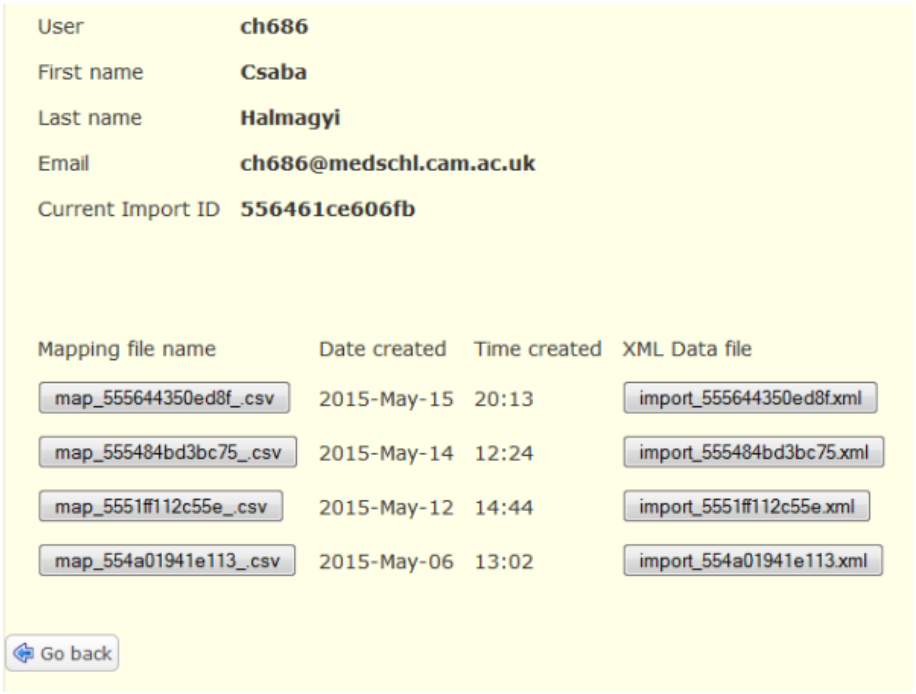
Every time you click "Create ODM XML" button, a mapping file will be created to store the association rules. These files contain three columns:

1. The first column is for storing the full OID reference of an item in OpenClinica. The OIDs are separated by double hashes. For example:

SE_BASICQUESTIONS##F_CHARBASICFOR_V11##IG_CHARB_BASICQUESTIONS##I_CHARB_GENHEALTH

2. The second column contains the event name, the form name and the item name.
3. The third column contains the CSV datafile header name associated to the item.

Once a mapping file has been created, it can be reused if you import data against the same items (overwriting previously imported data or importing data for new subjects). The previously saved mapping files can be downloaded via Clicking "profile" at ODIN home page.



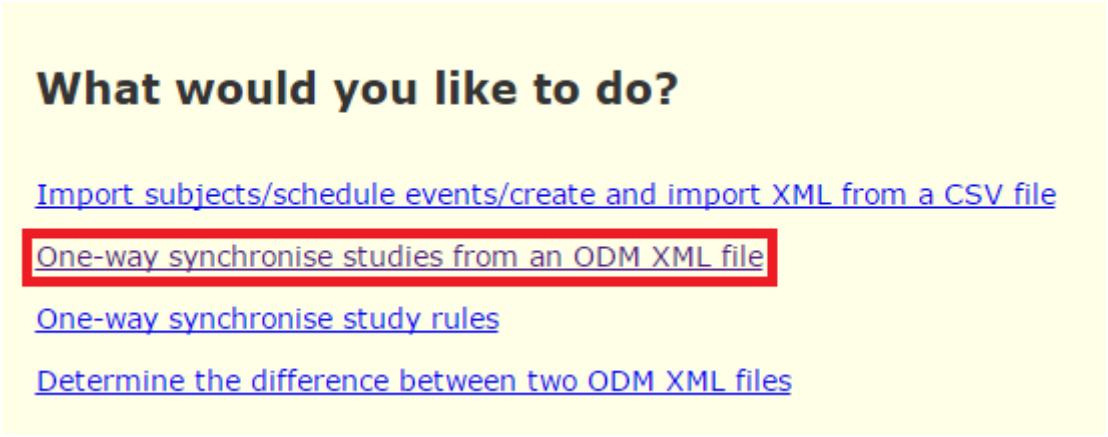
3.4 Data import to the RDCIT Secure Data Service

For data importing to the RDCIT SDS, the user will need a SSL VPN access to the RDCIT SDS. Only the SSL VPN access allows the user to import data to the RDCIT SDS, not GATEWAY and HTML 5 access. Please refer to RDCIT SDS Access Guide for more details.

The user needs to connect the SSL VPN (the traffic light) prior to import data to the RDCIT SDS. Once the traffic light is on, the user will be able to log on to the ODIN (http://methods.portal.rd.trc.nihr.ac.uk/odin/login.php) and start importing.

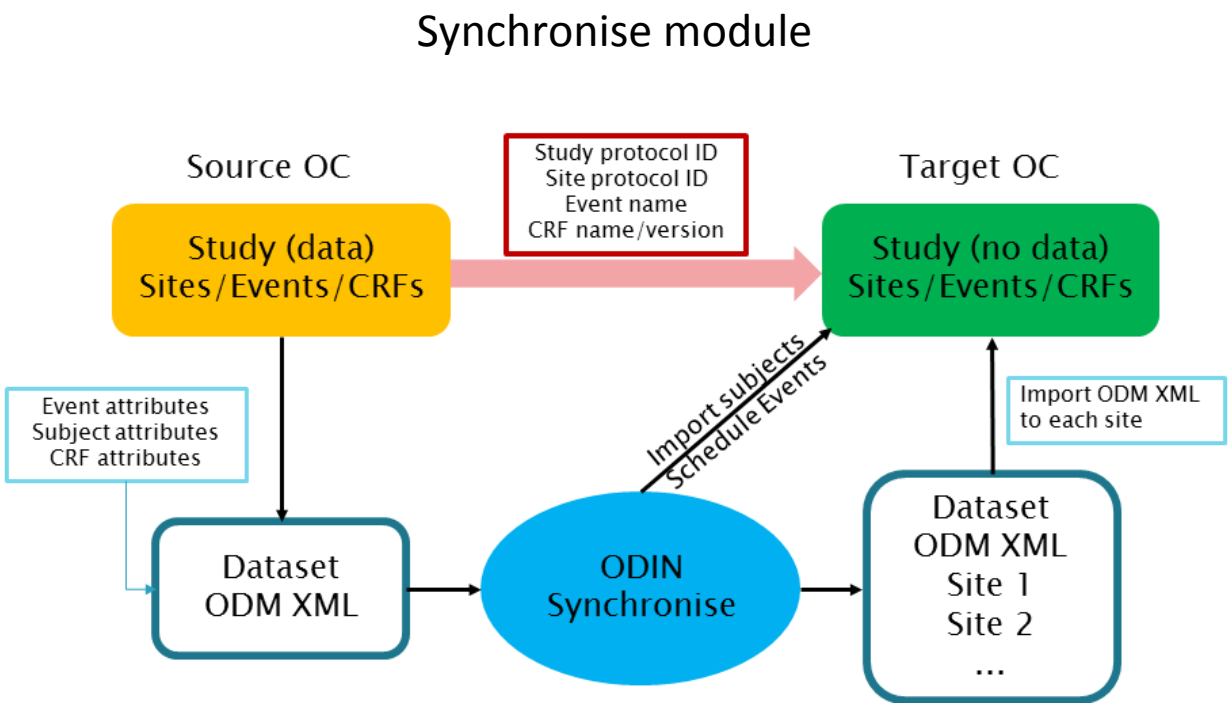
4 Synchronise studies

ODIN has the function to synchronise a study from one OpenClinica instance (Source OpenClinica) to another OpenClinica instance (Target OpenClinica) with the second option "*One_way synchronise studies from an ODM XML file*" on the task menu in ODIN.



4.1 Synchronise module

The below is the diagram explains how the synchronisation works: Create a data set for a study in the Source OpenClinica and use ODIN to import the dataset to the target OpenClinica.



Prior to synchronise study using ODIN, you will need to complete two requirements to prepare synchronisation: **create a new study in the target OpenClinica**, and **create an ODM XML file of your study in the source OpenClinica**

1. Create a study in the target OpenClinca

For study synchronisation, there are some rules for creating a new study in the target OpenClinica:

- *The Study protocol ID* in the Source OpenClinica is the same as the Study protocol ID in the target OpenClinica.
- *The Site protocol ID* in the Source OpenClinica is the same as the Site protocol ID in the target OpenClinica.
- *The Event name* in the Source OpenClinica is the same as the Event name in the target OpenClinica.
- *The CRF name and version* in the Source OpenClinica are the same as the CRF name and version in the target OpenClinica.

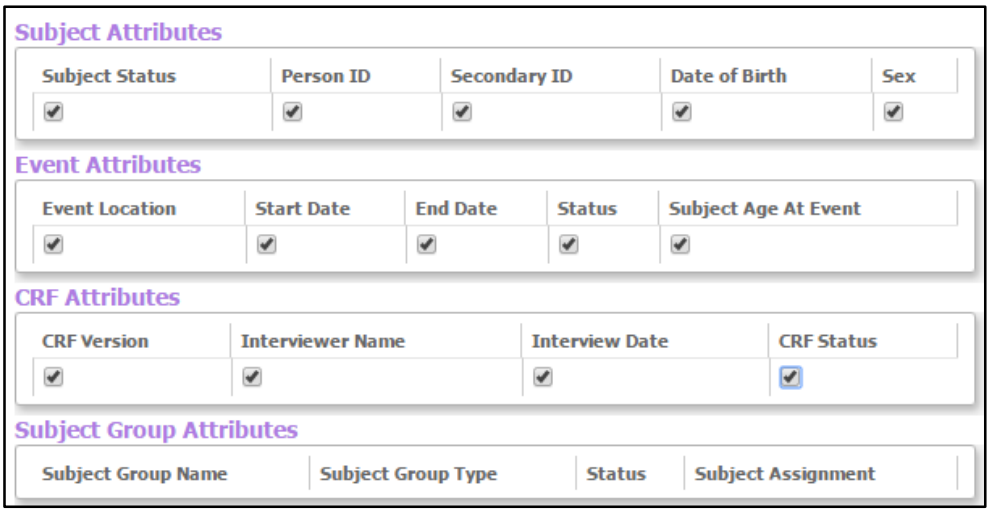
NB: The study name and site name in the source OpenClinica can be different from the target OpenClinca.

2. Create an ODM XML file of your study in the source OpenClinica

For the details on how to create a Dataset, please read **section 5** "Data Export from OpenClinica" for more information.

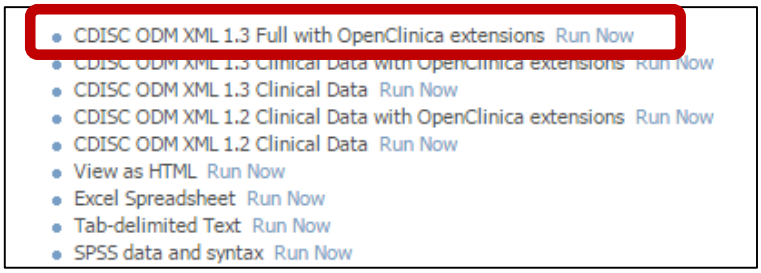
For study synchronisation, there are some rules for creating an ODM XML file:

- Include Event attributes, Subject attributes, and CRF attributes to the dataset.



- Select "Data from all Available CRFs" when specifying the dataset properties.

- Choose “CDISC ODM XML 1.3 Full with OpenClinica extensions” to download the file.

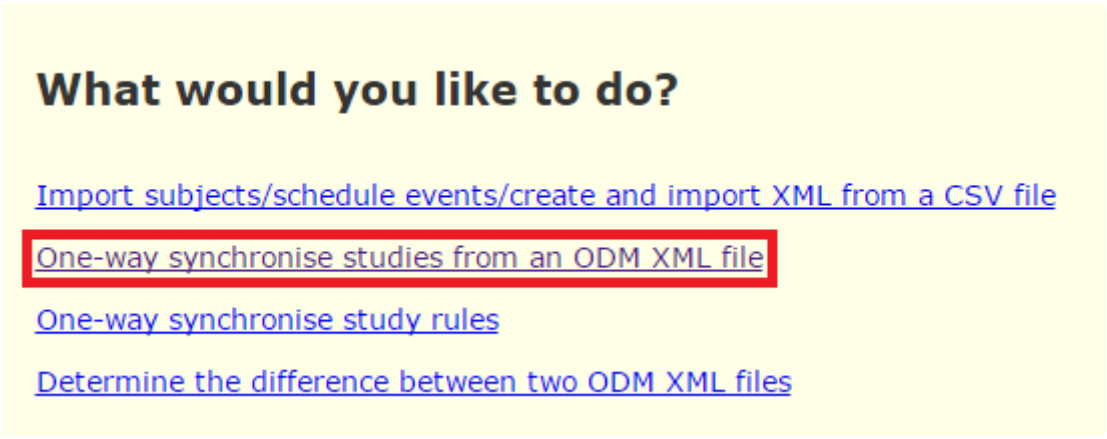


Once you finish creating the study and ODM XML file, you can start the process of synchronisation.

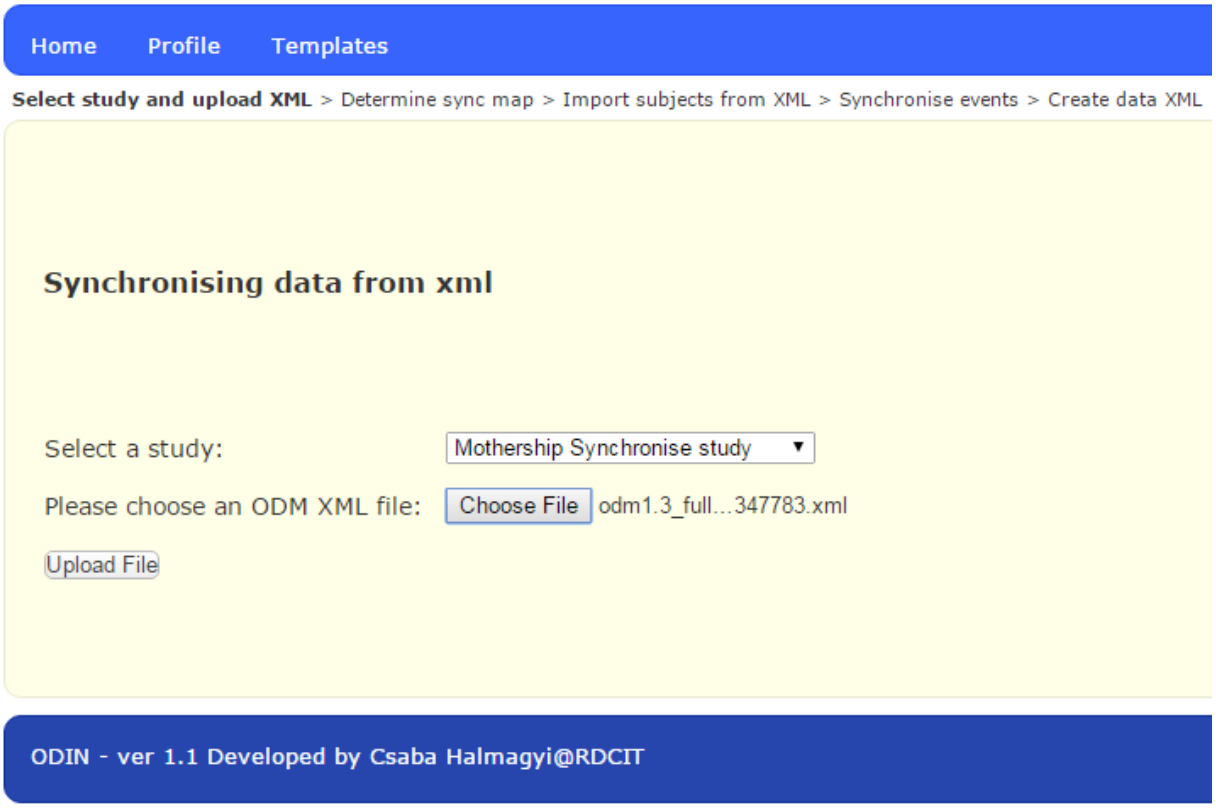
4.2 Synchronise Study using ODIN

1. Log on to ODIN, using your target OpenClinica credential.

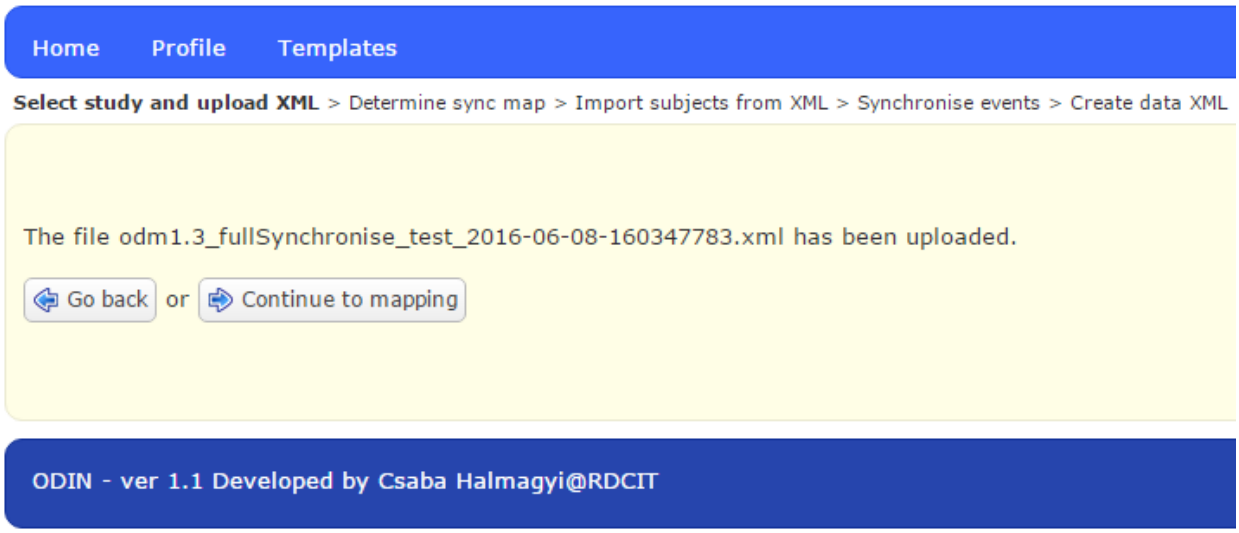
2. The task menu page opens, select the second option “one – way synchronise studies from an ODM XML file.



3. Select a study from the drop down list and upload the ODM XML file.

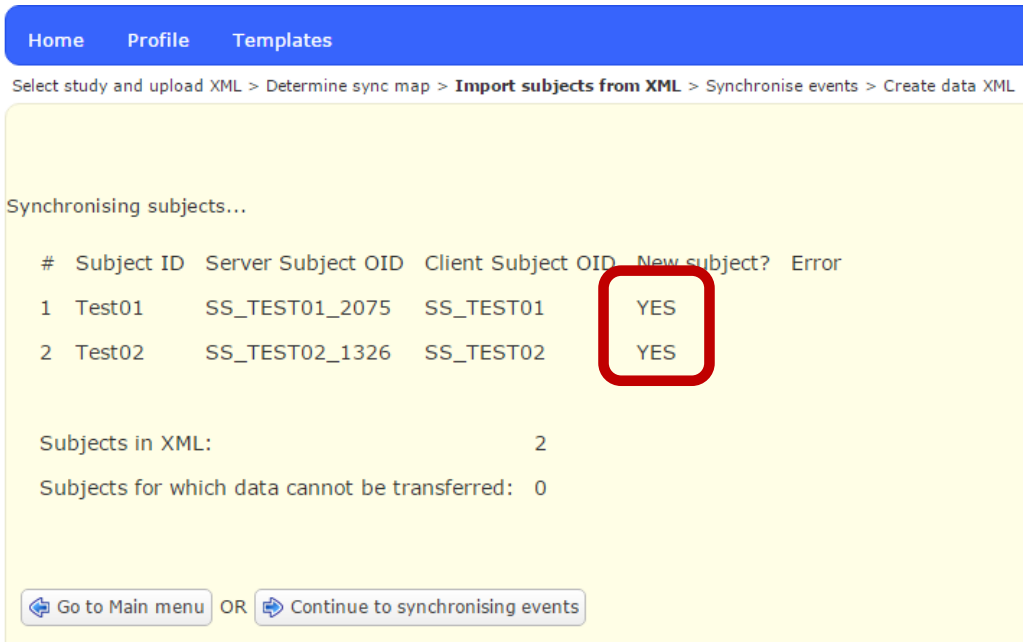


4. Continue to mapping:

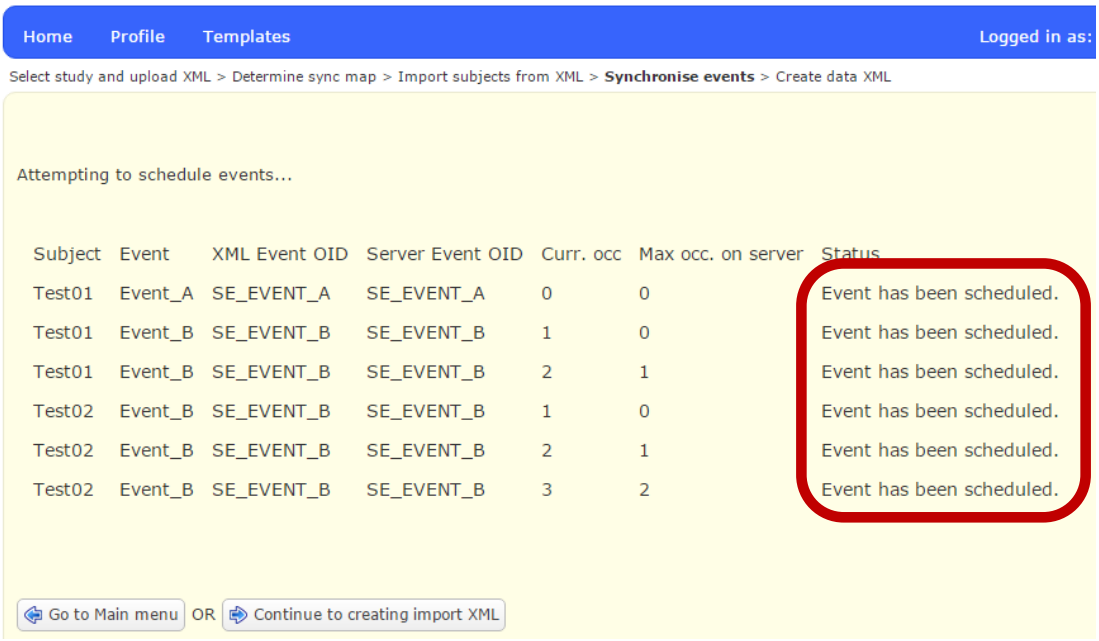


5. Verify the study unique protocol ID, Site unique protocol ID, Event name and CRFs name/version. if they are not the same as the study in the target OpenClinica, you will not be able to continue to synchronise the study.

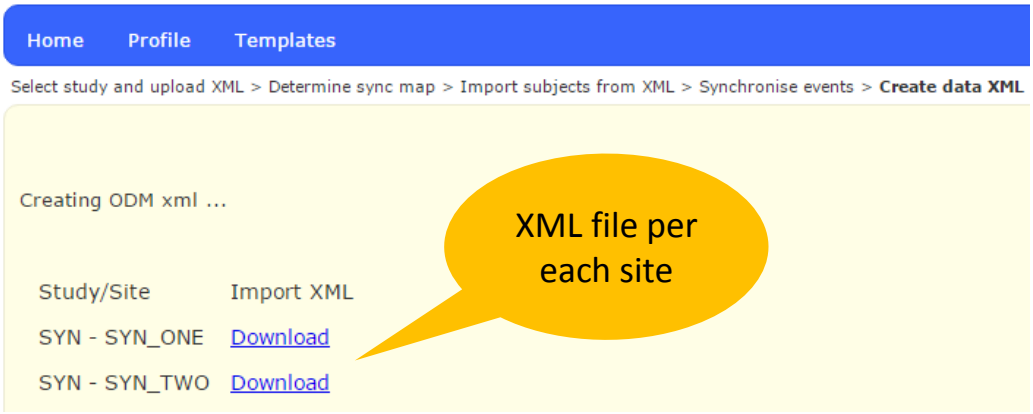
6. Import subjects:



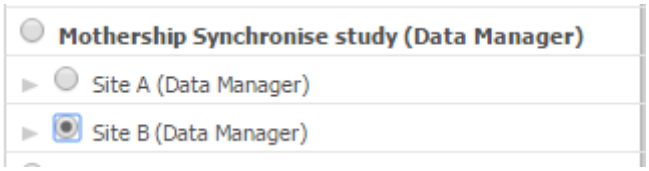
7. Schedule events:

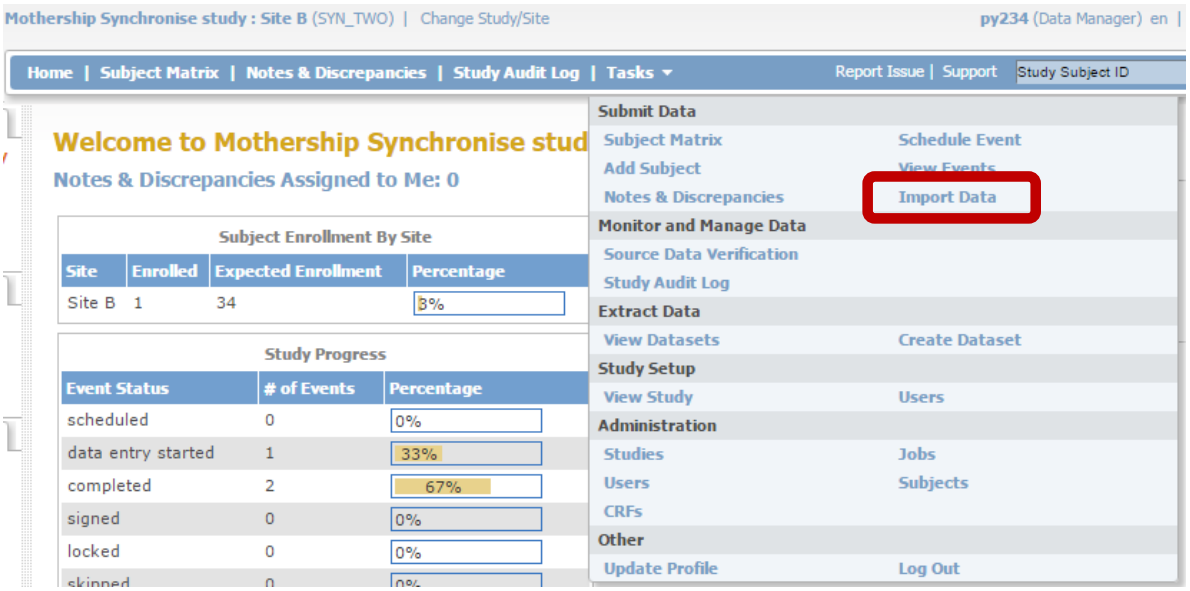


8. Download the XML file:

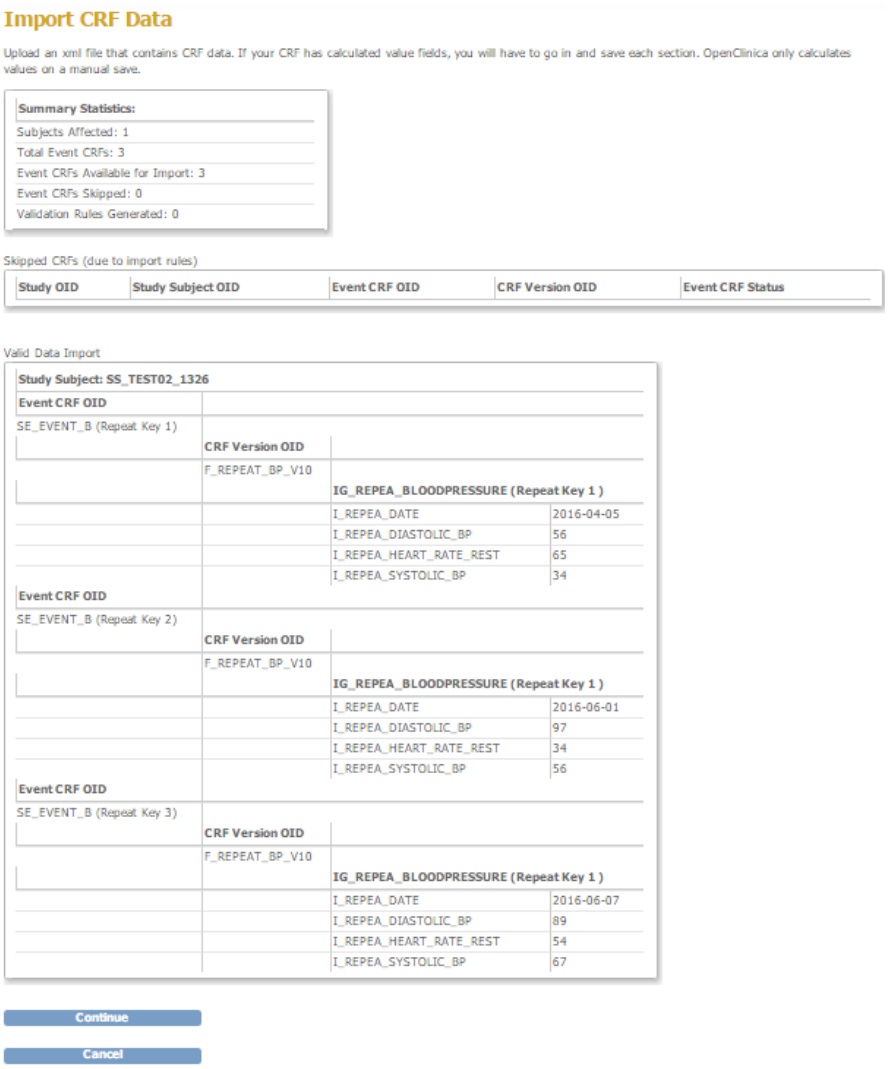


9. Log on to your target OpenClinica, select the correct site to import Data:

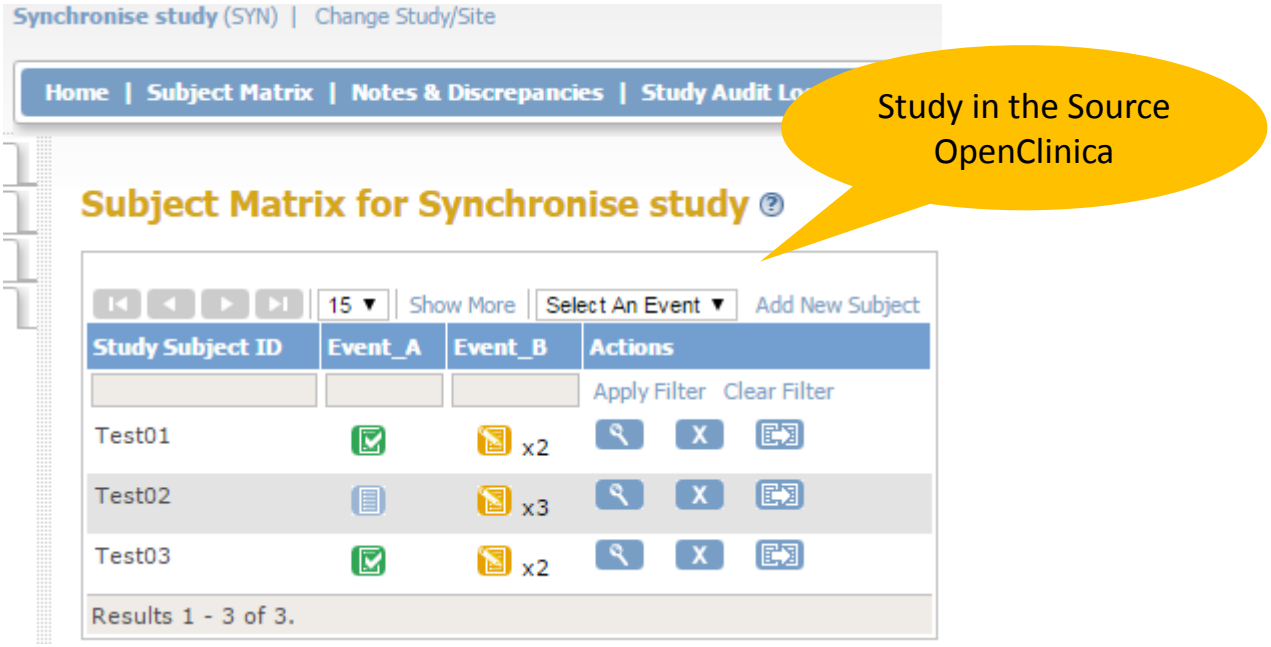


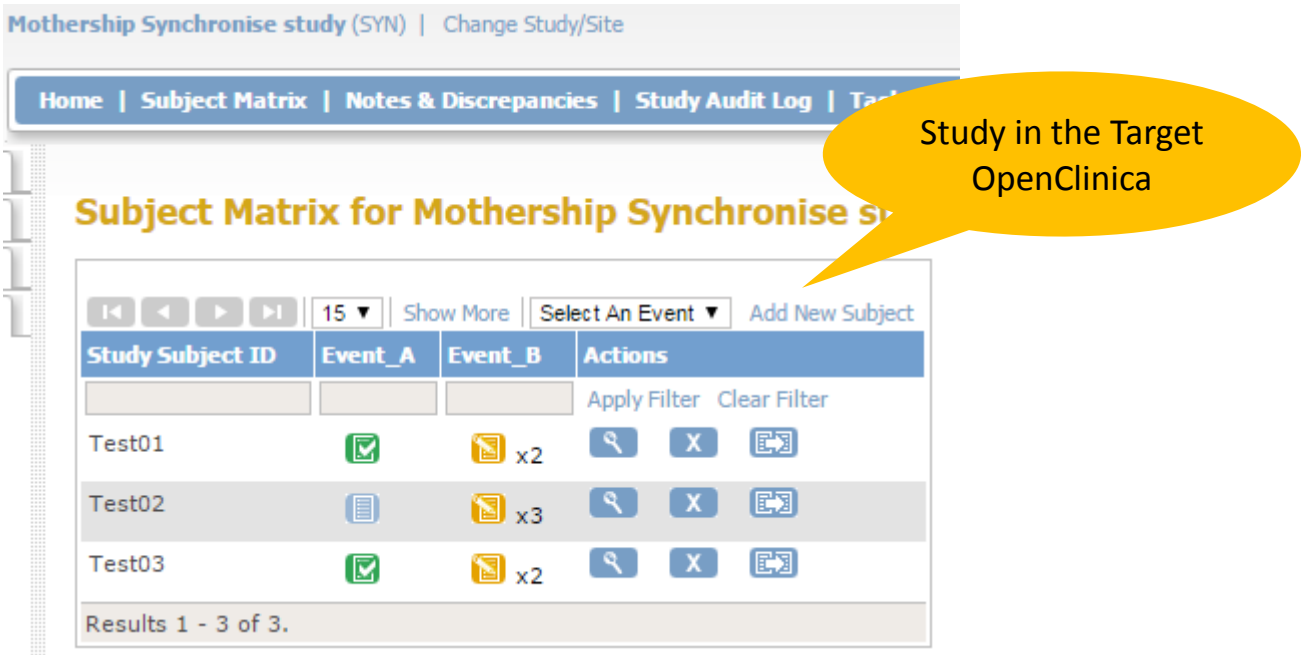


10. Upload the ODM XML file and import data:



Now you should have the study in the target OpenClinica which is the same in the Source OpenClinica.





5 Data Export from OpenClinica

5.1 Create a Dataset

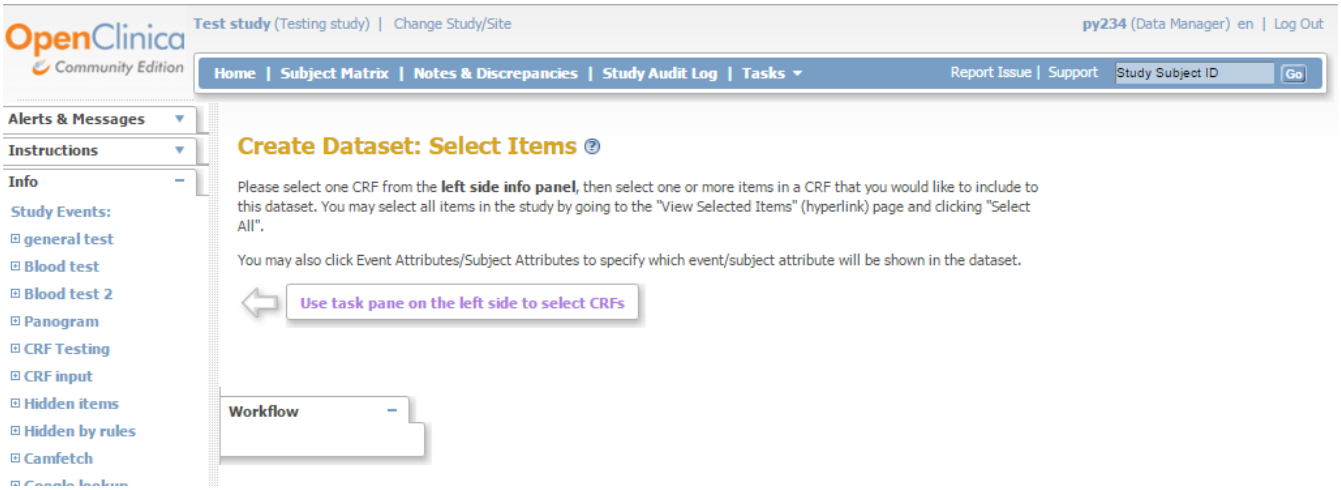
Dataset is a collection of data and metadata from records such as CRFs, Study Events, etc. that matches a dataset definition and is exported to a file whose format can be selected, for use in other applications.

To create a dataset, you first define the dataset (specify the criteria for the dataset) for the current Site or Study. Then you generate and download the file (extract data that matches the definition to a file format you select).

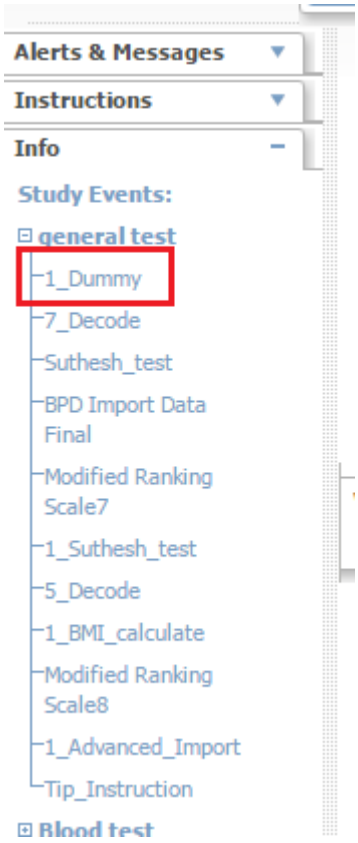
5.1.1 Define the dataset

To define a dataset for the current Study or Site, perform the following steps:

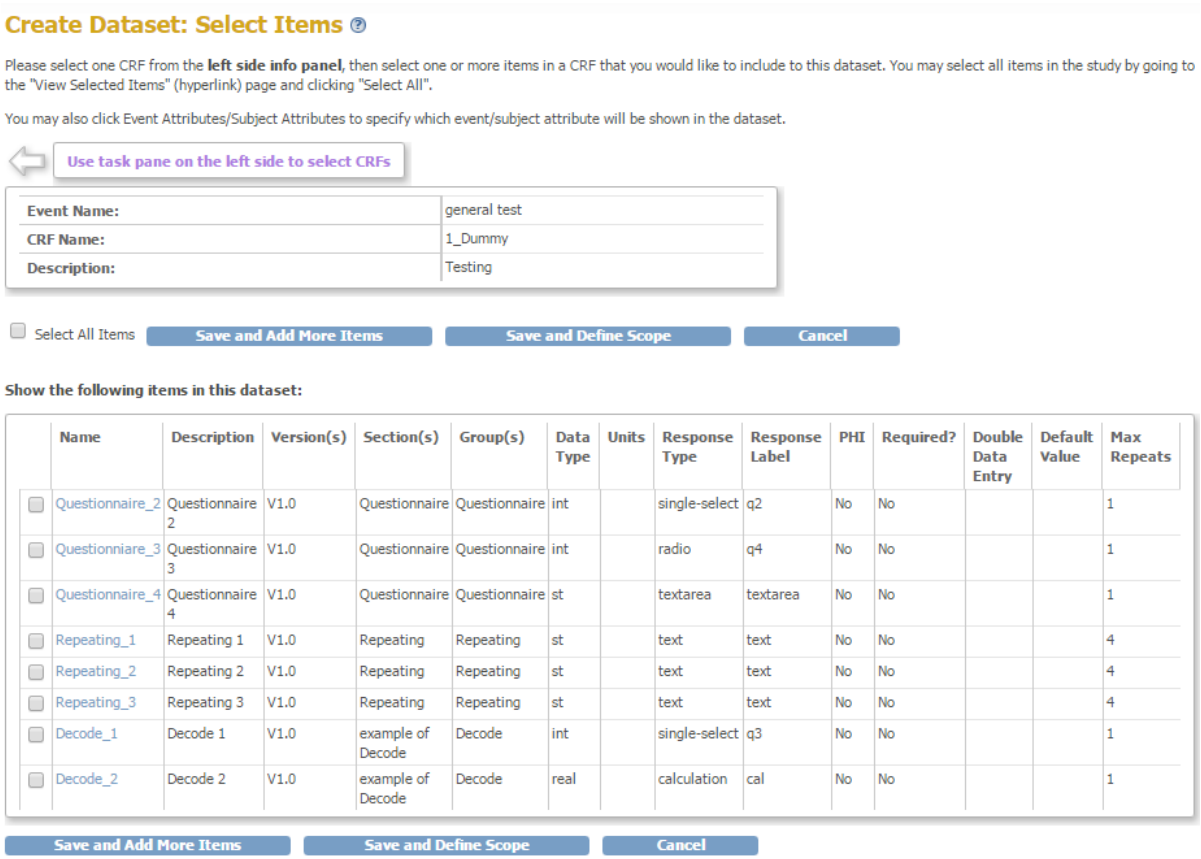
1. Select Tasks > Create Dataset.
   The Create Dataset page for the current Study or Site opens.
2. Click Proceed to Create a Dataset.
   The Create Dataset: Select Items page opens.



3. In the Info sidebar panel, click an Event, and then click a CRF for the Event that you want to include in the dataset definition. For the example shown, General Test Event was selected, and then the *1_Dummy* CRF for the Event was selected. You can add more Event CRFs in a later step.



After clicking an Event CRF, a table displays, listing all of the Event's CRF Items you can include in the dataset definition. It also lists the metadata for each item, which was defined when the CRF was created.



Create Dataset: Select Items

Please select one CRF from the **left side info panel**, then select one or more items in a CRF that you would like to include to this dataset. You may select all items in the study by going to the "View Selected Items" (hyperlink) page and clicking "Select All".

You may also click Event Attributes/Subject Attributes to specify which event/subject attribute will be shown in the dataset.

Use task pane on the left side to select CRFs

Event Name:	general test
CRF Name:	1_Dummy
Description:	Testing

Select All Items | Save and Add More Items | Save and Define Scope | Cancel

Show the following items in this dataset:

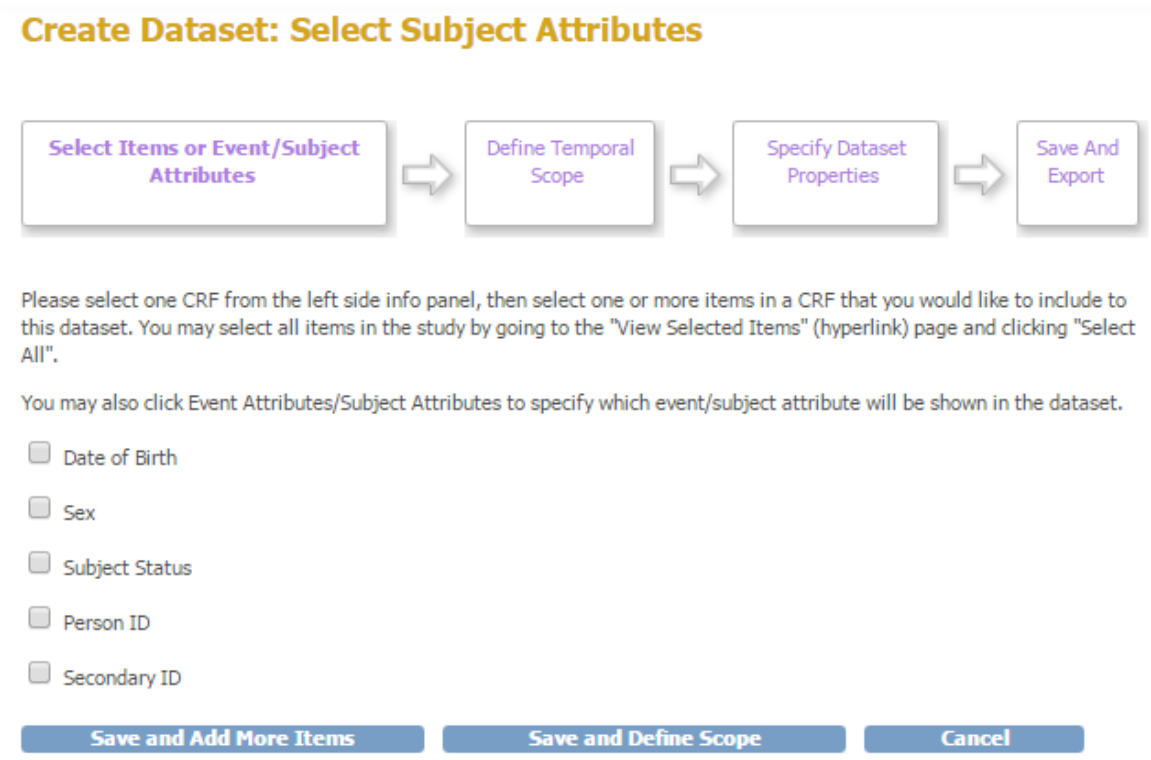
	Name	Description	Version(s)	Section(s)	Group(s)	Data Type	Units	Response Type	Response Label	PHI	Required?	Double Data Entry	Default Value	Max Repeats
☐	Questionnaire_2	Questionnaire 2	V1.0	Questionnaire	Questionnaire	int		single-select	q2	No	No			1
☐	Questionniare_3	Questionnaire 3	V1.0	Questionnaire	Questionnaire	int		radio	q4	No	No			1
☐	Questionnaire_4	Questionnaire 4	V1.0	Questionnaire	Questionnaire	st		textarea	textarea	No	No			1
☐	Repeating_1	Repeating 1	V1.0	Repeating	Repeating	st		text	text	No	No			4
☐	Repeating_2	Repeating 2	V1.0	Repeating	Repeating	st		text	text	No	No			4
☐	Repeating_3	Repeating 3	V1.0	Repeating	Repeating	st		text	text	No	No			4
☐	Decode_1	Decode 1	V1.0	example of Decode	Decode	int		single-select	q3	No	No			1
☐	Decode_2	Decode 2	V1.0	example of Decode	Decode	real		calculation	cal	No	No			1

Save and Add More Items | Save and Define Scope | Cancel

4. In the table, select the checkbox for each Item to include in the dataset definition, or select the Select All Items checkbox. If you are selecting many but not all Items in the CRF, you might find it easier to use Select All Items, then for each item you want to exclude, select its checkbox to clear it.
5. If you want to add more Event CRFs:
    - Click the Save and Add More Items button.
    - Repeat steps 3 to 5 to select the Event CRF Items to include in the dataset definition.
    - To see the Items already selected, click the View Selected Items link in the Info sidebar panel.

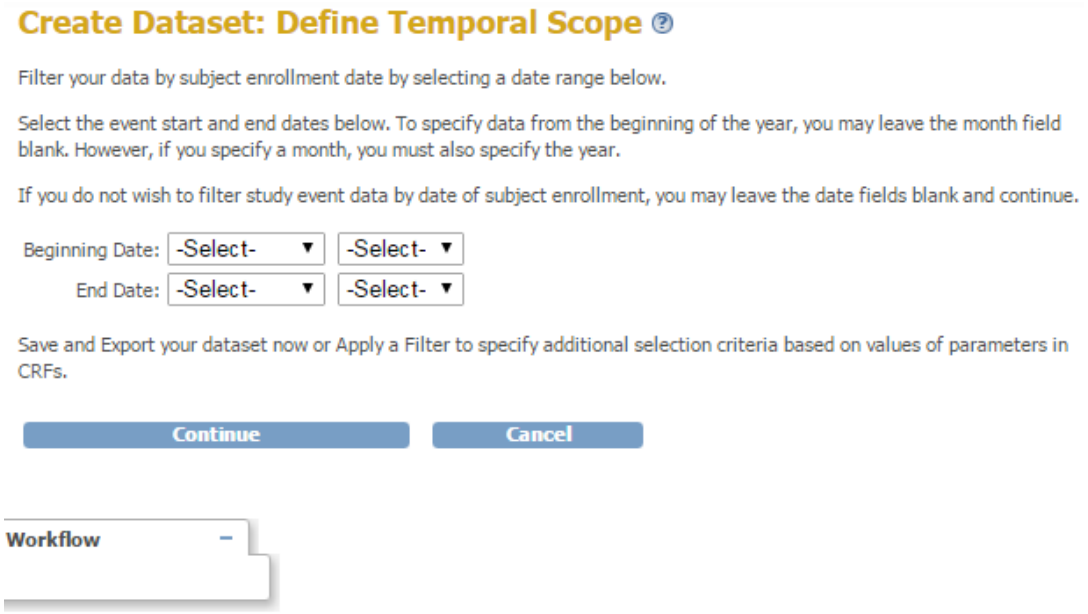
If you want to select all Items for all CRFs for all Events, instead of selecting each Event and CRF, instead click Select All Items in Study in the Info sidebar panel.

6. To select non-CRF Attributes to include in the dataset definition, click the link for it in the Info sidebar panel. The Create Dataset: Select Attributes page opens for the option you chose. For example, when you click Subject Attributes, the Create Dataset: Select Subject Attributes page opens.

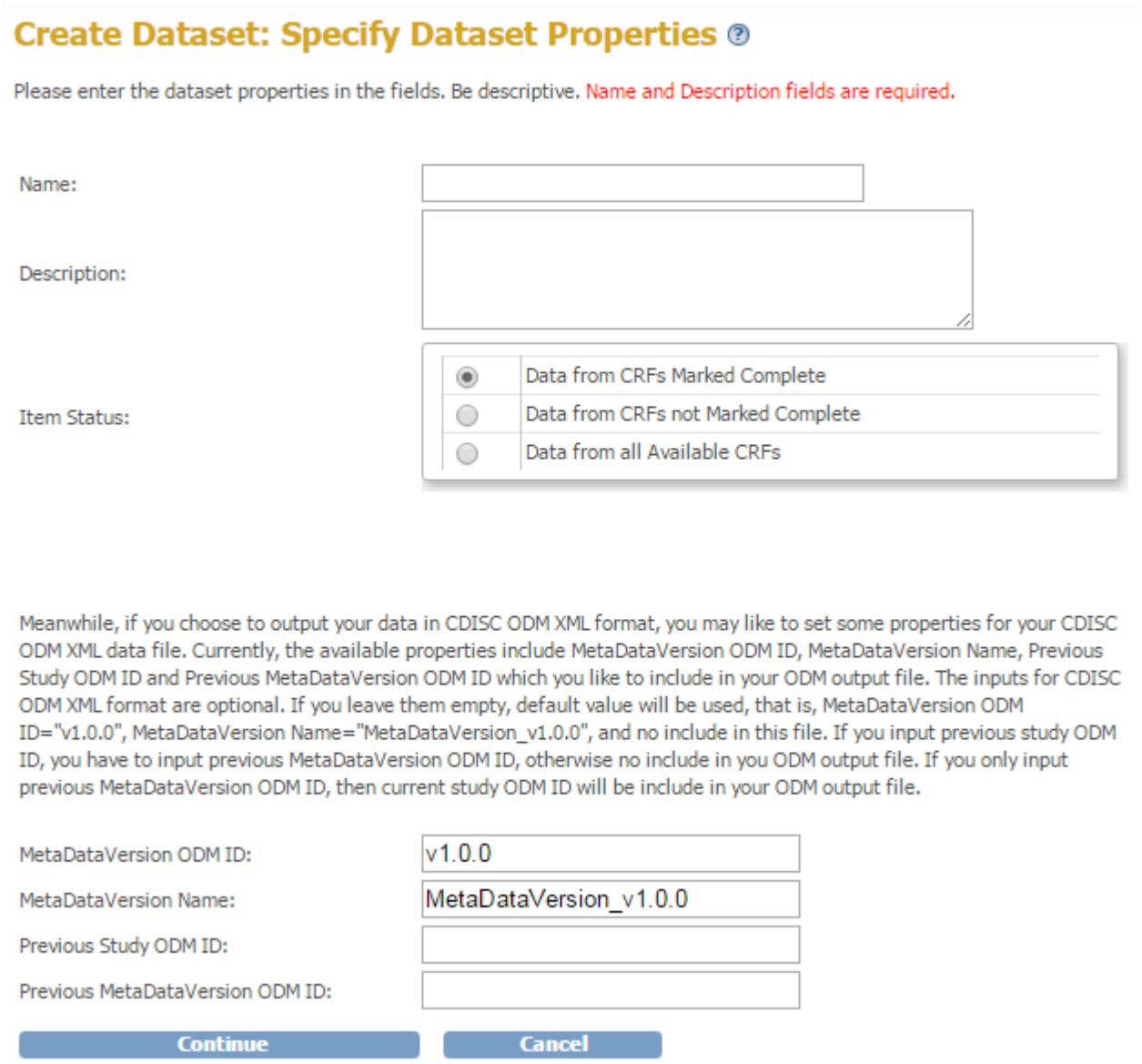


7. Select the checkbox for each Attribute to include.
8. If you want to include Attributes of a different type:
    - Click Save and Add More Items.
    - Repeat steps 6 to 8 to select additional Attributes to include in the dataset definition.
    - To see the Attributes already selected, click the View Selected Items link in the Info sidebar panel.

9. After selecting all Items and Attributes you want to include in the dataset, click the Save and Define Scope button. The Create Dataset: Define Temporal Scope page opens.



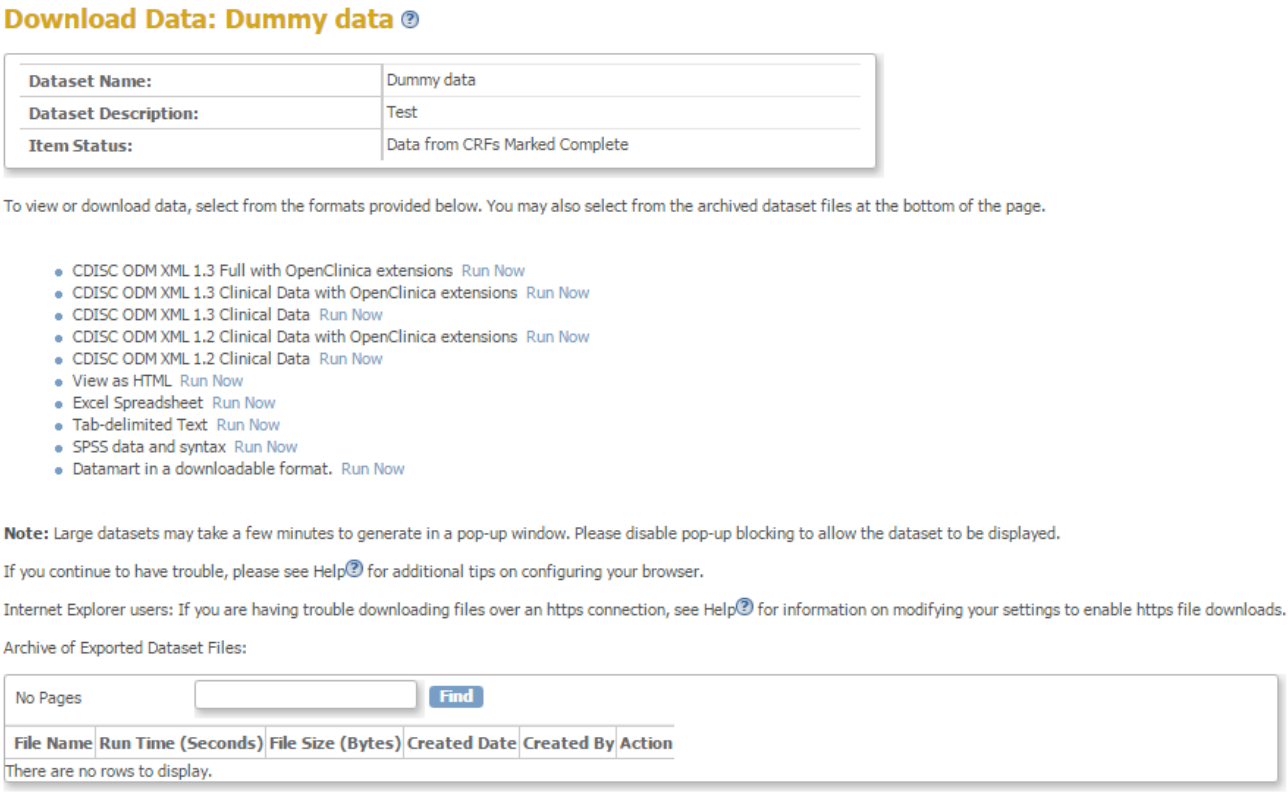
10. Set a date range for the dataset definition by following the onscreen instructions: the date you specify is the Subject enrollment date. For example, if you select 2011 for the year for the Beginning Date and Ending Date, the dataset will include all Subjects who enrolled in 2011. After specifying the date range, click the Continue button.
    The Create Dataset: Specify Dataset Properties page opens.



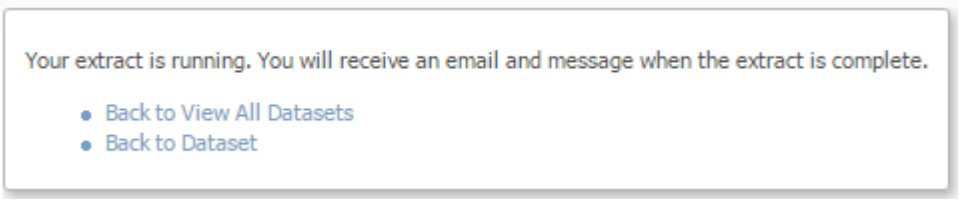
11. Provide a Name and Description for the dataset definition, and optionally, change the Item Status. You can also specify ODM ID values, which you might use if you export the dataset to an ODM file format. Then click Continue. The Create Dataset: Confirm Dataset Properties page opens.
12. Verify the Name and Description, then click 'Confirm' and Save.
    The Download Data page opens. Now that the dataset definition exists, at any time you can generate and download a file containing all data that matches the criteria you specified.

5.1.2 Generate and download Dataset

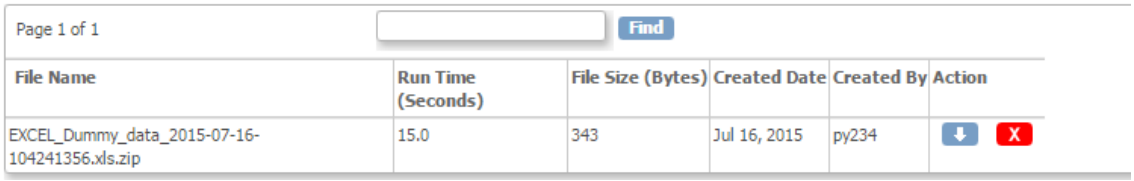
1. After creating, editing, or selecting a dataset definition, the Download Data page displays. On the page, click Run Now for the file format you want.



2. OpenClinica displays a page that reports your extract is running. OpenClinica maintains a queue of dataset extracts, including scheduled jobs and dataset extracts initiated by other users. New data extract requests are added to the end of the queue. You can view the queue to see where your data extract is in the list. After a few moments (or longer, for a large dataset or when there were already other requests in the queue), the dataset file is available. To view it, click Back to Dataset.



3. The Download Data page displays again, and the dataset you generated is listed in the Archive of Exported Dataset Files table. The example below shows a dataset generated with an Excel Spreadsheet format. Click the Download icon in the Actions column for the dataset.



The file downloads using your web browser's download features.

OpenClinica also sends an email to you that contains a link to the dataset file, providing another way for you to access the file, if that option was specified in the datainfo.properties file as part of your system configuration.

4. Open the file to verify that it contains the data you expected. If not, you can edit the dataset definition in OpenClinica and generate the dataset file again.

Data in the dataset file reflects the OpenClinica database at the time the dataset file is generated, not at the time you download the file. The dataset file name includes the date and time the file was generated.

The Archive of Exported Dataset Files table includes only one dataset file for each format for the dataset definition. When you generate a dataset file, if a dataset file having the same format already exists in the table, the file you generate replaces the existing file in the table. If you want both files, download the existing file, and then generate the dataset and download it using a different, or you can edit the dataset definition and save it using a different name and generate a the dataset for it.

6 Reference:

OpenClinica - Data Import & Export, V 1.1, Catherine Titterton. January 2015

ODIN Beta End User Documentation, V 0.8. Csaba Halmagyi. May 2015

OpenClinica User Guide, https://docs.openclinica.com/3.1/openclinica-user-guide/create-dataset