

ODM Generator

version 3.2

User Manual

Date last update: 2022-04-02

Introduction

The ODM Generator is a new tool from XML4Pharma to transform data from "flat files", such as CSV files, exports from Excel and other worksheets, and SAS Transport 5 ("XPT") files into CDISC ODM format. The software can as well be run in Graphical User Interface (GUI) mode as in Command Line Interface (CLI) mode, the latter making it possible to run in batch mode.

Version 3 also includes the "ODM Merger", allowing to combine ODM files generated with the "ODM Generator" into a single ODM file. The user manual for the latter comes as a separate PDF file.

ODM files can be further optimized using any ODM editor software, such as the ["ODM Designer"](#). These CDISC ODM files can then be used in other software tools, such as the popular [SDTM-ETL software for the generation of CDISC SDTM or SEND datasets](#) as well as the corresponding "define.xml" (coordinated approach). The ODM Generator has been developed on the request of SDTM-ETL users who had difficulties using non-ODM files (e.g. files send by laboratories) in the SDTM-ETL software. It enabled to move from a "puzzle of formats" of their clinical data and metadata to a single format that is at the same time the worldwide standard format for clinical data and metadata.

Requirements

The ODM Generator software can be run on any modern computer on which Java (at least version 1.8 or OpenJava version 11) has been installed.

In case of doubt run (double click on Windows) "check-java.bat" or "check-java.sh" (Mac, Linux) in the folder where you installed the software. If Java is installed, this will show something like:

```
C:\WINDOWS\system32\cmd.exe
java version "1.8.0_181"
Java(TM) SE Runtime Environment (build 1.8.0_181-b13)
Java HotSpot(TM) 64-Bit Server VM (build 25.181-b13, mixed mode)

If you see java version "1.8" or above then your Java
version is compatible with the ODMGenerator software.

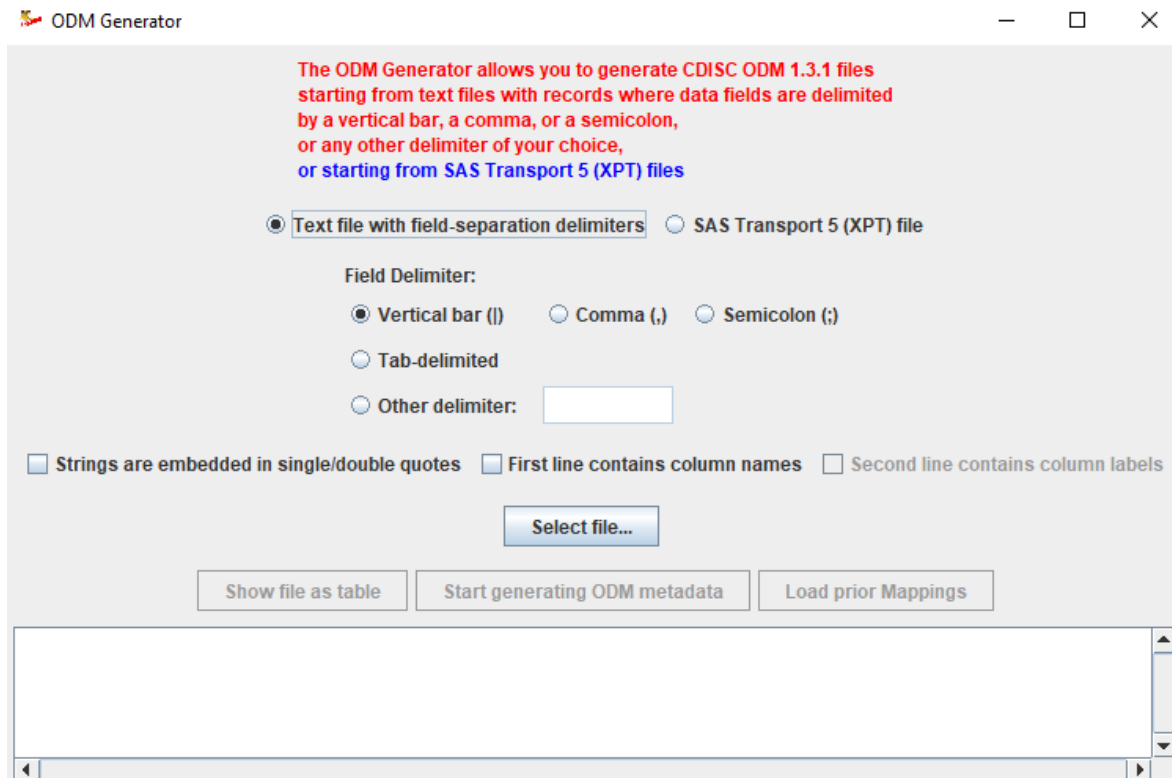
If not, you will need to upgrade your Java Runtime Environment (JRE).
For instructions, please go to http://java.sun.com/
or contact us at info@XML4Pharma.com

You may close this window now.
```

Starting the software (GUI mode) and first steps

Navigate to the folder/directory where you installed the software. On windows, double-click the icon with the name "ODMGenerator.bat". On Linux or Linux-based operating systems, use "ODMGenerator.sh".

The following screen is then displayed:



First case: Source files are "flat text files", such as CSV files

The software allows you to convert "flat text files" to CDISC ODM format, where data fields are separated by a delimiter such as a vertical bar, a comma or semicolon or by a tab ("CSV" files), or any other single character. In case the delimiter itself appears as text in a field, the text in the field (i.e. between the delimiters) should be in straight double quotes (no "skew" quotes). In that case, it is advised that one also checks the checkbox "Strings are embedded in double quotes". The software will normally try to detect whether strings are embedded in double quotes, but this cannot always be guaranteed.

For example, when you export data from an Excel file with the semicolon as the delimiter, cells that have a semicolon in the text will be exported with the cell content between quotes, such as:

a;b;"c;d";e;f

The ODM Generator knows and understands this mechanism and will take appropriate action.

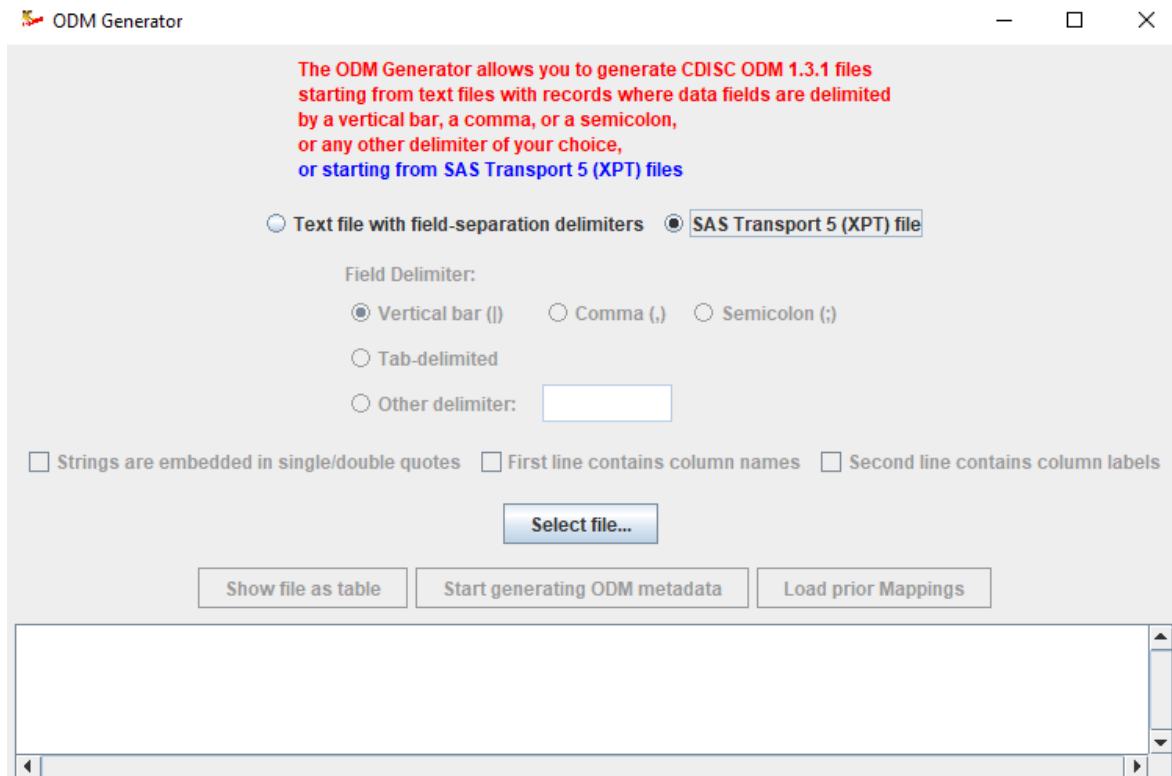
If in the data file, the first row contains the "column names", i.e. the first row should not be considered to be containing data, and should only be used for providing labels, then one should also check the checkbox "First line contains column names". Similarly, if there is a second line with header information, like variable labels, check the checkbox "Second line contains column labels". This checkbox is only enabled when the checkbox "First line contains column names" is checked.

Later, we will see that such variable labels go into the "Description" element in the generated ODM.

Second case: Source files are SAS Transport (XPT) files

New in version 3 is that one can start from a SAS Transport 5 ("XPT") file. Remark that other types of SAS files are not supported as their format is propriety.

If one would start from a SAS Transport 5 file, check the radiobutton "SAS Transport 5 (XPT) file". The fields regarding file delimiters will then be grayed out:



The screenshot shows the ODM Generator application window. At the top, there is a title bar with the text "ODM Generator" and standard window control buttons (minimize, maximize, close). Below the title bar, a red text box contains the following message: "The ODM Generator allows you to generate CDISC ODM 1.3.1 files starting from text files with records where data fields are delimited by a vertical bar, a comma, or a semicolon, or any other delimiter of your choice, or starting from SAS Transport 5 (XPT) files". Below this message, there are two radio buttons: "Text file with field-separation delimiters" (unselected) and "SAS Transport 5 (XPT) file" (selected). Under the "SAS Transport 5 (XPT) file" option, there is a "Field Delimiter:" section with four radio buttons: "Vertical bar (|)" (selected), "Comma (,)", "Semicolon (;)", and "Tab-delimited". Below these, there is an "Other delimiter:" label followed by an empty text input field. At the bottom of the field delimiter section, there are three checkboxes: "Strings are embedded in single/double quotes" (unchecked), "First line contains column names" (unchecked), and "Second line contains column labels" (unchecked). Below the checkboxes, there is a "Select file..." button. At the bottom of the window, there are three buttons: "Show file as table", "Start generating ODM metadata", and "Load prior Mappings". Below these buttons is a large empty text area with a scrollbar on the right side.

Loading your data file

After having selected the delimiter (or having it defined in the field "other delimiter"), or having selected "SAS Transport 5 (XPT) files, load the file that you want to have transformed into CDISC ODM.

In this user manual, we will use a sample "CDISC Lab v.1.0.1" file (see <https://www.cdisc.org/standards/transport/lab>) that uses the vertical bar "|" as delimiter. This example file can be found in the directory "TestFiles" that comes with the software. The file with the first row containing the column names is "Lab1-0-1-BaseSampleData_first_line_field_names.dat", the original one, without column names is "Lab1-0-1-BaseSampleData.dat"

In a later part of the manual, we will also elaborate an example that starts from a SAS Transport 5 (XPT) file. Here is a snapshot for the example "CDISC Lab v.1.0.1:

```

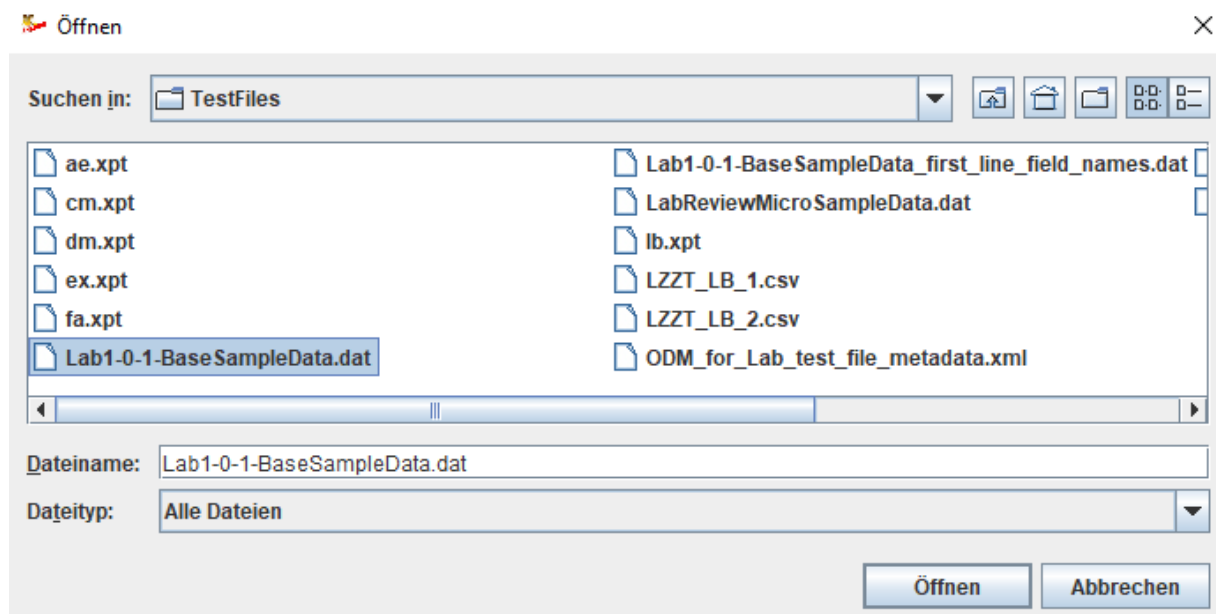
1 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
2 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
3 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
4 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
5 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
6 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
7 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
8 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
9 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
10 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
11 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
12 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
13 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
14 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
15 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
16 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
17 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
18 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
19 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
20 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
21 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|11|11|Joh
22 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|17|17|Sus
23 01-0-01|2003-08-07T14:16:29-05:00|A1234|Central Lab ABC|CDISC Test 1|CDISC Test 1|C|17|17|Sus

```

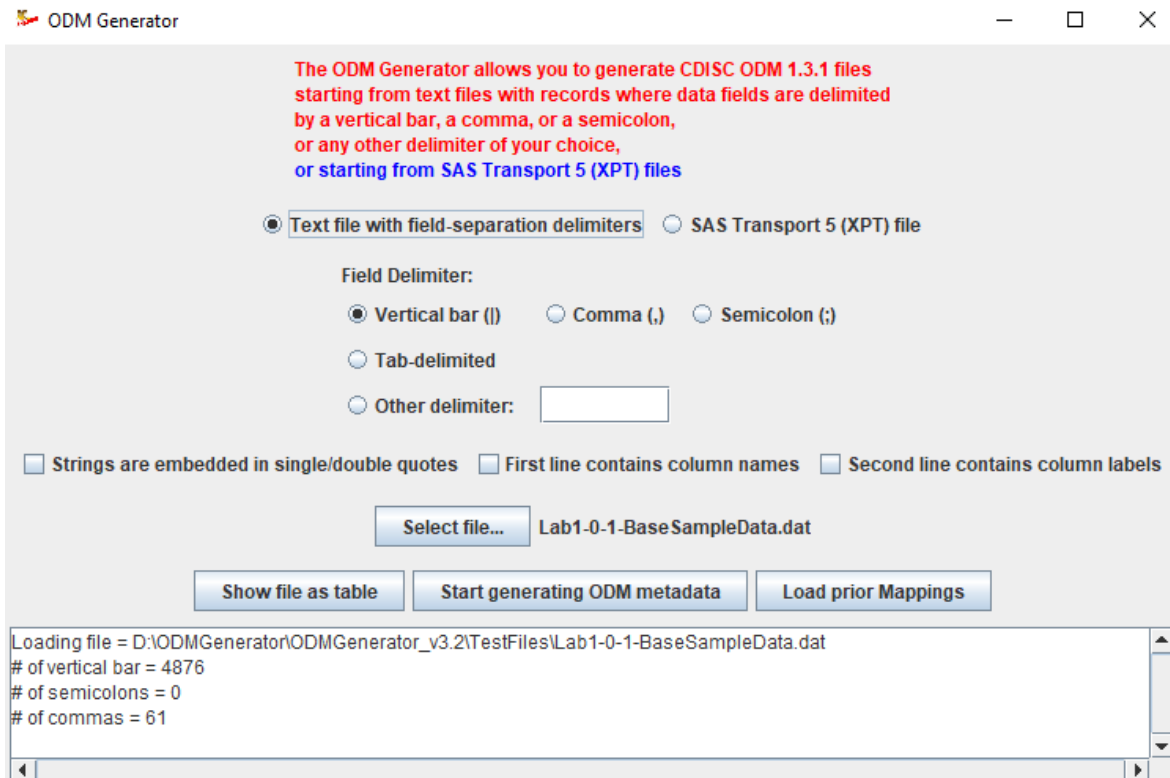
For a description of the fields, see <https://www.cdisc.org/standards/transport/lab>.

The application is not at all limited to Lab files, any type of "flat text" files with fields separated by a delimiter, including exports from Excel files, can be handled.

In order to load a file, use the button "Load file...". A file chooser shows up allowing you to select the file, e.g.:



After clicking "Ok", "Load", or the corresponding expression in your own language (depending on the operating system), the filename of the loaded file is displayed:



The lower part of the window shows the messages from the "logger". Each session is logged into a log file located in the directory "logs". This log file is also the first source when encountering problems or errors. In the latter case, when you contact us, always send the log file of that session with your mail.

Generating metadata

You can now do one of three things:

- a) Show the file as a table for inspection (recommended)
- b) Start generating the ODM metadata
- c) Load a file with earlier defined definitions of ODM metadata for this type of loaded file

When working with a "flat text" file type, you will probably first want to see the data in order to understand what the data is about. In that case, click the button "Show file as table". In our case, this leads to:

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	11	11	John Smith, M.D.	8222			ABC
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	17	17	Susan Jones, M.D.	8277			RST
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	17	17	Susan Jones, M.D.	8277			RST
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	17	17	Susan Jones, M.D.	8277			RST
01-0-01	2003-08-07T14:16:29-05:00	A1234	Central Lab ABC	CDISC Test 1	CDISC Test 1	C	17	17	Susan Jones, M.D.	8277			RST

We can already see that field 2 (F2) represents a date & timestamp, and that field 4 (F4) probably represents the name of the laboratory. We need to look in the specification document for the details however.

If however, the first line in the source file contains the column names, and you have checked the checkbox "First line contains column names", one will see the original column names as a tooltip on the column header. For example:

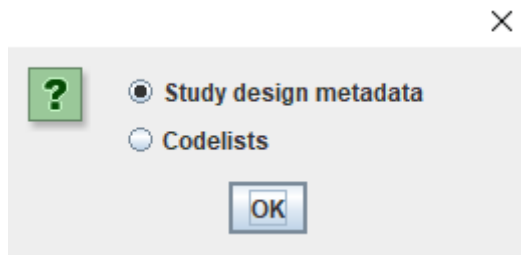
F9	F10	F11
11	John Smith, M.D.	8222
11	John Smith, M.D.	Investigator Name

In case of having loaded a SAS-XPT file, the table will look like:

ID	SALBEG	SEX	TIME	AGE	SALNOW	EDLEVEL	WORK	JOBCAT	MINORITY	SEXRACE
628	8400	0	81	28.5	16080	16	0.25	4	0	1
630	24000	0	73	40.33	41400	16	12.5	5	0	1
632	10200	0	83	31.08	21960	15	4.08	5	0	1
633	8700	0	93	31.17	19200	16	1.83	4	0	1
635	17400	0	83	41.92	28350	19	13	5	0	1
637	12996	0	80	29.5	27250	18	2.42	4	0	1
641	6900	0	79	28	16080	15	3.17	1	0	1
649	5400	0	67	28.75	14100	15	0.5	1	0	1
650	5040	0	96	27.42	12420	15	1.17	1	0	1
652	6300	0	77	52.92	12300	12	26.42	3	0	1
653	6300	0	84	33.5	15720	15	6	1	0	1

When later generating a mapping, it is always wise to keep this window open, as it helps you understand your data, helping you to make the right decisions.

When using the button "Start generating ODM metadata", the system will ask whether you want to generate "usual" metadata (i.e. definitions of visits, forms, ...) or that you want to generate a set of ODM codelists from the tabular data:



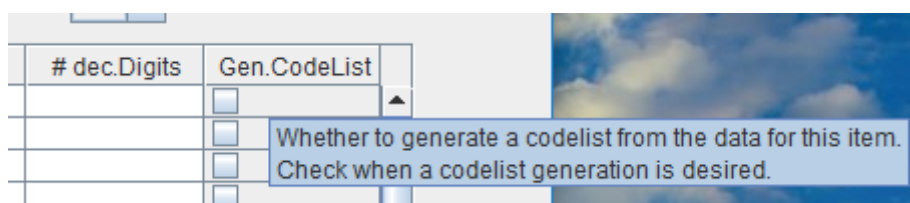
The latter is a special case, and is separately treated in the section "Generating ODM codelists".

We will here however continue with the case of generating "usual" metadata, i.e. definitions of visits, forms, ...

The system will now analyze the data and make proposals for its metadata, especially the datatype (according to the ODM standard) and the maximal field length. In our case with the CDISC-Lab pipe-delimited file, this leads to a new window being opened with the following proposed metadata:

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
F1	<input checked="" type="checkbox"/>	IT.F1	F1		text	7		<input type="checkbox"/>
F2	<input checked="" type="checkbox"/>	IT.F2	F2		datetime			<input type="checkbox"/>
F3	<input checked="" type="checkbox"/>	IT.F3	F3		text	5		<input type="checkbox"/>
F4	<input checked="" type="checkbox"/>	IT.F4	F4		text	15		<input type="checkbox"/>
F5	<input checked="" type="checkbox"/>	IT.F5	F5		text	12		<input type="checkbox"/>
F6	<input checked="" type="checkbox"/>	IT.F6	F6		text	12		<input type="checkbox"/>
F7	<input checked="" type="checkbox"/>	IT.F7	F7		text	1		<input type="checkbox"/>
F8	<input checked="" type="checkbox"/>	IT.F8	F8		integer	2		<input type="checkbox"/>
F9	<input checked="" type="checkbox"/>	IT.F9	F9		integer	2		<input type="checkbox"/>
F10	<input checked="" type="checkbox"/>	IT.F10	F10		text	17		<input type="checkbox"/>
F11	<input checked="" type="checkbox"/>	IT.F11	F11		integer	4		<input type="checkbox"/>
F12	<input checked="" type="checkbox"/>	IT.F12	F12		text	0		<input type="checkbox"/>
F13	<input checked="" type="checkbox"/>	IT.F13	F13		text	0		<input type="checkbox"/>
F14	<input checked="" type="checkbox"/>	IT.F14	F14		text	3		<input type="checkbox"/>
F15	<input checked="" type="checkbox"/>	IT.F15	F15		text	1		<input type="checkbox"/>
F16	<input checked="" type="checkbox"/>	IT.F16	F16		text	33		<input type="checkbox"/>
F17	<input checked="" type="checkbox"/>	IT.F17	F17		date			<input type="checkbox"/>
F18	<input checked="" type="checkbox"/>	IT.F18	F18		text	0		<input type="checkbox"/>
F19	<input checked="" type="checkbox"/>	IT.F19	F19		text	31		<input type="checkbox"/>
F20	<input checked="" type="checkbox"/>	IT.F20	F20		text	3		<input type="checkbox"/>
F21	<input checked="" type="checkbox"/>	IT.F21	F21		text	10		<input type="checkbox"/>
F22	<input checked="" type="checkbox"/>	IT.F22	F22		text	1		<input type="checkbox"/>
F23	<input checked="" type="checkbox"/>	IT.F23	F23		text	1		<input type="checkbox"/>
F24	<input checked="" type="checkbox"/>	IT.F24	F24		text	5		<input type="checkbox"/>
F25	<input checked="" type="checkbox"/>	IT.F25	F25		text	15		<input type="checkbox"/>
F26	<input checked="" type="checkbox"/>	IT.F26	F26		text	7		<input type="checkbox"/>
F27	<input checked="" type="checkbox"/>	IT.F27	F27		datetime			<input type="checkbox"/>
F28	<input checked="" type="checkbox"/>	IT.F28	F28		text	4		<input type="checkbox"/>
F29	<input checked="" type="checkbox"/>	IT.F29	F29		integer	1		<input type="checkbox"/>

Although the column names in this table are pretty self-explanatory, one can always obtain more information by holding the mouse over a column header. For example, for "Gen. CodeList":



The "Generate CodeList" feature is a new feature as of version 2 of the software. For each of the variables, when the checkbox is checked, the data is analyzed and a codelist is generated from those data. For example, for the field 15 (F15), from the specification and the data itself, we see that this represents the sex of the subject, with only two possible values: "F" (female) and "M" (male). If the

"Generate CodeList" checkbox for F15 is checked, a codelist will be generated containing these two values, and will be associated to the item definition¹.

In our example, other fields that come into consideration for assigning a codelist (and that are not set to be excluded) are: "subject race" (field 18), "visit type" (field 22), "subject age units" (field 42), "fasting status" (field 43), and all "flag" fields.

For the first field (F1), the system proposes to assign the datatype "text" with a maximal length of 7 (this is the length of the longest value for this field found in the file). For the second field, it proposes to assign the datatype "datetime". In that case the maximal length should not be assigned.

In case there is no header line in the input file, this is now the right time to at least edit the "Name" cells in the table, providing the information what the input field is about. This can be a bit tedious, but soon, you will be able to save all information to file, and reuse that later.

When the first line of the source file contains the column headers, and you checked the checkbox "First line contains column names", one will see the original column names as a tooltip on each of the first cells in the rows. For example:

F5	<input checked="" type="checkbox"/>		IT.F5
F6	<input checked="" type="checkbox"/>		IT.F6
F7	<input checked="" type="checkbox"/>		IT.F7
F8	<input checked="" type="checkbox"/>		IT.F8
F9	<input checked="" type="checkbox"/>		IT.F9

At the same time, the 4th field, field "Name" will be populated with the column name of the source file when the checkbox "First line contains column names" was checked. For example:



Field for Subject ID: <input type="text" value="NONE"/> Field for visit (StudyEvent): <input type="text" value="NONE"/> Field for form (Form): <input type="text"/>					
Field	Include	OID	Name	Label	Data Type
F1	<input checked="" type="checkbox"/>	IT.F1	Model Version		text
F2	<input checked="" type="checkbox"/>	IT.F2	File Creation Date and Time		datetime
F3	<input checked="" type="checkbox"/>	IT.F3	Transmission Source ID		text
F4	<input checked="" type="checkbox"/>	IT.F4	Transmission Source Name		text
F5	<input checked="" type="checkbox"/>	IT.F5	Study ID or Number		text
F6	<input checked="" type="checkbox"/>	IT.F6	Study Name		text
F7	<input checked="" type="checkbox"/>	IT.F7	Transmission Type		text
F8	<input checked="" type="checkbox"/>	IT.F8	Site ID or Number		integer
F9	<input checked="" type="checkbox"/>	IT.F9	Investigator ID or Number		integer
F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name		text

In case the second line of the input file contained variable labels, and the checkbox "Second line contains variable labels" was checked, also the cells in the column "Label" will be populated.

Remark: providing a label for each variable is always optional. In the generated ODM, the value will go into the "Description" child element of "ItemDef", which essentially defines a variable.

When a SAS Transport 5 file was loaded, "Label" will always be populated, as SAS Transport 5 has the severe limitation that variable names cannot be more than 8 characters, and thus cannot always be self-explanatory, so that also a variable label (also limited to 40 characters) is provided.

¹ As the file does not contain information that "F" means "Female", and "M" means "male", this is information that can later be added by any ODM editor or development software, such as the ["ODM Designer"](#).

In the case of having loaded a SAS-XPT file, this may e.g. look like:

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
ID	<input checked="" type="checkbox"/>	IT.ID	ID	Employee code	integer	4		<input type="checkbox"/>
SALBEG	<input checked="" type="checkbox"/>	IT.SALBEG	SALBEG	Beginning salary	integer	5		<input type="checkbox"/>
SEX	<input checked="" type="checkbox"/>	IT.SEX	SEX	Sex of employee	integer	1		<input type="checkbox"/>
TIME	<input checked="" type="checkbox"/>	IT.TIME	TIME	Job seniority	integer	2		<input type="checkbox"/>
AGE	<input checked="" type="checkbox"/>	IT.AGE	AGE	Age of employee	float	5		<input type="checkbox"/>
SALNOW	<input checked="" type="checkbox"/>	IT.SALNOW	SALNOW	Current salary	integer	5		<input type="checkbox"/>
EDLEVEL	<input checked="" type="checkbox"/>	IT.EDLEVEL	EDLEVEL	Educational level	integer	2		<input type="checkbox"/>
WORK	<input checked="" type="checkbox"/>	IT.WORK	WORK	Work experience	float	19		<input type="checkbox"/>
JOB CAT	<input checked="" type="checkbox"/>	IT.JOB CAT	JOB CAT	Employment category	integer	1		<input type="checkbox"/>
MINORITY	<input checked="" type="checkbox"/>	IT.MINORITY	MINORITY	Minority classification	integer	1		<input type="checkbox"/>
SEX RACE	<input checked="" type="checkbox"/>	IT.SEX RACE	SEX RACE	Sex & race classification	integer	1		<input type="checkbox"/>

CDISC ODM uses "OIDs" which are object identifiers. As a first guess, in case of loading a character-delimited file (so not a SAS Transport 5 file), the identifier "IT.F" followed by the field number is initially assigned, and "Fx" (where x being the field number) is initially assigned for the "Name". If we e.g. know that the second field represents the "file creation datetime", we can make the changes as follows:

Field	Include	OID	Name
F1	<input checked="" type="checkbox"/>	IT.F1	F1
F2	<input checked="" type="checkbox"/>	IT.FileCreationDatetime	File creation datetime
F3	<input checked="" type="checkbox"/>	IT.F3	F3

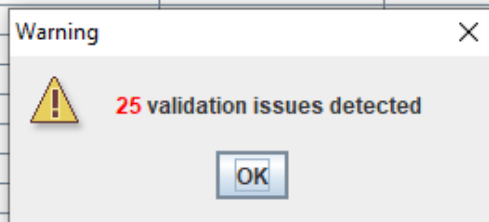
Remark that OIDs are arbitrary and only used as identifiers in references. So, if you want to use "abracadabra" for the OID for the second field, that's OK. I personally prefer to assign something meaningful to OIDs. Please just take into account that having blanks in OIDs is considered bad practice. One can e.g. replace such by an underscore "_".

In the above case, one can of course also populate "Label" manually, but as we already changed the "Name" into something that well describes what the field is about, this is usually not necessary. Having labels is mostly very useful when data from a SAS Transport 5 file is loaded.

In case you need to process different files, e.g. originating from different tabs in an Excel file, or from a set of SAS-XPT files, it is also a good idea to edit/rename the OIDs, as each OID in ODM needs to have a unique meaning. So, having "IT.F9" for different file mappings, may later lead to difficulties when wanting to merge generated ODM files (see section "Merging ODM files"). So, for example, when one has loaded an XPT file with vital signs, one can best assign OIDs like "IT.VS.1", "IT.VS.2" etc..

You can now start assigning metadata by editing the table. As there are certain rules in ODM, you can always do a validation on your assignments using the "Validate" button. In our example this will lead to a warning:

Field	Include	OID	Name	Label
F1	<input checked="" type="checkbox"/>	IT.F1	F1	
F2	<input checked="" type="checkbox"/>	IT.FileCreatiomDatetime	File creation datetime	
F3	<input checked="" type="checkbox"/>	IT.F3	F3	
F4	<input checked="" type="checkbox"/>	IT.F4		
F5	<input checked="" type="checkbox"/>	IT.F5		
F6	<input checked="" type="checkbox"/>	IT.F6		
F7	<input checked="" type="checkbox"/>	IT.F7		
F8	<input checked="" type="checkbox"/>	IT.F8		
F9	<input checked="" type="checkbox"/>	IT.F9		
F10	<input checked="" type="checkbox"/>	IT.F10		
F11	<input checked="" type="checkbox"/>	IT.F11	F11	
F12	<input checked="" type="checkbox"/>	IT.F12	F12	



After clicking "OK", the assignments for which there is an issue will be highlighted, and a tooltip with give some details. For example:

F10		text	33	
F17		date		
F18		text	0	
F19		text	31	
F20		text	3	Length must be a positive integer
F21		text	10	
F22		text	1	

In this case, the underlying reason is that there were no values at all for field 19 in the file, so a maximum length could not be assigned. CDISC ODM however always want us to have a maximum length assigned when the datatype is "text"².

One can now edit cells to which an error is assigned by a simple click. For fields that have no data, one should think about whether these should be included anyway in the output. If not, one should then uncheck the checkbox "Include". Also see the section "Filtering".

Storing generated/developed metadata to file

We can always store our assignments to a file for later use and reuse. So, these assignments need to be done only once for a specific file type, and can be reused over and over again. When clicking the button "**Save mappings**", a file chooser will pop up to choose a file allowing you to save everything to a text file. In our case, the contents of this file look like:

² This has to do with the ODM metadata being used as a specification for relational database table, i.e. assigning the VARCHAR(n) datatype.

```

1 NumFields=92
2 SkipFirstLine=false
3 Field=F1 Include=true OID="IT.F1" Name="F1" DataType=text Length=7 GenerateCodeList=false
4 Field=F2 Include=true OID="IT.FileCreationDateTime" Name="File creation datetime" DataType=datetime GenerateCodeList=false
5 Field=F3 Include=true OID="IT.F3" Name="F3" DataType=text Length=5 GenerateCodeList=false
6 Field=F4 Include=true OID="IT.F4" Name="F4" DataType=text Length=15 GenerateCodeList=false
7 Field=F5 Include=true OID="IT.F5" Name="F5" DataType=text Length=12 GenerateCodeList=false
8 Field=F6 Include=true OID="IT.F6" Name="F6" DataType=text Length=12 GenerateCodeList=false
9 Field=F7 Include=true OID="IT.F7" Name="F7" DataType=text Length=1 GenerateCodeList=false
10 Field=F8 Include=true OID="IT.F8" Name="F8" DataType=integer Length=2 GenerateCodeList=false
11 Field=F9 Include=true OID="IT.F9" Name="F9" DataType=integer Length=2 GenerateCodeList=false
12 Field=F10 Include=true OID="IT.F10" Name="F10" DataType=text Length=17 GenerateCodeList=false
13 Field=F11 Include=true OID="IT.F11" Name="F11" DataType=integer Length=4 GenerateCodeList=false
14 Field=F12 Include=true OID="IT.F12" Name="F12" DataType=text Length=0 GenerateCodeList=false
15 Field=F13 Include=true OID="IT.F13" Name="F13" DataType=text Length=0 GenerateCodeList=false
16 Field=F14 Include=true OID="IT.F14" Name="F14" DataType=text Length=3 GenerateCodeList=false
17 Field=F15 Include=true OID="IT.F15" Name="F15" DataType=text Length=1 GenerateCodeList=true
18 Field=F16 Include=true OID="IT.F16" Name="F16" DataType=text Length=33 GenerateCodeList=false
19 Field=F17 Include=true OID="IT.F17" Name="F17" DataType=date GenerateCodeList=false
20 Field=F18 Include=true OID="IT.F18" Name="F18" DataType=text Length=0 GenerateCodeList=false
21 Field=F19 Include=true OID="IT.F19" Name="F19" DataType=text Length=31 GenerateCodeList=false
22 Field=F20 Include=true OID="IT.F20" Name="F20" DataType=text Length=3 GenerateCodeList=false
23 Field=F21 Include=true OID="IT.F21" Name="F21" DataType=text Length=10 GenerateCodeList=false

```

It has a very simple format, so can also be edited offline, outside the tool.

In our case, the sample file is a "flat text" file that obtains the "CDISC Lab v.1.0.1" standard. In your case, this may be a company-internal standard, or a standardized format from your provider.

We will now continue with a "mappings" file where all the mappings with "CDISC Lab v.1.0.1" standard. As already stated, these assignments can either be done within the tool (followed by saving the mappings to file), or by editing an exported "mappings file".

The example "mappings file" for this example can be found in the distribution at:

Mappings/Lab_1-0-1_example_mappings.dat

In our case, our file with assignments and mappings looks like:

```

Field=F1 Include=false OID="IT.F1" Name="Model Version" DataType=text Length=7 GenerateCodeList=false
Field=F2 Include=false OID="IT.F2" Name="File Creation Date and Time" DataType=datetime GenerateCodeList=false
Field=F3 Include=false OID="IT.F3" Name="Transmission Source ID" DataType=text Length=5 GenerateCodeList=false
Field=F4 Include=false OID="IT.F4" Name="Transmission Source Name" DataType=text Length=15 GenerateCodeList=false
Field=F5 Include=true OID="IT.F5" Name="Study ID or Number" DataType=text Length=12 GenerateCodeList=false
Field=F6 Include=false OID="IT.F6" Name="Study Name" DataType=text Length=12 GenerateCodeList=false
Field=F7 Include=false OID="IT.F7" Name="Transmission Type" DataType=text Length=1 GenerateCodeList=false
Field=F8 Include=true OID="IT.F8" Name="Site ID or Number" DataType=integer Length=2 GenerateCodeList=false
Field=F9 Include=true OID="IT.F9" Name="Investigator ID or Number" DataType=integer Length=2 GenerateCodeList=false
Field=F10 Include=true OID="IT.F10" Name="Investigator Name" DataType=text Length=17 GenerateCodeList=false
Field=F11 Include=true OID="IT.F11" Name="Screen ID or Number" DataType=integer Length=4 GenerateCodeList=false
Field=F12 Include=false OID="IT.F12" Name="Subject ID or Number" DataType=text Length=0 GenerateCodeList=false
Field=F13 Include=false OID="IT.F13" Name="Spare subject level ID or Number" DataType=text Length=0 GenerateCodeList=false
Field=F14 Include=false OID="IT.F14" Name="Subject Initials" DataType=text Length=3 GenerateCodeList=false
Field=F15 Include=true OID="IT.F15" Name="Subject Sex" DataType=text Length=1 GenerateCodeList=true
Field=F16 Include=false OID="IT.F16" Name="Subject Sex Code List ID" DataType=text Length=33 GenerateCodeList=false
Field=F17 Include=true OID="IT.F17" Name="Subject Date Of Birth" DataType=date GenerateCodeList=false
Field=F18 Include=false OID="IT.F18" Name="Subject Race" DataType=text Length=0 GenerateCodeList=true
Field=F19 Include=false OID="IT.F19" Name="Subject Race Code List ID" DataType=text Length=31 GenerateCodeList=false
Field=F20 Include=true OID="IT.F20" Name="Visit ID or Number" DataType=text Length=3 GenerateCodeList=false
Field=F21 Include=false OID="IT.F21" Name="Visit Name" DataType=text Length=10 GenerateCodeList=false
Field=F22 Include=false OID="IT.F22" Name="Visit Type" DataType=text Length=1 GenerateCodeList=true

```

Reusing mappings

This file with mappings can now be loaded (or reloaded) by the software using the button "Load prior mappings":

The ODM Generator allows you to generate CDISC ODM 1.3.1 files starting from text files with records where data fields are delimited by a vertical bar, a comma, or a semicolon, or any other delimiter of your choice, or starting from SAS Transport 5 (XPT) files

Text file with field-separation delimiters
 SAS Transport 5 (XPT) file

Field Delimiter:

Vertical bar (|)
 Comma (,)
 Semicolon (;)

Tab-delimited

Other delimiter:

Strings are embedded in single/double quotes
 First line contains column names
 Second line contains column labels

Select file... Lab1-0-1-Base SampleData.dat

Which will display our assignments as:

Field for Subject ID: F11 Field for visit (StudyEvent): F20 Field for form (Form): NONE

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
F1	<input type="checkbox"/>	IT.F1	Model Version		text	7		<input type="checkbox"/>
F2	<input type="checkbox"/>	IT.F2	File Creation Date a...		datetime			<input type="checkbox"/>
F3	<input type="checkbox"/>	IT.F3	Transmission Sour...		text	5		<input type="checkbox"/>
F4	<input type="checkbox"/>	IT.F4	Transmission Sour...		text	15		<input type="checkbox"/>
F5	<input checked="" type="checkbox"/>	IT.F5	Study ID or Number		text	12		<input type="checkbox"/>
F6	<input type="checkbox"/>	IT.F6	Study Name		text	12		<input type="checkbox"/>
F7	<input type="checkbox"/>	IT.F7	Transmission Type		text	1		<input type="checkbox"/>
F8	<input checked="" type="checkbox"/>	IT.F8	Site ID or Number		integer	2		<input type="checkbox"/>
F9	<input checked="" type="checkbox"/>	IT.F9	Investigator ID or Nu...		integer	2		<input type="checkbox"/>
F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name		text	17		<input type="checkbox"/>
F12	<input type="checkbox"/>	IT.F12	Subject ID or Number		text	0		<input type="checkbox"/>
F13	<input type="checkbox"/>	IT.F13	Spare subject level l...		text	0		<input type="checkbox"/>
F14	<input type="checkbox"/>	IT.F14	Subject Initials		text	3		<input type="checkbox"/>
F15	<input checked="" type="checkbox"/>	IT.F15	Subject Sex		text	1		<input checked="" type="checkbox"/>
F16	<input type="checkbox"/>	IT.F16	Subject Sex Code Li...		text	33		<input type="checkbox"/>
F17	<input checked="" type="checkbox"/>	IT.F17	Subject Date Of Birth		date			<input type="checkbox"/>
F18	<input type="checkbox"/>	IT.F18	Subject Race		text	0		<input checked="" type="checkbox"/>
F19	<input type="checkbox"/>	IT.F19	Subject Race Code ...		text	31		<input type="checkbox"/>
F21	<input type="checkbox"/>	IT.F21	Visit Name		text	10		<input type="checkbox"/>
F22	<input type="checkbox"/>	IT.F22	Visit Type		text	1		<input checked="" type="checkbox"/>
F23	<input type="checkbox"/>	IT.F23	Visit Type Modifier		text	1		<input type="checkbox"/>
F24	<input checked="" type="checkbox"/>	IT.F24	Central Laboratory ID		text	5		<input type="checkbox"/>
F25	<input type="checkbox"/>	IT.F25	Central Laboratory ...		text	15		<input type="checkbox"/>
F26	<input checked="" type="checkbox"/>	IT.F26	Accession ID or Nu...		text	7		<input type="checkbox"/>
F27	<input type="checkbox"/>	IT.F27	Last Active Date and...		datetime			<input type="checkbox"/>
F28	<input type="checkbox"/>	IT.F28	Record Extension T...		text	4		<input type="checkbox"/>
F29	<input checked="" type="checkbox"/>	IT.F29	Specimen ID or Nu...		integer	1		<input type="checkbox"/>
F30	<input checked="" type="checkbox"/>	IT.F30	Actual Collection Da...		datetime			<input type="checkbox"/>
F31	<input type="checkbox"/>	IT.F31	Planned Collection ...		text	9		<input type="checkbox"/>

Showing that some for some fields (15, 18, 22), a codelist will be generated from the data, and that for some fields (1-4, 6-7, ...) the information will not be included in the ODM file later. See the section "Filtering" for further information. The choice which to have a codelist generated for, is of course yours!

It is always a good idea to use the "Validate" button at this time.

As already stated, you will only need to develop these assignments only once for each "flat text" file type in your organization.

Reuse of mappings will later also be of importance when generating clinical data, e.g. in the case the input file was updated with new data. This also means that you already want to develop your mappings even before all data are available. Generating clinical data can later also be done in "batch" mode, see section "Running the software in batch mode".

Assigning fields for visit ID or number, form ID or number, and subform ID or number

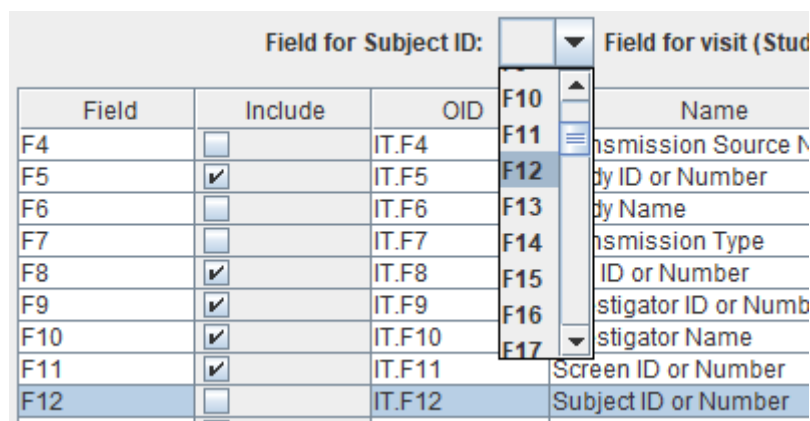
ODM organizes the data per visit (in ODM named "StudyEvent"), per form, and per subform (named "ItemGroup"). We now need to assign (some of these) from the fields in the source file.

Adding this information essentially allows to go from a flat 2-dimensional representation of the data (our source data) to a multidimensional representation: the CDISC ODM format.

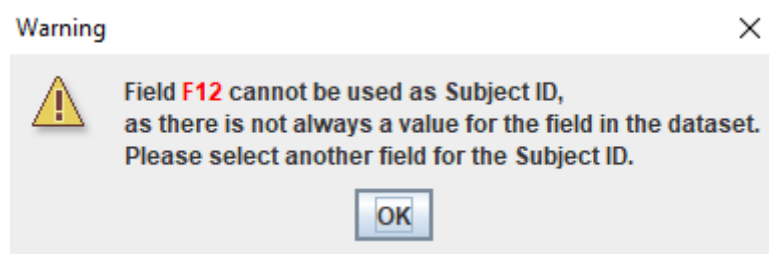
First of all, we must assign one of the fields to the "Subject ID". If this is not done, it will not be possible to generate an ODM file. In our lab data example, the Subject ID can be found in field 12:

F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name	text	17
F11	<input checked="" type="checkbox"/>	IT.F11	Screen ID or Number	integer	4
F12	<input type="checkbox"/>	IT.F12	Subject ID or Number	text	0
F13	<input type="checkbox"/>	IT.F13	Spare subject level ID or Number	text	0
F14	<input type="checkbox"/>	IT.F14	Subject Initials	text	3

So, in the graphical user interface, we select "F12" for the field to be used for the Subject ID:



We do however get a warning:



Telling us that there is not always data in this field.

Inspection of the data reveals that only field 11 ("Screen ID") is always populated, field 12 isn't:



F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	11	11	John S...	8222			ABC	M
CDISC...	CDISC...	C	17	17	Susan...	8277			RST	F
CDISC...	CDISC...	C	17	17	Susan...	8277			RST	F
CDISC...	CDISC...	C	17	17	Susan...	8277			RST	F

So it looks here as the "Screen ID" was used as the subject identifier, and the "Subject ID" was left empty. So, we will use field 11 for the identifier of the subject (Subject ID):

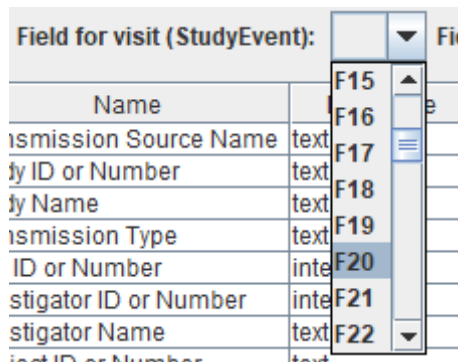
Field for Subject ID:			Field for visit (StudyEvent):
Field	Include	OID	Name
F4	<input type="checkbox"/>	IT.F4	
F5	<input checked="" type="checkbox"/>	IT.F5	
F6	<input type="checkbox"/>	IT.F6	Transmission Source Name
F7	<input type="checkbox"/>	IT.F7	Study ID or Number
F8	<input checked="" type="checkbox"/>	IT.F8	Study Name
F9	<input checked="" type="checkbox"/>	IT.F9	Transmission Type
F10	<input checked="" type="checkbox"/>	IT.F10	Site ID or Number
F11	<input checked="" type="checkbox"/>	IT.F11	Investigator ID or Number
F12	<input type="checkbox"/>	IT.F12	Investigator Name
F13	<input type="checkbox"/>	IT.F13	Screen ID or Number
F14	<input type="checkbox"/>	IT.F14	Subject ID or Number
F15	<input type="checkbox"/>	IT.F15	Subject Sex

When done, we will notice that the row for F11 disappears from the table, as it is "promoted" to be the first level organizer in the data:

F8	<input checked="" type="checkbox"/>	IT.F8	Site ID or Number
F9	<input checked="" type="checkbox"/>	IT.F9	Investigator ID or Number
F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name
F12	<input type="checkbox"/>	IT.F12	Subject ID or Number
F13	<input type="checkbox"/>	IT.F13	Spare subject level ID or Number
F14	<input type="checkbox"/>	IT.F14	Subject Initials
F15	<input checked="" type="checkbox"/>	IT.F15	Subject Sex

Secondly, we want to assign a field for the "visit ID or number". If we don't, all the data will be assigned to a single, "default" visit. In some seldom cases, where there was only one visit, this can indeed be the case. In the CDISC Lab Standard however, field 20 "Visit ID or number" is however

exact meant for storing this information. So, in the graphical user interface, we select "F20" for the "Field for visit (StudyEvent)":

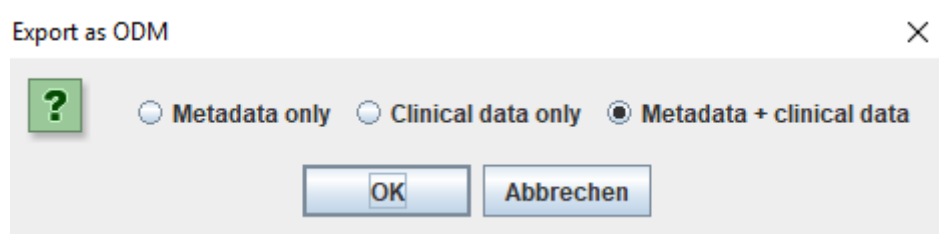


Also here, we then find that the row F20 disappears from the table, as it has been "promoted".

We would now also need to assign a field for the "form" and for the "subform" (ItemGroup). However, we do not find a suitable field for these! The reason is simply that all data are supposed to come from a single form, the "laboratory form"³. The same applies for the "subform". So we will leave the assignments for these unused. The software will then later generate a single default form and subform in the metadata, which we will be enabled to give an identifier and a name.

After assign a field for the subject ID and for the visit ID, we are now ready to generate the ODM file. We can first save the mappings using the "Save Mappings" button, so that all this information can be reloaded later, e.g. when new, additional data comes in and we either want to revisit the process and mappings, or generate new or additional clinical data in ODM format, for which we need this information.

After clicking the "Export as ODM" button, the following dialog appears:



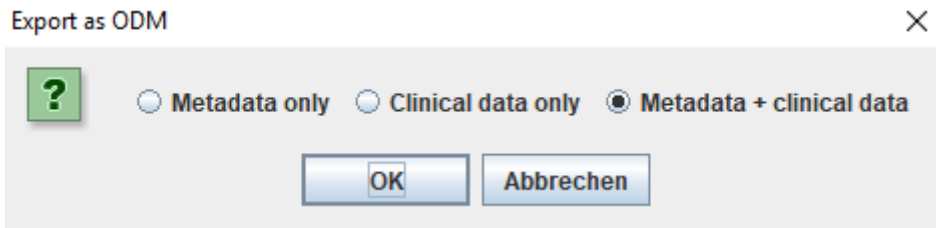
Remark that the software will first do a validation. If issues are reported, one may choose to correct these first, or ignore these (which essentially might lead to a non-conform ODM file), and continue.

In our case, we will generate as well metadata as clinical data. The selection will usually depend on what one will use the ODM file for. When later new data comes in, one can use the mappings again and then e.g. only export clinical data. So, saving the mappings is "highly recommended"

Exporting to CDISC ODM

You can now start exporting the data and metadata as CDISC ODM using the "Export as ODM" button. The following dialog is displayed:

³ A possibility here would be to assign field 44 ("battery ID") to "Form". This would however generate different forms for "Urinalysis", "Urine Drug Screen", "Hematology", etc. Of course this can be a good idea, when such a split up is desired.

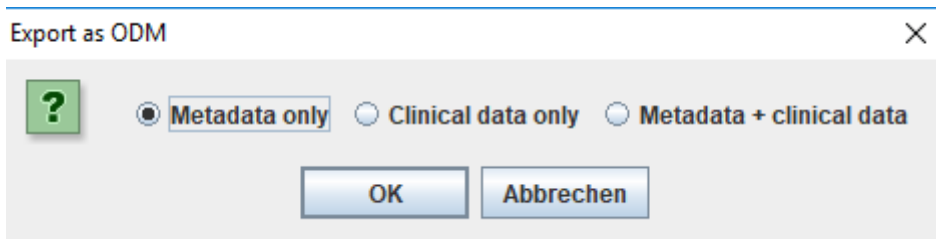


Depending on what you would like to do with the ODM data, you may choose between:

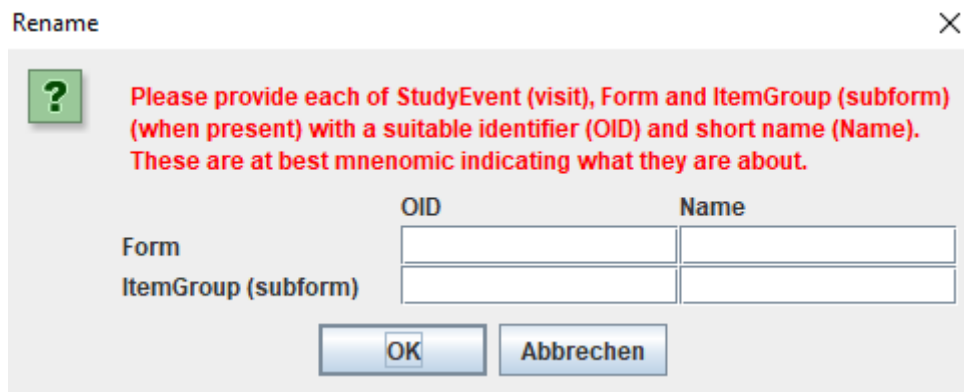
- a) Metadata only: the ODM file will be a file with study design information only
- b) Clinical data only: the data from your "flat text" file will be converted to CDISC "ODM ClinicalData" without the metadata. This is an option that is recommended only in the case that you have generated a "metadata" file separately.
- c) Metadata + clinical data: An ODM file will be generated containing as well the data as metadata (this is never a bad choice)

Remark that the software will first do a validation. If issues are reported, one may choose to correct these first, or ignore these (which essentially might lead to a non-conform ODM file), and continue.

Let us first generate a file with "metadata only":



After clicking "OK", you will first need to add some extra information:



As we did not assign fields for the form and subform levels, the software will generate a single "default" form, and a single "default" subform. We are now asked to provide an identifier (OID) and name for these. It is recommended to use mnemonic values for this. As our data is about lab data, we e.g. use "FO.LABORATORY" and "Laboratory form", and "IG.LABORATORY" and "Laboratory subform" for these, where "FO." and "IG." are pretty usual prefixes used for the identifiers in ODM. So we enter:

Rename ×

? Please provide each of StudyEvent (visit), Form and ItemGroup (subform) (when present) with a suitable identifier (OID) and short name (Name). These are at best mnemonic indicating what they are about.

	OID	Name
Form	FO.LABORATORY	Laboratory form
ItemGroup (subform)	IG.LABORATORY	Laboratory subform

When transforming different files, as will usually be the case with SAS Transport 5 files, it is advised to use the same "style" for the OIDs and Names.

Also remark that when one would later want to merge the generated ODM files, using the "ODM Merger", one must assign the same fields and give the same OID for "Subject" and "StudyEvent" (visit). If each file represents a single form, then assure that for each of them, the assigned OID is different.

After clicking OK, the system asks us for some additional information:

Study OID ×

? Please provide an OID (identifier) for the study

MyStudy

We give "MyStudy" as the identifier (code) for the study. It is highly recommended to use the company internal code for the study for this. It can be that the value for it is already in the input file. Remark that it is a very good idea to use the "STUDYID" for SDTM or SEND as the OID in this dialog. This will make mapping using [SDTM-ETL](#) even easier.

After clicking OK, a new dialog is displayed:

Global Variables ×

? Please provide content for the following ODM fields


Study Name:

Study Description:

Protocol Name:

Which must be filled as these are required fields in the ODM, but is not often used afterwards. It is recommended to fill these fields with information from the protocol. For example:

Global Variables ×

 **Please provide content for the following ODM fields**

Study Name:


Study Description:

Protocol Name:

After clicking OK, the system will start generating the ODM, and will then display a file chooser, allowing you to assign and name of the ODM file.

If "Metadata + clinical data" was selected, the system will also display the following dialog:

ODM Optimization ×

 **If your source data was **not** organized per subject and visit, it is **recommended** to do an optimization step now. This step will organize all output data per subject, visit (StudyEvent) and form in the output ODM file.**

If your source data was already organized per subject and visit, the optimization step is not necessary (but doesn't harm either).

Perform optimization now

If the source data was not already organized per subject and visit, check the checkbox "Perform optimization now". The reason is that the ODM Generator will each time it encounters a new subject (from the field for "subject"), it will start a new ODM "SubjectData" element. The same applies to "StudyEvent", when a new "visit" is encountered. This is perfect when the source data was already organized per subject and visit. When not, and additional, new data for a specific subject is encountered after there has been data for another subject, there will be two (or more) "SubjectData" elements in the output ODM file for the same subject. Using "Perform optimization now", reorganizes the data in such a way that all data for a single unique subject goes into a single "SubjectData", all data for a single "StudyEvent" (visit) within a subject goes into a single "StudyEventData" element, and so on.

When in doubt, it is better to check "Perform optimization now". Especially in case of larger files, this will however increase processing time, as can be expected. Optimizing however never hurts ...

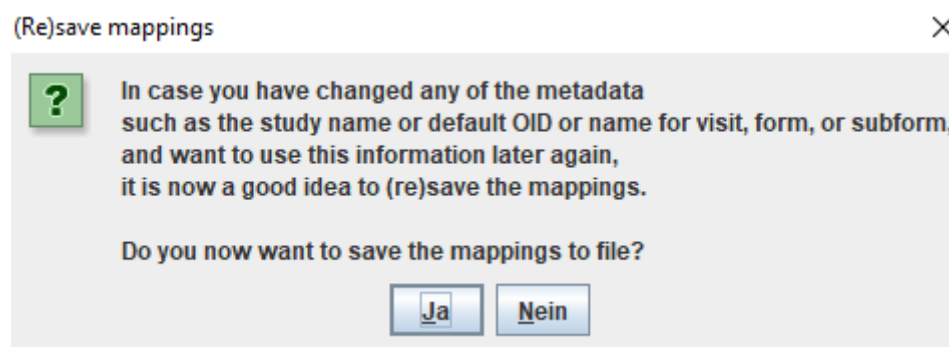
The contents of the output ODM file will then look like:

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:xlink="http://www.w3.org/1999/xlink"
3   CreationDateTime="2018-09-02T19:03:50.613"
4   Description="Generated by the XML4Pharma ODMCreator from file C:\CDISC_Standards\CDISC_Lab_1_0_1_
5   FileOID="MyStudy" FileType="Snapshot" Granularity="Metadata" ODMVersion="1.3.1"
6   SourceSystem="XML4Pharma ODMCreator" SourceSystemVersion="2017R1">
7   <Study OID="MyStudy">
8     <GlobalVariables>
9       <StudyName>My super study</StudyName>
10      <StudyDescription>Clinical study for my super new medication</StudyDescription>
11      <ProtocolName>PRQ-1234</ProtocolName>
12    </GlobalVariables>
13    <MetaDataVersion OID="MV.TestStudy" Name="Test study metadata version 1">
14      <Protocol>
15        <StudyEventRef StudyEventOID="SE.01" Mandatory="No"/>
16        <StudyEventRef StudyEventOID="SE.01R" Mandatory="No"/>
17        <StudyEventRef StudyEventOID="SE.02" Mandatory="No"/>
18      </Protocol>
19      <StudyEventDef OID="SE.01" Name="01" Repeating="No" Type="Scheduled">
20        <FormRef FormOID="FO.LABORATORY" Mandatory="No"/>
21      </StudyEventDef>
22      <StudyEventDef OID="SE.01R" Name="01R" Repeating="No" Type="Scheduled">
23        <FormRef FormOID="FO.LABORATORY" Mandatory="No"/>
24      </StudyEventDef>
25      <StudyEventDef OID="SE.02" Name="02" Repeating="No" Type="Scheduled">
26        <FormRef FormOID="FO.LABORATORY" Mandatory="No"/>
27      </StudyEventDef>
28      <FormDef OID="FO.LABORATORY" Name="Laboratory form" Repeating="No">
29        <ItemGroupRef ItemGroupOID="IG.LABORATORY" Mandatory="No"/>
30      </FormDef>
31      <ItemGroupDef OID="IG.LABORATORY" Name="Laboratory subform" Repeating="No">
32        <ItemRef ItemOID="IT.F5" Mandatory="No"/>
33        <ItemRef ItemOID="IT.F8" Mandatory="No"/>
34        <ItemRef ItemOID="IT.F9" Mandatory="No"/>

```

You can always validate this ODM file for compliance with the standard, using the "[XML4Pharma CDISC ODM Checker](#)", which is free of charge for CDISC members.

In case you have changed (or added new) any of the metadata, such as the study name, or any of the "default" OID or Name for visit (StudyEvent), form or subform (ItemGroup), you may now want to save the mappings (again) to use them in future. So, the software will ask you:



If you click "Yes", the software will ask you where (in which file) you want to save the mappings, just like in the case when you use the "save mappings" button.

Filtering

In many cases, you will not want to export the content of all the fields of your "flat text" file to CDISC ODM. In our case, there are 92 fields, and we only want to retain these that are relevant for a mapping to the SDTM LB (Laboratory Test Results) domain.

In order to select the fields that need to be exported as CDISC ODM, one should use the "Include" checkboxes on the left of the metadata mapping table. For example:

Field	Include	OID	Name	Label	Data Type	
F1	<input type="checkbox"/>	IT.F1	Model Version		text	7
F2	<input type="checkbox"/>	IT.F2	File Creation Date and Time		datetime	
F3	<input type="checkbox"/>	IT.F3	Transmission Source ID		text	5
F4	<input type="checkbox"/>	IT.F4	Transmission Source Name		text	15
F5	<input checked="" type="checkbox"/>	IT.F5	Study ID or Number		text	12
F6	<input checked="" type="checkbox"/>	IT.F6	Study Name		text	12
F7	<input type="checkbox"/>	IT.F7	Transmission Type		text	1
F8	<input checked="" type="checkbox"/>	IT.F8	Site ID or Number		integer	2
F9	<input checked="" type="checkbox"/>	IT.F9	Investigator ID or Number		integer	2
F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name		text	17

In this case, we do not want any export for the 4 first fields, as this information does not go into the SDTM "LB" (Laboratory Test Results) datasets, and is not relevant for review by the regulatory authorities. We however do want to keep the information about site and investigator.

We then select those fields only that are relevant. Whether a field is selected for export is also kept in the file with assignments and mappings that can be exported and imported again ("Load prior mappings"), so that the same mappings and assignment, including the filtering, can be used over and over again when new "flat text" files are received.

Starting from SAS Transport 5 files

New in version 3 is the ability of starting from a SAS Transport 5 ("XPT") file. In many cases, a single such file will represent an ODM form.

The procedure is almost exactly the same as for "flat text" files, except that no field delimiter need of course to be assigned. Just for completeness, we will go through the procedure for an XPT file "lb.xpt" that can be found in the directory "TestFiles". As this is a somewhat larger file (3488 rows, 2.7MB), expect longer loading and transformation times...

Here is a screenshot from when the "lb.xpt" is selected:

The ODM Generator allows you to generate CDISC ODM 1.3.1 files starting from text files with records where data fields are delimited by a vertical bar, a comma, or a semicolon, or any other delimiter of your choice, or starting from SAS Transport 5 (XPT) files

Text file with field-separation delimiters SAS Transport 5 (XPT) file

Field Delimiter:

Vertical bar (|) Comma (,) Semicolon (;)
 Tab-delimited
 Other delimiter:

Strings are embedded in single/double quotes First line contains column names Second line contains column labels

Select file... lb.xpt

Show file as table

Start generating ODM metadata

Load prior Mappings

Number of XPT records read = 3200
 Number of XPT records read = 3300
 Number of XPT records read = 3400
 Total number of XPT records read = 3488

and the window that is showing after using "Show file as table":

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRL0	LBORNRH1	LBSTRESC	LBSTRESN	LBSTRESU	LBSTNRLO	LBSTNRH1	LBNRIND
CDISCPIL...	LB	CDISC001	1	ALB	Albumin	CHEMISTRY	3.9	g/dL	3.5	4.6	39	39	g/L	35	46	NORMAL
CDISCPIL...	LB	CDISC001	2	ALP	Alkaline Ph...	CHEMISTRY	93	U/L	35	115	93	93	U/L	35	115	NORMAL
CDISCPIL...	LB	CDISC001	3	ALT	Alanine Am...	CHEMISTRY	18	U/L	6	35	18	18	U/L	6	35	NORMAL
CDISCPIL...	LB	CDISC001	4	AST	Aspartate A...	CHEMISTRY	26	U/L	11	36	26	26	U/L	11	36	NORMAL
CDISCPIL...	LB	CDISC001	5	BASO	Basophils	HEMATOL...	0.03	10 ⁹ /L	0	0.2	0.03	0.03	10 ⁹ /L	0	0.2	NORMAL
CDISCPIL...	LB	CDISC001	6	BLI	Bilirubin	CHEMISTRY	0.5	mg/dL	0.2	1.2	8.55	8.5499999...	umol/L	3	21	NORMAL
CDISCPIL...	LB	CDISC001	7	UREAN	Urea Nitrog...	CHEMISTRY	21	mg/dL	4	24	7.497	7.497	mmol/L	1.4	8.6	NORMAL
CDISCPIL...	LB	CDISC001	8	CA	Calcium	CHEMISTRY	9.1	mg/dL	8.4	10.3	2.27045	2.27045	mmol/L	2.1	2.57	NORMAL
CDISCPIL...	LB	CDISC001	9	CHOL	Cholesterol	CHEMISTRY	254	mg/dL	149	286	6.56844	6.56844	mmol/L	3.85	7.4	NORMAL
CDISCPIL...	LB	CDISC001	10	CK	Creatine Ki...	CHEMISTRY	79	U/L	22	198	79	79	U/L	22	198	NORMAL
CDISCPIL...	LB	CDISC001	11	CL	Chloride	CHEMISTRY	102	mEq/L	94	112	102	102	mmol/L	94	112	NORMAL
CDISCPIL...	LB	CDISC001	12	COLOR	Color	URINALYSIS	NORMAL				NORMAL					NORMAL
CDISCPIL...	LB	CDISC001	13	CREAT	Creatinine	CHEMISTRY	1.3	mg/dL	0.8	1.6	114.92	114.92	umol/L	71	141	NORMAL
CDISCPIL...	LB	CDISC001	14	EOS	Eosinophils	HEMATOL...	0.08	10 ⁹ /L	0	0.57	0.08	0.08	10 ⁹ /L	0	0.57	NORMAL
CDISCPIL...	LB	CDISC001	15	GGT	Gamma Gl...	CHEMISTRY	31	U/L	10	50	31	31	U/L	10	50	NORMAL
CDISCPIL...	LB	CDISC001	16	GLUC	Glucose	CHEMISTRY	74	mg/dL	50	250	4.10774	4.10774	mmol/L	2.8	13.9	NORMAL
CDISCPIL...	LB	CDISC001	17	HCT	Hematocrit	HEMATOL...	46.0	%	37	51	0.46	0.46		0.37	0.51	NORMAL
CDISCPIL...	LB	CDISC001	18	HGB	Hemoglobin	HEMATOL...	14.9	g/dL	12.5	17	9.24694	9.2469399...	mmol/L	7.76	10.55	NORMAL
CDISCPIL...	LB	CDISC001	19	K	Potassium	CHEMISTRY	4.3	mEq/L	3.4	5.4	4.3	4.3	mmol/L	3.4	5.3999999...	NORMAL
CDISCPIL...	LB	CDISC001	20	KETONES	Ketones	URINALYSIS	0				0	0				NORMAL
CDISCPIL...	LB	CDISC001	21	LYM	Lymphocytes	HEMATOL...	1.20	10 ⁹ /L	0.8	3	1.2	1.2	10 ⁹ /L	0.8	3	NORMAL
CDISCPIL...	LB	CDISC001	22	MCH	Ery. Mean ...	HEMATOL...	30	pg	26	34	1.8618	1.8618	fmol	1.6	2.1	NORMAL
CDISCPIL...	LB	CDISC001	23	MCHC	Ery. Mean ...	HEMATOL...	33	g/dL	31	38	20.4798	20.4798	mmol/L	19	24	NORMAL
CDISCPIL...	LB	CDISC001	24	MCV	Ery. Mean ...	HEMATOL...	92	fL	80	100	92	92	fL	80	100	NORMAL
CDISCPIL...	LB	CDISC001	25	MONO	Monocytes	HEMATOL...	0.38	10 ⁹ /L	0.12	0.92	0.38	0.38	10 ⁹ /L	0.12	0.92	NORMAL
CDISCPIL...	LB	CDISC001	26	SODIUM	Sodium	CHEMISTRY	139	mEq/L	135	145	139	139	mmol/L	135	145	NORMAL
CDISCPIL...	LB	CDISC001	27	PH	pH	URINALYSIS	5		5	8	5	5		5	8	NORMAL
CDISCPIL...	LB	CDISC001	28	PHOS	Phosphate	CHEMISTRY	3.4	mg/dL	2.2	5.1	1.09786	1.09786	mmol/L	0.71	1.65	NORMAL
CDISCPIL...	LB	CDISC001	29	PLAT	Platelets	HEMATOL...	150	10 ⁹ /L	130	394	150	150	10 ⁹ /L	130	394	NORMAL
CDISCPIL...	LB	CDISC001	30	PROT	Protein	CHEMISTRY	7	g/dL	6	8	67	67	g/L	60	80	NORMAL
CDISCPIL...	LB	CDISC001	31	RBC	Erythrocytes	HEMATOL...	5.00	10 ¹² /L	4	5.8	5	5	10 ¹² /L	4	5.8	NORMAL
CDISCPIL...	LB	CDISC001	32	SFGRAV	Specific Gr...	URINALYSIS	1.017		1.006	1.03	1.017	1.017		1.006	1.03	NORMAL
CDISCPIL...	LB	CDISC001	33	TSH	Thyrotropin	OTHER	2.63	mIU/L	32	5	2.63	2.63	mIU/L	0.32	5	NORMAL
CDISCPIL...	LB	CDISC001	34	URATE	Urate	CHEMISTRY	4.9	mg/dL	2.5	7.5	291.452	291.452	umol/L	149	446	NORMAL
CDISCPIL...	LB	CDISC001	35	UROBIL	Urobilinogen	URINALYSIS	0		0	0	0	0		0	0	NORMAL
CDISCPIL...	LB	CDISC001	36	VITB12	Vitamin B12	OTHER	641	ng/L	200	900	472.9298	472.9298	pmol/L	148	664	NORMAL

The example is far from optimal, as it comes from an SDTM package⁴, but it is just for demoing ...

One can now generate proposed metadata using the button "Start generating ODM metadata". The result is a new window:

⁴ One may ask why this is not used to generate "define.xml" files starting from SAS-SDTM-XPT files. It is, but that is just another product: the "Define-XML Designer".

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
STUDYID	<input checked="" type="checkbox"/>	IT.STUDYID	STUDYID	Study Identifier	text	12		<input type="checkbox"/>
DOMAIN	<input checked="" type="checkbox"/>	IT.DOMAIN	DOMAIN	Domain Abbreviation	text	2		<input type="checkbox"/>
USUBJID	<input checked="" type="checkbox"/>	IT.USUBJID	USUBJID	Unique Subject Identifier	text	8		<input type="checkbox"/>
LBSEQ	<input checked="" type="checkbox"/>	IT.LBSEQ	LBSEQ	Sequence Number	integer	3		<input type="checkbox"/>
LBTESTCD	<input checked="" type="checkbox"/>	IT.LBTESTCD	LBTESTCD	Lab Test or Examination Short Name	text	7		<input type="checkbox"/>
LBTEST	<input checked="" type="checkbox"/>	IT.LBTEST	LBTEST	Lab Test or Examination Name	text	39		<input type="checkbox"/>
LBCAT	<input checked="" type="checkbox"/>	IT.LBCAT	LBCAT	Category for Lab Test	text	10		<input type="checkbox"/>
LBORRES	<input checked="" type="checkbox"/>	IT.LBORRES	LBORRES	Result or Finding in Original Units	text	6		<input type="checkbox"/>
LBORRESU	<input checked="" type="checkbox"/>	IT.LBORRESU	LBORRESU	Original Units	text	7		<input type="checkbox"/>
LBORNRL0	<input checked="" type="checkbox"/>	IT.LBORNRL0	LBORNRL0	Reference Range Lower Limit in Orig Unit	float	5		<input type="checkbox"/>
LBORNRLHI	<input checked="" type="checkbox"/>	IT.LBORNRLHI	LBORNRLHI	Reference Range Upper Limit in Orig Unit	float	4		<input type="checkbox"/>
LBSTRESC	<input checked="" type="checkbox"/>	IT.LBSTRESC	LBSTRESC	Character Result/Finding in Std Format	text	8		<input type="checkbox"/>
LBSTRESN	<input checked="" type="checkbox"/>	IT.LBSTRESN	LBSTRESN	Numeric Result/Finding in Standard Units	float	20		<input type="checkbox"/>
LBSTRESU	<input checked="" type="checkbox"/>	IT.LBSTRESU	LBSTRESU	Standard Units	text	7		<input type="checkbox"/>
LBSTNRLO	<input checked="" type="checkbox"/>	IT.LBSTNRLO	LBSTNRLO	Reference Range Lower Limit-Std Units	float	19		<input type="checkbox"/>
LBSTNRHI	<input checked="" type="checkbox"/>	IT.LBSTNRHI	LBSTNRHI	Reference Range Upper Limit-Std Units	float	18		<input type="checkbox"/>
LBNRIND	<input checked="" type="checkbox"/>	IT.LBNRIND	LBNRIND	Reference Range Indicator	text	8		<input type="checkbox"/>
LBLOBXFL	<input checked="" type="checkbox"/>	IT.LBLOBXFL	LBLOBXFL	Last Observation Before Exposure Flag	text	1		<input type="checkbox"/>
VISITNUM	<input checked="" type="checkbox"/>	IT.VISITNUM	VISITNUM	Visit Number	float	4		<input type="checkbox"/>
VISIT	<input checked="" type="checkbox"/>	IT.VISIT	VISIT	Visit Name	text	27		<input type="checkbox"/>
EPOCH	<input checked="" type="checkbox"/>	IT.EPOCH	EPOCH	Epoch	text	9		<input type="checkbox"/>
LBDTCT	<input checked="" type="checkbox"/>	IT.LBDTCT	LBDTCT	Date/Time of Specimen Collection	partialDateTi...			<input type="checkbox"/>
LBDY	<input checked="" type="checkbox"/>	IT.LBDY	LBDY	Study Day of Specimen Collection	integer	3		<input type="checkbox"/>

As one sees, also the "Label" column is fully populated, as SAS Transport 5 (due to the 8- and 40 character limitations) always has a label for each variable.

And one can start again filtering which fields will make it in the ODM file, for which fields a codelist must be generated, etc. An example mappings file can be found in the directory "Mappings", named "LB_XPT_mappings.dat". Similar example mapping files for the other sample XPT files can be found in the same directory.

If the file "LB_XPT_mappings.dat" is used (using "Load prior Mappings" in the menu), the following is shown:

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
STUDYID	<input checked="" type="checkbox"/>	IT.STUDYID	STUDYID	Study Identifier	text	12		<input type="checkbox"/>
DOMAIN	<input checked="" type="checkbox"/>	IT.DOMAIN	DOMAIN	Domain Abbreviation	text	2		<input type="checkbox"/>
USUBJID	<input checked="" type="checkbox"/>	IT.USUBJID	USUBJID	Unique Subject Identifier	text	8		<input type="checkbox"/>
LBSEQ	<input checked="" type="checkbox"/>	IT.LBSEQ	LBSEQ	Sequence Number	integer	3		<input type="checkbox"/>
LBTESTCD	<input checked="" type="checkbox"/>	IT.LBTESTCD	LBTESTCD	Lab Test or Examination Short Name	text	7		<input checked="" type="checkbox"/>
LBTEST	<input checked="" type="checkbox"/>	IT.LBTEST	LBTEST	Lab Test or Examination Name	text	39		<input checked="" type="checkbox"/>
LBCAT	<input checked="" type="checkbox"/>	IT.LBCAT	LBCAT	Category for Lab Test	text	10		<input type="checkbox"/>
LBORRES	<input checked="" type="checkbox"/>	IT.LBORRES	LBORRES	Result or Finding in Original Units	text	6		<input type="checkbox"/>
LBORRESU	<input checked="" type="checkbox"/>	IT.LBORRESU	LBORRESU	Original Units	text	7		<input checked="" type="checkbox"/>
LBORNRL0	<input checked="" type="checkbox"/>	IT.LBORNRL0	LBORNRL0	Reference Range Lower Limit in Orig Unit	float	5		<input type="checkbox"/>
LBORNRLHI	<input checked="" type="checkbox"/>	IT.LBORNRLHI	LBORNRLHI	Reference Range Upper Limit in Orig Unit	float	4		<input type="checkbox"/>
LBSTRESC	<input checked="" type="checkbox"/>	IT.LBSTRESC	LBSTRESC	Character Result/Finding in Std Format	text	8		<input type="checkbox"/>
LBSTRESN	<input checked="" type="checkbox"/>	IT.LBSTRESN	LBSTRESN	Numeric Result/Finding in Standard Units	float	20		<input checked="" type="checkbox"/>
LBSTRESU	<input checked="" type="checkbox"/>	IT.LBSTRESU	LBSTRESU	Standard Units	text	7		<input type="checkbox"/>
LBSTNRLO	<input checked="" type="checkbox"/>	IT.LBSTNRLO	LBSTNRLO	Reference Range Lower Limit-Std Units	float	19		<input type="checkbox"/>
LBSTNRHI	<input checked="" type="checkbox"/>	IT.LBSTNRHI	LBSTNRHI	Reference Range Upper Limit-Std Units	float	18		<input type="checkbox"/>
LBNRIND	<input checked="" type="checkbox"/>	IT.LBNRIND	LBNRIND	Reference Range Indicator	text	8		<input type="checkbox"/>
LBLOBXFL	<input checked="" type="checkbox"/>	IT.LBLOBXFL	LBLOBXFL	Last Observation Before Exposure Flag	text	1		<input checked="" type="checkbox"/>
VISIT	<input checked="" type="checkbox"/>	IT.VISIT	VISIT	Visit Name	text	27		<input type="checkbox"/>
EPOCH	<input checked="" type="checkbox"/>	IT.EPOCH	EPOCH	Epoch	text	9		<input type="checkbox"/>
LBDTCT	<input checked="" type="checkbox"/>	IT.LBDTCT	LBDTCT	Date/Time of Specimen Collection	partialDateTi...			<input type="checkbox"/>
LBDY	<input checked="" type="checkbox"/>	IT.LBDY	LBDY	Study Day of Specimen Collection	integer	3		<input type="checkbox"/>

One sees that the field for selecting the subject is chosen to be the SAS field "USUBJID" and for selecting the StudyEvent (visit) is chosen to be "VISITNUM". Makes sense isn't it?

Exporting the ODM file as "Metadata + clinical data" then generates an ODM file with the following content:

```

<?xml version="1.0" encoding="UTF-8"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:xlink="http://www.w3.org/1999/xlink" CreationDateTime="20
<Study OID="MyStudy">
  <GlobalVariables>
    <StudyName>Test study protocol name</StudyName>
    <StudyDescription>Test study description</StudyDescription>
    <ProtocolName>Test study protocol name</ProtocolName>
  </GlobalVariables>
  <MetaDataVersion OID="MV.MyStudy" Name="MetaDataVersion for Study MyStudy">
    <Protocol>
      <StudyEventRef StudyEventOID="SE.1" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.4" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.5" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.5.01" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.7" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.8" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.9" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.10" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.11" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.12" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.13" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.1.01" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.7.01" Mandatory="No"/>
      <StudyEventRef StudyEventOID="SE.8.01" Mandatory="No"/>
    </Protocol>
    <StudyEventDef OID="SE.1" Name="1" Repeating="No" Type="Scheduled">
      <FormRef FormOID="FO.LB" Mandatory="No"/>
    </StudyEventDef>
    <StudyEventDef OID="SE.4" Name="4" Repeating="No" Type="Scheduled">
      <FormRef FormOID="FO.LB" Mandatory="No"/>
    </StudyEventDef>
    <StudyEventDef OID="SE.5" Name="5" Repeating="No" Type="Scheduled">
      <FormRef FormOID="FO.LB" Mandatory="No"/>
    </StudyEventDef>
    <StudyEventDef OID="SE.5.01" Name="5.01" Repeating="No" Type="Scheduled">

```

And the "ClinicalData" part:

```

<ClinicalData StudyOID="MyStudy" MetaDataVersionOID="MV.MyStudy">
  <SubjectData SubjectKey="CDISCP001">
    <StudyEventData StudyEventOID="SE.1">
      <FormData FormOID="FO.LB">
        <ItemGroupData ItemGroupOID="IG.LB">
          <ItemData ItemOID="IT.STUDYID" Value="CDISCP001"/>
          <ItemData ItemOID="IT.DOMAIN" Value="LB"/>
          <ItemData ItemOID="IT.LBSEQ" Value="1"/>
          <ItemData ItemOID="IT.LBTESTCD" Value="ALB"/>
          <ItemData ItemOID="IT.LBTEST" Value="Albumin"/>
          <ItemData ItemOID="IT.LBCAT" Value="CHEMISTRY"/>
          <ItemData ItemOID="IT.LBORRES" Value="3.9"/>
          <ItemData ItemOID="IT.LBORRESU" Value="g/dL"/>
          <ItemData ItemOID="IT.LBORNRLO" Value="3.5"/>
          <ItemData ItemOID="IT.LBORNRHI" Value="4.6"/>
          <ItemData ItemOID="IT.LBSTRES" Value="39"/>
          <ItemData ItemOID="IT.LBSTRESN" Value="39"/>
          <ItemData ItemOID="IT.LBSTRESU" Value="g/L"/>
          <ItemData ItemOID="IT.LBSTNRLO" Value="35"/>
          <ItemData ItemOID="IT.LBSTNRHI" Value="46"/>
          <ItemData ItemOID="IT.LBNRIND" Value="NORMAL"/>
          <ItemData ItemOID="IT.LBLOBXFL" Value="Y"/>
          <ItemData ItemOID="IT.VISIT" Value="SCREENING 1"/>
          <ItemData ItemOID="IT.EPOCH" Value="SCREENING"/>
          <ItemData ItemOID="IT.LBDTC" Value="2012-11-23T11:20"/>
          <ItemData ItemOID="IT.LBDY" Value="-7"/>
        </ItemGroupData>
        <ItemGroupData ItemGroupOID="IG.LB">
          <ItemData ItemOID="IT.STUDYID" Value="CDISCP001"/>
          <ItemData ItemOID="IT.DOMAIN" Value="LB"/>
          <ItemData ItemOID="IT.LBSEQ" Value="2"/>
          <ItemData ItemOID="IT.LBTESTCD" Value="ALP"/>

```

All example output files of this manual can be found in the directory "TestResults" in the distribution.

Editing OIDs

New starting from v.3.2 is that one can change OIDs more easily in "replace all" mode.

This is especially interesting when one needs to generate different ODM files and later wants to merge them together in a single ODM file (or one file with metadata and one with clinical data) - see section "Merging ODM files". In such a case, OIDs must be synchronized between files. For example, each generated ODM should then have the same OID for the item "visit number". On the other hand, different ODM files may then not have the same OID for items that differ between files. Especially when starting from character-delimited files, this will become problematic when the originally assigned OIDs like "IT.F1" etc. are maintained, as the meaning of the first field may considerably differ between files.

In such cases, it is good custom to change the initially assigned OIDs in such a way that meanings of items that are believed to be unique, get a unique OID, also between files.

For examples, when one has a file that contains all vital signs data, one may change the initially assigned OIDs like "IT.Fx" (where "x" is a number) into "IT.VITALS.x" (where "x" is a number, or a designation, like "_SYSBP", but then assign the OID for "visit number" to "IT.VISITNUM" and do the latter in all mappings, as the visit number is something that repeats over files.

Editing OIDs manually in the "Metadata Editor" can be tedious if the number of items is large, so we added a new feature to do a "replace all". This can be done using the "Edit OIDs" button.

For example, for our lab data example, we have:

Field	Include	OID	Name	Label	Data Type	Length	# dec.Digits	Gen.CodeList
F1	<input type="checkbox"/>	IT.F1	Model Version		text	7		<input type="checkbox"/>
F2	<input type="checkbox"/>	IT.F2	File Creation Date a...		datetime			<input type="checkbox"/>
F3	<input type="checkbox"/>	IT.F3	Transmission Sour...		text	5		<input type="checkbox"/>
F4	<input type="checkbox"/>	IT.F4	Transmission Sour...		text	15		<input type="checkbox"/>
F5	<input checked="" type="checkbox"/>	IT.F5	Study ID or Number		text	12		<input type="checkbox"/>
F6	<input type="checkbox"/>	IT.F6	Study Name		text	12		<input type="checkbox"/>
F7	<input type="checkbox"/>	IT.F7	Transmission Type		text	1		<input type="checkbox"/>
F8	<input checked="" type="checkbox"/>	IT.F8	Site ID or Number		integer	2		<input type="checkbox"/>
F9	<input checked="" type="checkbox"/>	IT.F9	Investigator ID or Nu...		integer	2		<input type="checkbox"/>
F10	<input checked="" type="checkbox"/>	IT.F10	Investigator Name		text	17		<input type="checkbox"/>
F12	<input type="checkbox"/>	IT.F12	Subject ID or Number		text	0		<input type="checkbox"/>
F13	<input type="checkbox"/>	IT.F13	Spare subject level L...		text	0		<input type="checkbox"/>
F14	<input type="checkbox"/>	IT.F14	Subject Initials		text	3		<input type="checkbox"/>
F15	<input checked="" type="checkbox"/>	IT.F15	Subject Sex		text	1		<input checked="" type="checkbox"/>
F16	<input type="checkbox"/>	IT.F16	Subject Sex Code Li...		text	33		<input type="checkbox"/>
F17	<input checked="" type="checkbox"/>	IT.F17	Subject Date Of Birth		date			<input type="checkbox"/>
F18	<input type="checkbox"/>	IT.F18	Subject Race		text	0		<input checked="" type="checkbox"/>
F19	<input type="checkbox"/>	IT.F19	Subject Race Code ...		text	31		<input type="checkbox"/>
F21	<input type="checkbox"/>	IT.F21	Visit Name		text	10		<input type="checkbox"/>
F22	<input type="checkbox"/>	IT.F22	Visit Type		text	1		<input checked="" type="checkbox"/>
F23	<input type="checkbox"/>	IT.F23	Visit Type Modifier		text	1		<input type="checkbox"/>
F24	<input checked="" type="checkbox"/>	IT.F24	Central Laboratory ID		text	5		<input type="checkbox"/>
F25	<input type="checkbox"/>	IT.F25	Central Laboratory ...		text	15		<input type="checkbox"/>
F26	<input checked="" type="checkbox"/>	IT.F26	Accession ID or Nu...		text	7		<input type="checkbox"/>
F27	<input type="checkbox"/>	IT.F27	Last Active Date and...		datetime			<input type="checkbox"/>
F28	<input type="checkbox"/>	IT.F28	Record Extension T...		text	4		<input type="checkbox"/>
F29	<input checked="" type="checkbox"/>	IT.F29	Specimen ID or Nu...		integer	1		<input type="checkbox"/>
F30	<input checked="" type="checkbox"/>	IT.F30	Actual Collection Da...		datetime			<input type="checkbox"/>
F31	<input type="checkbox"/>	IT.F31	Planned Collection ...		text	9		<input type="checkbox"/>

and we want to change all OIDs in the form "IT.LAB.Fx" where "x" is the field number. This can easily be done manually in the table, but it is a lot of work, as there are 92 items in the list.

We can however use the button "Edit OIDs". If it is clicked, the following window is displayed:

✕

Edit/Replace OIDs

? For all OIDs:

replace:

by:

ATTENTION: replacement is case-sensitive

copy from Name

with prefix:

without prefix

Replace blanks by:

Preview

I can now ask to replace all "IT." into e.g. "IT.LAB." by filling the fields:

Edit/Replace OIDs ×

? For all OIDs:

replace:

by:

ATTENTION: replacement is case-sensitive

copy from Name

with prefix:

without prefix

Replace blanks by:

Preview

It is then always a good idea to use the "Preview" button, which leads to a new window:

i

Old OID	New OID
IT.F1	IT.LAB.F1
IT.F2	IT.LAB.F2
IT.F3	IT.LAB.F3
IT.F4	IT.LAB.F4
IT.F5	IT.LAB.F5
IT.F6	IT.LAB.F6
IT.F7	IT.LAB.F7
IT.F8	IT.LAB.F8
IT.F9	IT.LAB.F9
IT.F10	IT.LAB.F10
IT.F11	IT.LAB.F11
IT.F12	IT.LAB.F12
IT.F13	IT.LAB.F13
IT.F14	IT.LAB.F14
IT.F15	IT.LAB.F15
IT.F16	IT.LAB.F16
IT.F17	IT.LAB.F17
IT.F18	IT.LAB.F18
IT.F19	IT.LAB.F19
IT.F20	IT.LAB.F20
IT.F21	IT.LAB.F21
IT.F22	IT.LAB.F22
IT.F23	IT.LAB.F23
IT.F24	IT.LAB.F24

After closing this window by clicking "OK", if one is then satisfied, one can click the "OK" button, or, when decides not to change anything anyway, one clicks the "Cancel" button:

Edit/Replace OIDs ×

? For all OIDs:

replace:

by:

ATTENTION: replacement is case-sensitive

copy from Name

with prefix:

without prefix

replace blanks by

When one decides to apply the changes, clicking "OK" leads to:

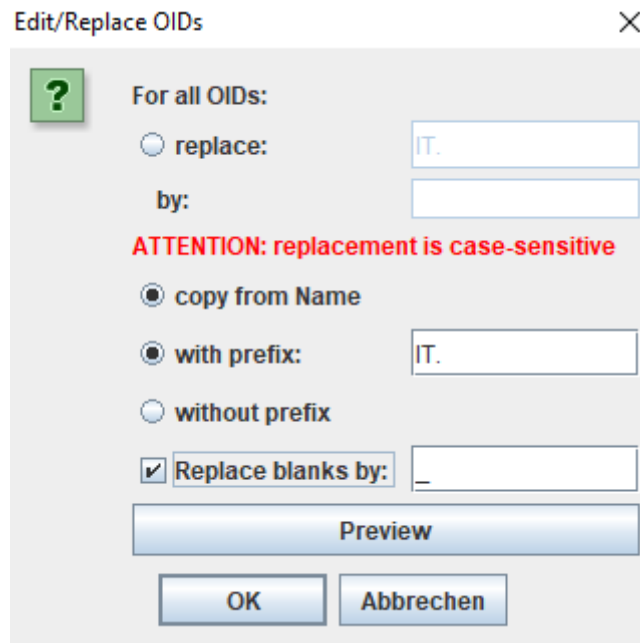
Field	Include	OID	Name	Label	Data Type	Length
F1	<input type="checkbox"/>	IT.LAB.F1	Model Version		text	7
F2	<input type="checkbox"/>	IT.LAB.F2	File Creation Date a...		datetime	
F3	<input type="checkbox"/>	IT.LAB.F3	Transmission Sour...		text	5
F4	<input type="checkbox"/>	IT.LAB.F4	Transmission Sour...		text	15
F5	<input checked="" type="checkbox"/>	IT.LAB.F5	Study ID or Number		text	12
F6	<input type="checkbox"/>	IT.LAB.F6	Study Name		text	12
F7	<input type="checkbox"/>	IT.LAB.F7	Transmission Type		text	1
F8	<input checked="" type="checkbox"/>	IT.LAB.F8	Site ID or Number		integer	2
F9	<input checked="" type="checkbox"/>	IT.LAB.F9	Investigator ID or Nu...		integer	2
F10	<input checked="" type="checkbox"/>	IT.LAB.F10	Investigator Name		text	17
F12	<input type="checkbox"/>	IT.LAB.F11	Subject ID or Number		text	0
F13	<input type="checkbox"/>	IT.LAB.F12	Spare subject level l...		text	0
F14	<input type="checkbox"/>	IT.LAB.F13	Subject Initials		text	3
F15	<input checked="" type="checkbox"/>	IT.LAB.F14	Subject Sex		text	1
F16	<input type="checkbox"/>	IT.LAB.F15	Subject Sex Code Li...		text	33
F17	<input checked="" type="checkbox"/>	IT.LAB.F16	Subject Date Of Birth		date	
F18	<input type="checkbox"/>	IT.LAB.F17	Subject Race		text	1

One can then still change the OID in individual items, especially those that are "generic", i.e. also will occur in other files. A typical example is the "Site ID or Number", field F8, for which we then change the OID from "IT.LAB.F8" into "IT.SITE":

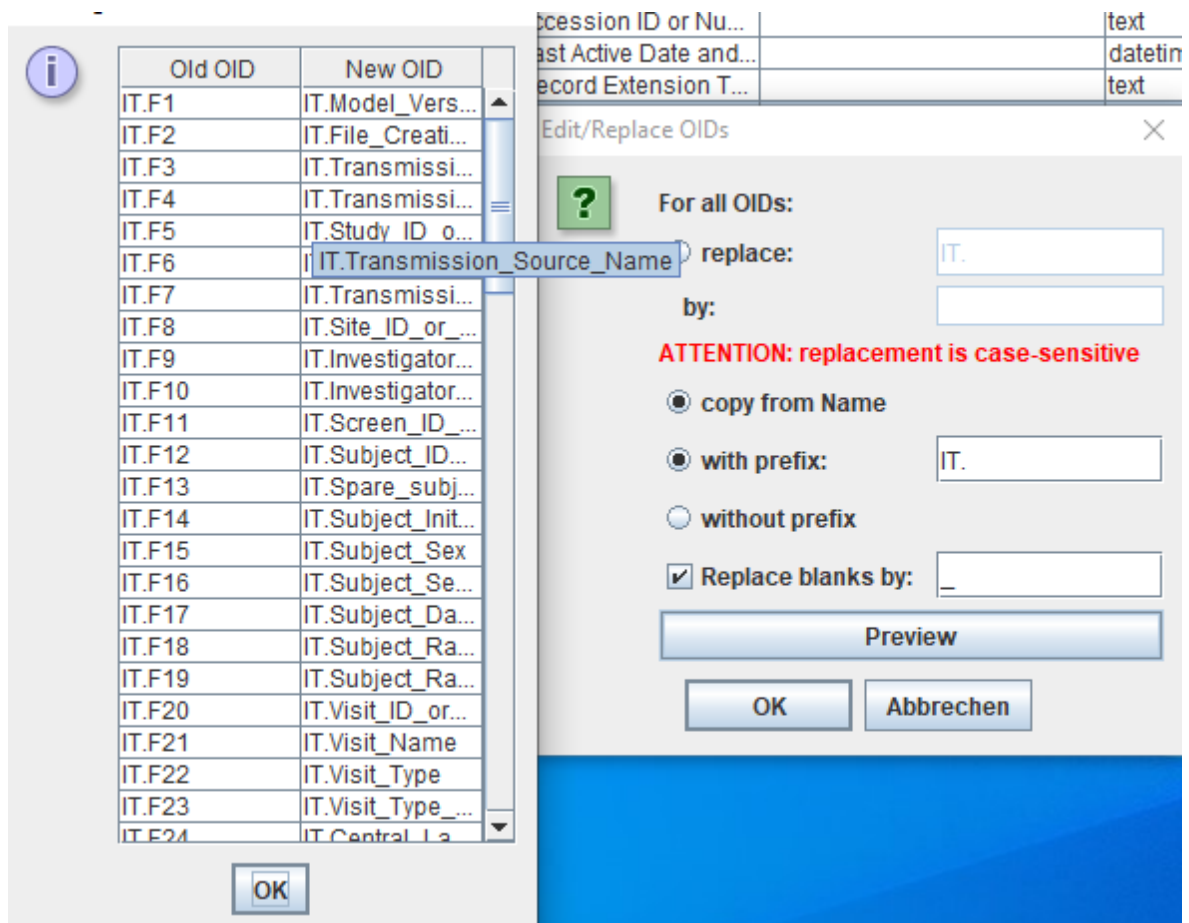
F6	<input type="checkbox"/>	IT.LAB.F6	Study Name
F7	<input type="checkbox"/>	IT.LAB.F7	Transmission Type
F8	<input checked="" type="checkbox"/>	IT.SITE	Site ID or Number
F9	<input checked="" type="checkbox"/>	IT.LAB.F9	Investigator ID or Nu...
F10	<input checked="" type="checkbox"/>	IT.LAB.F10	Investigator Name

If one wants to keep these important changes, do not forget to use the "Save Mappings" button!

The second possibility is to use (parts) of the "Name" for the OID, if there is such one. When doing so, it is good practice (but not obligation) to replace blanks by another character, e.g. an underscore ("_"). In order to use this possibility, first click "Edit OIDs", and then check the radiobutton "copy from Name":



In our case, as we have blanks in the names, we will replace them by an underscore, and add a prefix "IT.". Again, we can use the "Preview" button leading to:



When fine, one can execute the changes by clicking "OK", leading to:

Field	Include	OID	Name	Label	Dat...	Len...	# de...	Gen...
F13	<input type="checkbox"/>	IT.Spare_subject_level_ID_or_Number	Spare subject level ID or Number		text	0		
F14	<input type="checkbox"/>	IT.Subject_Initials	Subject Initials		text	3		
F15	<input checked="" type="checkbox"/>	IT.Subject_Sex	Subject Sex		text	1		
F16	<input type="checkbox"/>	IT.Subject_Sex_Code_List_ID	Subject Sex Code List ID		text	33		
F17	<input checked="" type="checkbox"/>	IT.Subject_Date_Of_Birth	Subject Date Of Birth		date			
F18	<input type="checkbox"/>	IT.Subject_Race	Subject Race		text	0		
F19	<input type="checkbox"/>	IT.Subject_Race_Code_List_ID	Subject Race Code List ID		text	31		
F20	<input type="checkbox"/>	IT.Visit_ID_or_Number	Visit ID or Number		text	3		
F21	<input type="checkbox"/>	IT.Visit_Name	Visit Name		text	10		
F22	<input type="checkbox"/>	IT.Visit_Type	Visit Type		text	1		
F23	<input type="checkbox"/>	IT.Visit_Type_Modifier	Visit Type Modifier		text	1		
F24	<input checked="" type="checkbox"/>	IT.Central_Laboratory_ID	Central Laboratory ID		text	5		
F25	<input type="checkbox"/>	IT.Central_Laboratory_Name	Central Laboratory Name		text	15		
F26	<input type="checkbox"/>	IT.Accession_ID_or_Number	Accession ID or Number		text	7		
F27	<input checked="" type="checkbox"/>	IT.Last_Active_Date_and_Time	Last Active Date and Time		date...			
F28	<input checked="" type="checkbox"/>	IT.Record_Extension_Type	Record Extension Type		text	4		
F29	<input type="checkbox"/>	IT.Specimen_ID_or_Number	Specimen ID or Number		inte...	1		
F30	<input checked="" type="checkbox"/>	IT.Actual_Collection_Date_and_Time	Actual Collection Date and Time		date...			
F31	<input checked="" type="checkbox"/>	IT.Planned_Collection_Time_Elapsed	Planned Collection Time Elapsed		text	9		
F32	<input checked="" type="checkbox"/>	IT.Planned_Collection_Time_Elapsed_D...	Planned Collection Time Elapsed Descrip...		text	9		
F33	<input checked="" type="checkbox"/>	IT.Collection_End_Date_and_Time	Collection End Date and Time		text	0		
F34	<input checked="" type="checkbox"/>	IT.Received_Date_and_Time	Received Date and Time		date...			
F35	<input checked="" type="checkbox"/>	IT.Specimen_Condition	Specimen Condition		text	17		
F36	<input checked="" type="checkbox"/>	IT.Lab_-_Specimen_Comments	Lab - Specimen Comments		text	0		
F37	<input checked="" type="checkbox"/>	IT.Investigator_-_Specimen_Comments	Investigator - Specimen Comments		text	0		
F38	<input checked="" type="checkbox"/>	IT.Specimen_Material_ID	Specimen Material ID		text	4		
F39	<input checked="" type="checkbox"/>	IT.Specimen_Material_Code_List_ID	Specimen Material Code List ID		text	35		
F40	<input checked="" type="checkbox"/>	IT.Specimen_Material_Name	Specimen Material Name		text	11		
F41	<input checked="" type="checkbox"/>	IT.Subject_Age_at_Collection	Subject Age at Collection		inte...	2		

Field for Subject ID: NONE Field for visit (StudyEvent): NONE Field for form (Form): NONE

Edit OIDs Validate Save Mappings Generate ODM

WARNING: OIDs must be unique in ODM, Names not.

This means that when you have two items, and they have the same name, and you use the name to generate better OIDs, this will result in two items with the same OID, which will surely lead to problems later. But of course, you can always still edit the OIDs one by one before starting generating the ODM.

Running the software in batch mode

Once a file with "mappings" has been generated, it is also possible to use this file CLI or in batch mode, i.e. without the use of the graphical user interface. This is especially useful when new clinical data is generated in the source format, and an ODM file with clinical data needs to be generated from that and the previously developed mappings between source data and ODM data.

At this moment, running in batch mode only support generation of "clinical data only", as this is the most common use case for running in batch. We are considering to also implement the two other use cases "metadata only" and "metadata + clinical data" in the future. These are however only minor use cases.

An example of the processing instructions in a "batch file" is found in the file "ODMGenerator3_batch_CDISC_Lab_example.bat" in the folder "Batch".

Here are the contents:

```
1 @ECHO OFF
2 java -jar ..\ODMGenerator_batch.jar
3   -inputfilelocation=D:\CDISC_Standards\CDISC_Lab_1_0_1_final\Lab1-0-1-BaseSampleData.dat
4   -odmoutputfilelocation=D:\temp\ODM_test_export.xml
5   -mappingsfilelocation=D:\ODMGenerator\ODMGenerator_v3\Mappings\Lab_1-0-1_example_mappings.dat
6
7
8
```

Lines 2 to 5 should essentially be one single line, but we have split them here for better display.

The first line is essentially an old DOS command (for those who remember the pre-Windows era) stating that the lines from the batch file need not be repeated in the output.

Lines 2 and 3 tells the system where to find the executables of the software. You may need to adapt the second part of the third line depending on where you installed the software.

Line 4 than contains the command to do the execution. Every statement that starts with a dash ("-") is a parameter-value pair, and all of these are necessary.

The parameters are the following:

Parameter	Meaning
-inputfilelocation	location (path) and name of the source data (usually CSV or other "flat" data), e.g. exported from Excel
-outputfilelocation	location (path) and name of the ODM file that will be generated
-mappingsfilelocation	location (path) and name of the file with the mappings that were stored when using the graphical user interface

TODO: add parameter for "Optimization".

When then executing this "batch" file, a window will open and display the progress of the execution. For example:

```
Windows PowerShell (x86)
PS D:\ODMGenerator\ODMGenerator_v3.2\Batch> .\ODMGenerator3_batch_CDISC_Lab_example.bat
Logging goes to: logs\ODMGenerator_LOG_2022_3_18_10-31-13.txt
2022-03-18 10:31:13,430 INFO ODM Generator
Version 3.0
by XML4Pharma
Licensed to:
Jozef Aerts
Expires 04/12/2022 18:41:15
2022-03-18 10:31:13,431 INFO Using input file = D:\CDISC_Standards\CDISC_Lab_1_0_1_final\Lab1-0-1-BaseSampleData.dat
2022-03-18 10:31:13,431 INFO Using mappings file = D:\ODMGenerator\ODMGenerator_v3\Mappings\Lab_1-0-1_example_mappings.dat
2022-03-18 10:31:13,431 INFO Now starting executing mappings using follow input:
Input file = D:\CDISC_Standards\CDISC_Lab_1_0_1_final\Lab1-0-1-BaseSampleData.dat
Output ODM file = D:\temp\ODM_test_export.xml
Mappings file = D:\ODMGenerator\ODMGenerator_v3\Mappings\Lab_1-0-1_example_mappings.dat
First line in input file will be skipped = false
2022-03-18 10:31:13,436 INFO # of lines with mapping instructions = 107
2022-03-18 10:31:13,436 INFO # of fields will be = 92
2022-03-18 10:31:13,436 INFO delimiter = |
2022-03-18 10:31:13,436 INFO STUDY OID = TestStudy
2022-03-18 10:31:13,523 INFO Data input file size = 34225
2022-03-18 10:31:13,523 INFO Starting generating clinical data
2022-03-18 10:31:13,524 INFO Skipping first record with field names
2022-03-18 10:31:13,541 INFO Creating new SubjectData with SubjectKey = 8277
2022-03-18 10:31:13,551 INFO Creating new SubjectData with SubjectKey = 8222
2022-03-18 10:31:13,556 INFO Total # lines = 53
2022-03-18 10:31:13,556 INFO ClinicalData have been written to file = D:\temp\ODM_test_export.xml
2022-03-18 10:31:13,556 INFO All Done!
PS D:\ODMGenerator\ODMGenerator_v3.2\Batch>
```

The generated ODM file with clinical data then looks like:

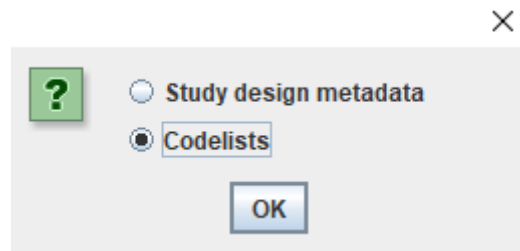
```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" CreationDateTime="2018-09-03T19:17:07.088" Description="Ge
3 <ClinicalData StudyOID="TestStudy" MetaDataVersionOID="MV.TestStudy">
4 <SubjectData SubjectKey="8222">
5 <StudyEventData StudyEventOID="SE.01">
6 <FormData FormOID="FO.Laboratory">
7 <ItemGroupData ItemGroupOID="IG.LABORATORY">
8 <ItemData ItemOID="IT.F5" Value="CDISC Test 1"/>
9 <ItemData ItemOID="IT.F8" Value="11"/>
10 <ItemData ItemOID="IT.F9" Value="11"/>
11 <ItemData ItemOID="IT.F10" Value="John Smith, M.D."/>
12 <ItemData ItemOID="IT.F15" Value="M"/>
13 <ItemData ItemOID="IT.F17" Value="1968-08-12"/>
14 <ItemData ItemOID="IT.F24" Value="C1234"/>
15 <ItemData ItemOID="IT.F26" Value="C434382"/>
16 <ItemData ItemOID="IT.F29" Value="6"/>
17 <ItemData ItemOID="IT.F30" Value="2001-05-09T10:55:00-05:00"/>
18 <ItemData ItemOID="IT.F40" Value="Urine"/>
19 <ItemData ItemOID="IT.F41" Value="32"/>
20 <ItemData ItemOID="IT.F42" Value="Y"/>
21 <ItemData ItemOID="IT.F46" Value="L1234"/>
22 <ItemData ItemOID="IT.F47" Value="Central Lab ABC - Chicago NA"/>
23 <ItemData ItemOID="IT.F48" Value="CMT5"/>
24 <ItemData ItemOID="IT.F49" Value="Urine Glucose"/>
25 <ItemData ItemOID="IT.F52" Value="2349-9"/>
26 <ItemData ItemOID="IT.F59" Value="Neg"/>
27 <ItemData ItemOID="IT.F67" Value="Neg"/>
28 <ItemData ItemOID="IT.F75" Value="Neg"/>
29 <ItemData ItemOID="IT.F91" Value="2001-05-10T10:19:32-05:00"/>
30 </ItemGroupData>
31 <ItemGroupData ItemGroupOID="IG.LABORATORY">
32 <ItemData ItemOID="IT.F5" Value="CDISC Test 1"/>
```

You may then want to use the for CDISC members freely available "[ODMChecker](#)" to validate the contents of this file with clinical data for ODM compliance and against the file with metadata that you generated using the graphical user interface.

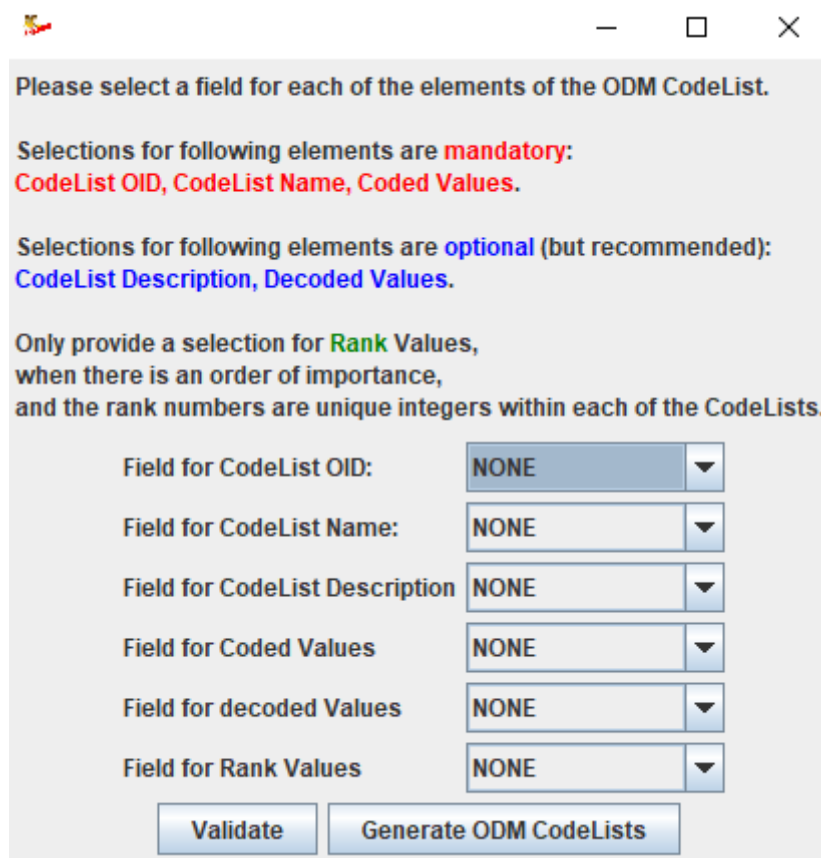
Generating ODM codelists

It may be that you have "flat" files that contain the codelists separately. In that case, you might want to generate the ODM codelists separately, and then later merge them with the file with the other metadata.

If you have a "flat file" with codelist definitions, with one codelist item definition per record (or row), after having clicked the button "Start generating ODM metadata", select the radiobutton "codelists":



After clicking "OK", a new window will appear:

A window titled "Please select a field for each of the elements of the ODM CodeList." It contains the following text: "Selections for following elements are **mandatory**: CodeList OID, CodeList Name, Coded Values." and "Selections for following elements are **optional** (but recommended): CodeList Description, Decoded Values." Below this is a note: "Only provide a selection for Rank Values, when there is an order of importance, and the rank numbers are unique integers within each of the CodeLists." There are six dropdown menus, each with "NONE" selected. The labels for the dropdowns are: "Field for CodeList OID:", "Field for CodeList Name:", "Field for CodeList Description", "Field for Coded Values", "Field for decoded Values", and "Field for Rank Values". At the bottom are two buttons: "Validate" and "Generate ODM CodeLists".

You can then starting selecting the fields to at least generate the ODM CodeList OID, a name for the codelist (ODM CodeList "Name" attribute), the field for the coded values.

If there is an additional description of the codelist in one of the fields, you may also make a selection in "Field for CodeList Description". That value will then go into the ODM CodeList/Description/TranslatedText element.

If your codes have "decodes", providing a "meaning" for each of the codes, you should also select the

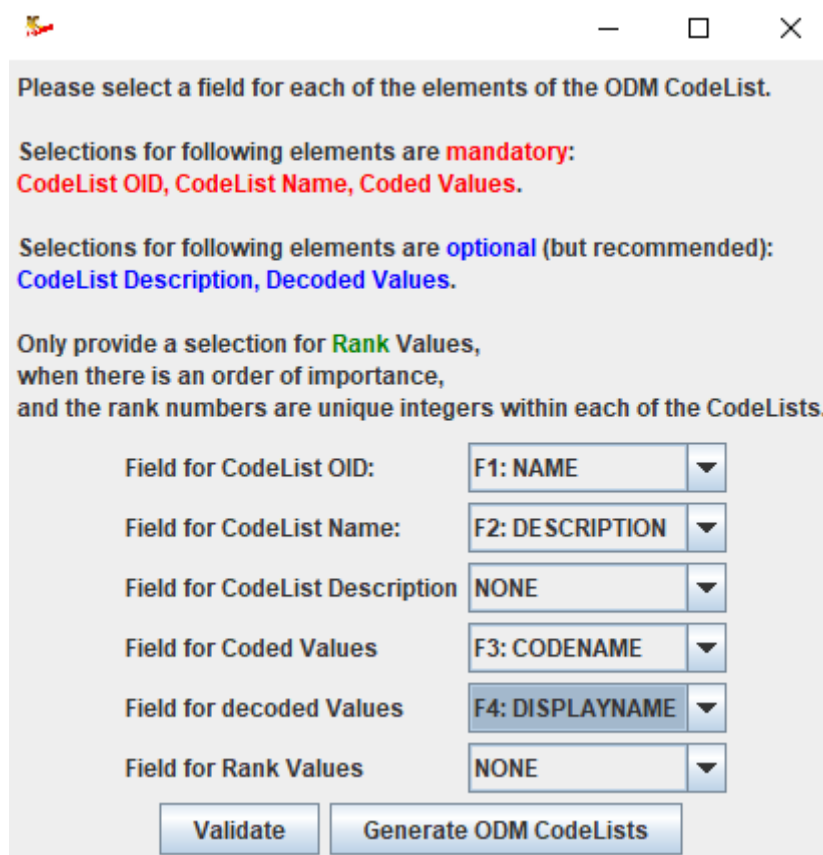
field for that. This will essentially be the case when the coded value is a number, or a single character or abbreviation (like "0" or "M" for "male", and "1" or "F" for "female").

If there is a field for the "order of importance", i.e. the "rank", you can also make a selection for "Field for Rank Values". One should however be extremely careful with this.

A typical value where also adding the "rank" is "adverse event severity", where the rank (which must be a unique integer value within the codelist for each code) is e.g. "1" for "mild", "2" for "moderate", and "3" for severe. Adding a value for "Rank" for items that have the same importance (e.g. "male", "female") doesn't make sense and will be confusing for the use of the codelist.

Also be aware that something like a "display order" is something completely different!

In our case, we make the selection:



The screenshot shows a dialog box with the following content:

Please select a field for each of the elements of the ODM CodeList.

Selections for following elements are mandatory:
CodeList OID, CodeList Name, Coded Values.

Selections for following elements are optional (but recommended):
CodeList Description, Decoded Values.

Only provide a selection for Rank Values,
when there is an order of importance,
and the rank numbers are unique integers within each of the CodeLists.

Field for CodeList OID:	F1: NAME
Field for CodeList Name:	F2: DESCRIPTION
Field for CodeList Description:	NONE
Field for Coded Values:	F3: CODENAME
Field for decoded Values:	F4: DISPLAYNAME
Field for Rank Values:	NONE

Buttons: Validate, Generate ODM CodeLists


where the field names (like "NAME") have been taken from the first line in the file (i.e. first line contains the column names). If there is no such line, the choices will simply be "F1", "F2" etc..

It is always wise to then click the button "Validate". A validation is then performed whether the minimum amount of information is provided, and when e.g. a field for "Rank Values" is provided, the values for the latter is always an integer and unique within the codelist.

When everything is fine, click the "Generate ODM CodeLists" button.

One is then first prompted for providing the Study ID (which will go into the "OID" attribute of the "Study" element in the ODM).

Study OID ×

 Please provide an OID (identifier) for the study

Please ensure that you provide the same value as when generate the other metadata of your study.

Followed by a file chooser asking where to write the generate ODM file with CodeList definitions for your study. You can then later merge that with the "usual" metadata of your study.

An example of a result is given below:

```
<CodeList OID="CL.CHILDPOT" Name="Childbearing Potential" DataType="integer">
  <CodeListItem CodedValue="1">
    <Decode>
      <TranslatedText>Childbearing potential</TranslatedText>
    </Decode>
  </CodeListItem>
  <CodeListItem CodedValue="2">
    <Decode>
      <TranslatedText>Post-menopausal</TranslatedText>
    </Decode>
  </CodeListItem>
  <CodeListItem CodedValue="3">
    <Decode>
      <TranslatedText>Non-childbearing potential</TranslatedText>
    </Decode>
  </CodeListItem>
</CodeList>
```

which was generated from 3 rows in the source "flat" file.

Merging ODM files

If you need to combine data from different "flat" files, e.g. from different "tabs" in your Excel file, you will first need to generate different ODM files, one for each input file that you need, and then you will want to combine these into a single ODM file for further processing.

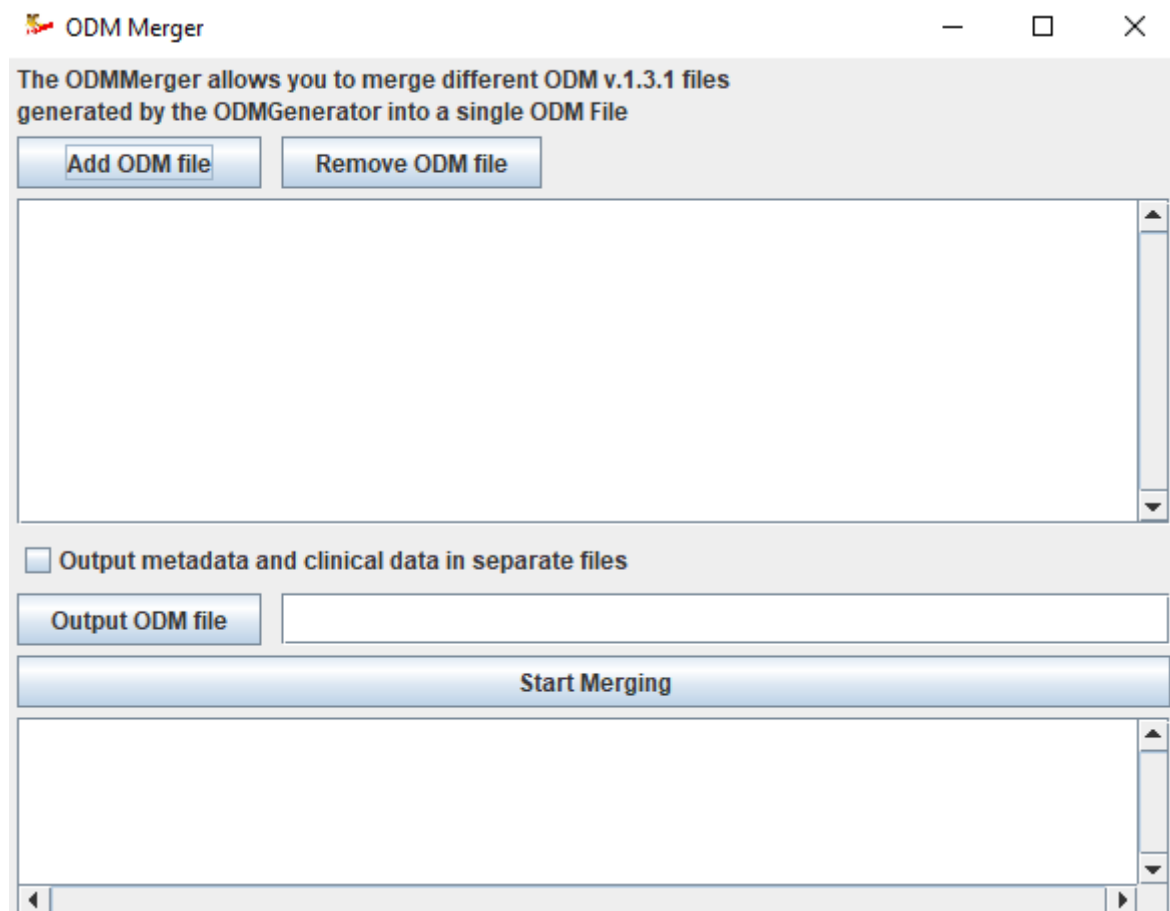
For this, we generated the "ODM Merger" software.

It comes as part of the "ODM Generator" distribution as of v.3.1 of the latter.

To start the "ODM Merger", double click "ODMMerger.bat" (on Windows, or use "ODMMerger.sh" in case of Mac or Linux):

TestResults	13.03.2022 17:36	Dateiordner	
check-java.bat	24.09.2017 10:09	Windows-Batchda...	1 KB
Distribution.jar	17.10.2005 16:06	Executable Jar File	49 KB
license.dat	13.03.2022 17:41	DAT-Datei	1 KB
log4j-1.2.13.jar	04.12.2005 18:00	Executable Jar File	350 KB
ODMGenerator.bat	29.08.2020 17:31	Windows-Batchda...	1 KB
ODMGenerator.jar	13.03.2022 18:21	Executable Jar File	289 KB
ODMGenerator_batch.jar	18.03.2022 08:57	Executable Jar File	142 KB
ODMMerger.bat	18.03.2022 10:42	Windows-Batchda...	1 KB

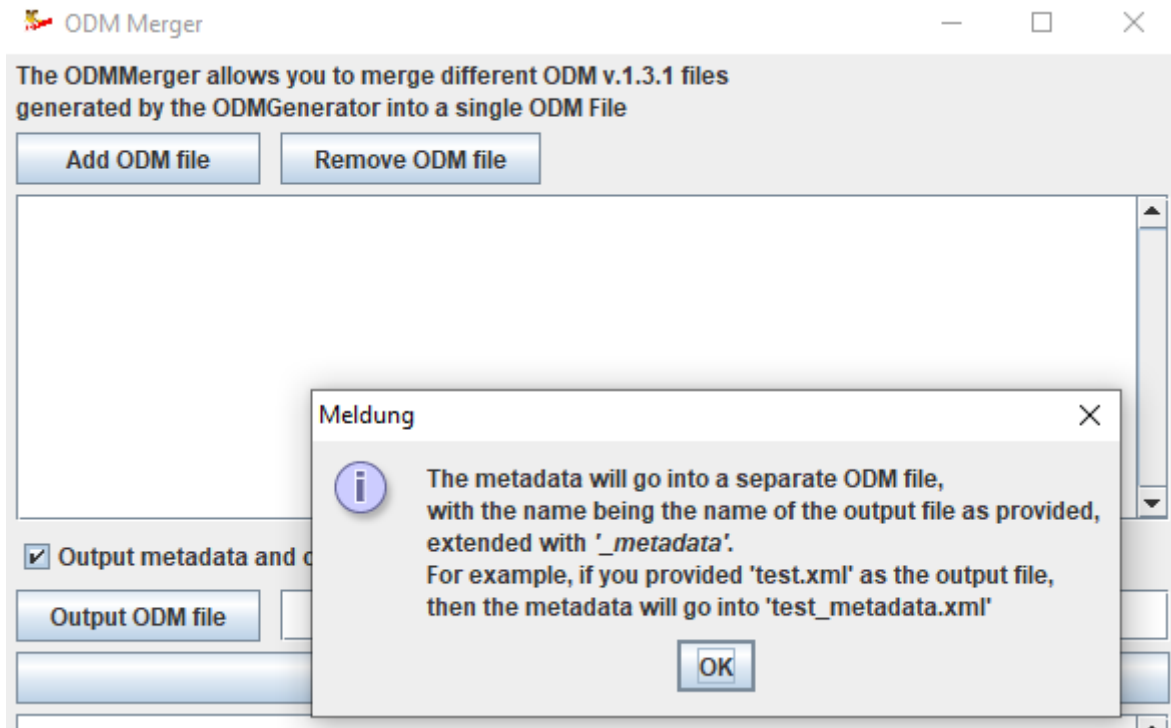
The license file will then be checked, and then the following windows will be displayed:



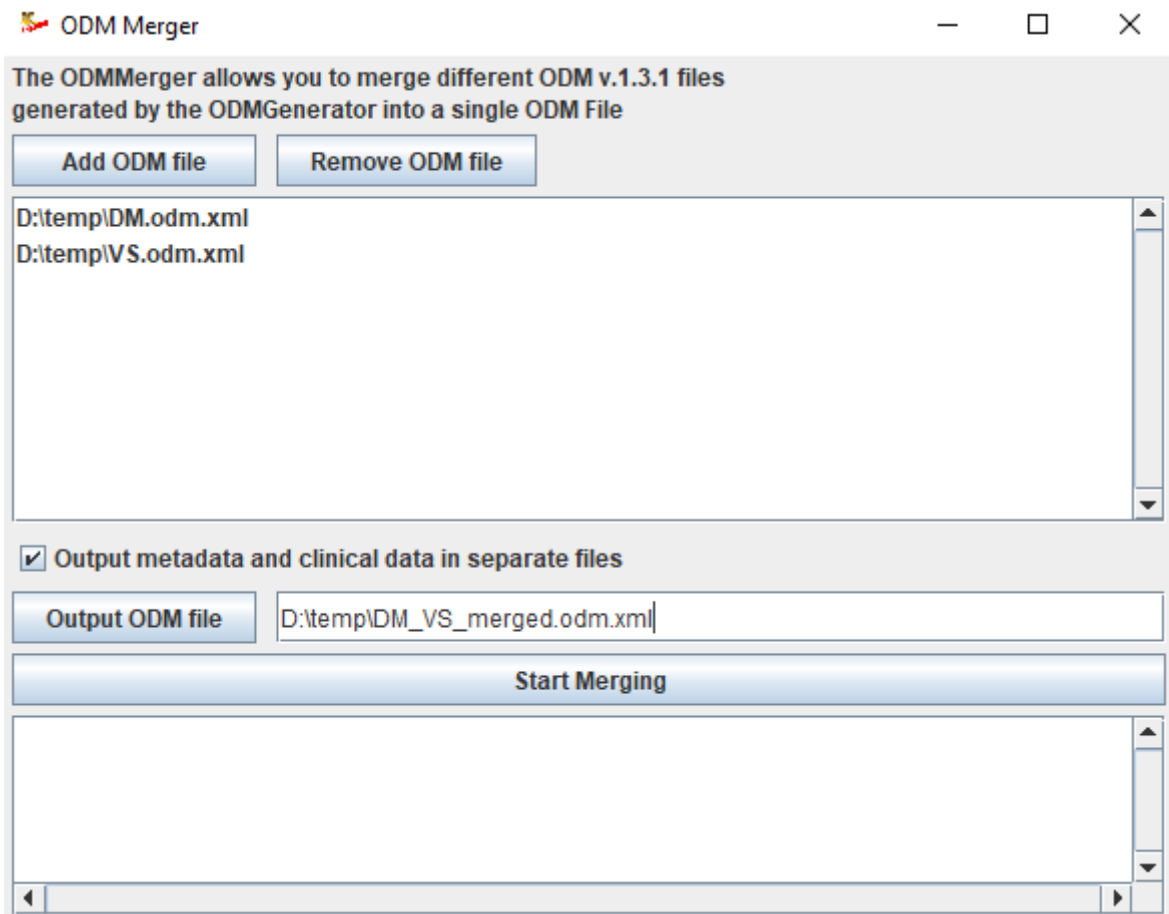
One can then add and remove ODM files using the buttons "Add ODM file" and "Remove ODM file".

Before starting merging, one can decide whether, as the output, the metadata needs to go into a separate file, so that as a result, one will have two files, one with metadata and one with clinical data. The default is that everything will be put in a single file.

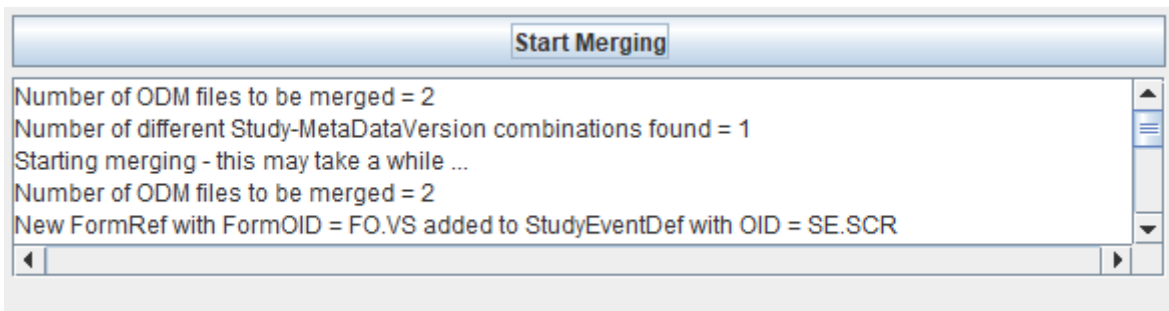
If one checks the checkbox "Output metadata and clinical data in separate files", the following message will appear:



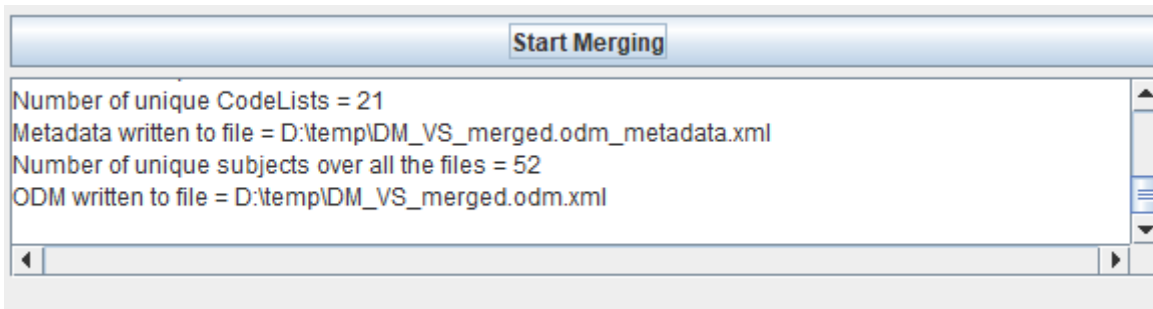
As an example, before using the "Start merging" button, we may have:



When then clicking "Start Merging", the ODM files are merged. In the lower text area, the progress of the process can be followed. That information is also written to a log file, which can be found in the "logs" folder.



and at the end:



The ODM file with combined metadata (DM_VS_merged.odm.xml) then looks like:

```
1 <?xml version="1.0" encoding="UTF-8"?>
2
3 <ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xs="http://www.w3.org/2001/XMLSchema"
4 <Study OID="MyStudy">
5   <GlobalVariables>
6     <StudyName>MyStudy protocol</StudyName>
7     <StudyDescription>MyStudy study</StudyDescription>
8     <ProtocolName>MyStudy protocol</ProtocolName>
9   </GlobalVariables>
10  <MetaDataVersion OID="MV.MyStudy" Name="MetaDataVersion for Study MyStudy">
11    <Protocol>
12      <StudyEventRef StudyEventOID="SE.SCR" Mandatory="No"/>
13      <StudyEventRef StudyEventOID="SE.VISIT1" Mandatory="No"/>
14      <StudyEventRef StudyEventOID="SE.EOS" Mandatory="No"/>
15      <StudyEventRef StudyEventOID="SE.VISIT3" Mandatory="No"/>
16      <StudyEventRef StudyEventOID="SE.VISIT2" Mandatory="No"/>
17    </Protocol>
18    <StudyEventDef OID="SE.SCR" Name="SCR" Type="Scheduled" Repeating="No">
19      <FormRef FormOID="FO.DM" Mandatory="No"/>
20      <FormRef FormOID="FO.VS" Mandatory="No"/>
21    </StudyEventDef>
22    <StudyEventDef OID="SE.VISIT1" Name="VISIT1" Type="Scheduled" Repeating="No">
23      <FormRef FormOID="FO.VS" Mandatory="No"/>
24    </StudyEventDef>
25    <StudyEventDef OID="SE.EOS" Name="EOS" Type="Scheduled" Repeating="No">
26      <FormRef FormOID="FO.VS" Mandatory="No"/>
27    </StudyEventDef>
28    <StudyEventDef OID="SE.VISIT3" Name="VISIT3" Type="Scheduled" Repeating="No">
29      <FormRef FormOID="FO.VS" Mandatory="No"/>
30    </StudyEventDef>
31    <StudyEventDef OID="SE.VISIT2" Name="VISIT2" Type="Scheduled" Repeating="No">
32      <FormRef FormOID="FO.VS" Mandatory="No"/>
33    </StudyEventDef>
34    <FormDef OID="FO.DM" Name="DM" Repeating="No">
35      <ItemGroupRef ItemGroupOID="IG.DM" Mandatory="No"/>
36    </FormDef>
37    <FormDef OID="FO.VS" Name="VS" Repeating="No">
38      <ItemGroupRef ItemGroupOID="IG.VS" Mandatory="No"/>
39    </FormDef>
40    <ItemGroupDef OID="IG.DM" Name="Demographics" Repeating="Yes">
41      <ItemRef ItemOID="IT.DM.F1" Mandatory="No"/>
42    </ItemGroupDef>
43  </MetaDataVersion>
44 </Study>
```

and the file with clinical data (DM_VS_merged.odm.xml):

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:xlink="http://www.w3.org/1999/xlink"
3   xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
4   Description="Generated by the XML4Pharma ODMCreator"
5   FileType="Snapshot" Granularity="AllClinicalData" FileOID="null"
6   CreationDateTime="2022-03-17T17:23:09.764" ODMVersion="1.3.1"
7   Originator="XML4Pharma ODMGenerator/ODMMerger" SourceSystem="XML4Pharma ODMCreator"
8   SourceSystemVersion="2021v3.1">
9   <ClinicalData StudyOID="MyStudy" MetaDataVersionOID="MV.MyStudy">
10     <SubjectData SubjectKey="US-601-0005">
11       <StudyEventData StudyEventOID="SE.SCR">
12         <FormData FormOID="FO.DM">
13           <ItemGroupData ItemGroupOID="IG.DM" ItemGroupRepeatKey="1">
14             <ItemData ItemOID="IT.DM.F1" Value="MyStudy"/>
15             <ItemData ItemOID="IT.DM.F2" Value="United States"/>
16             <ItemData ItemOID="IT.DM.F3" Value="US-601"/>
17             <ItemData ItemOID="IT.DM.F5" Value="Screening"/>
18             <ItemData ItemOID="IT.DM.F6" Value="Screening"/>
19             <ItemData ItemOID="IT.DM.F8" Value="2022-01-07"/>
20             <ItemData ItemOID="IT.DM.F9" Value="1"/>
21             <ItemData ItemOID="IT.DM.F10" Value="Demographics"/>
22             <ItemData ItemOID="IT.DM.F12" Value="1"/>
23             <ItemData ItemOID="IT.DM.F13" Value="1"/>
24             <ItemData ItemOID="IT.DM.F14" Value="2022-03-02T17:20Z"/>
25             <ItemData ItemOID="IT.DM.F15" Value="OPT0000000JS017"/>
26             <ItemData ItemOID="IT.DM.F16" Value="V500000000BP028"/>
27             <ItemData ItemOID="IT.DM.F17" Value="2022-03-02T20:51Z"/>
28             <ItemData ItemOID="IT.DM.F18" Value="1967-01-23"/>
29             <ItemData ItemOID="IT.DM.F19" Value="55"/>
30             <ItemData ItemOID="IT.DM.F20" Value="Male"/>
31             <ItemData ItemOID="IT.DM.F21" Value="Male"/>
32             <ItemData ItemOID="IT.DM.F22" Value="NOT HISPANIC OR LATINO"/>
33             <ItemData ItemOID="IT.DM.F23" Value="Not Hispanic or Latino"/>

```

Important things to take care of when merging ODM files

When merging ODM files generated by the "ODM Generator" for the same study, there are a few things one should take care of:

- ensure that the OID (object identifier) of the Study is identical in all ODM files. These can be found in the metadata parts as the "OID" attribute of the "Study" element , i.e. Study/@OID. In the clinical data, it can be found as the "StudyOID" attribute on the "ClinicalData" element. For example:

```

<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:xli
  <Study OID="MyStudy">
    <GlobalVariables>
      <StudyName>MyStudy protocol</StudyName>

```

and

```

SourceSystemVersion="2021v3.1">
  <ClinicalData StudyOID="MyStudy" MetaDataVersionOID="MV.MyStudy">
    <SubjectData SubjectKey="US-601-0005">
      <StudyEventData StudyEventOID="SE.SCR">
        <FormData FormOID="FO.DM">

```

Of course, this is something one should already take care of when generating the ODM files using the "ODM Merger".

- ensure that the subject identifiers match over all the files.

When generating the ODM files using the "ODM Generator", one must always assign a field for the subject identifier. This does not necessarily need to be the same field in each of the source files, but one of course should take care that the assignments are correct, so that during the merger, one indeed gets all unique subjects, and one "SubjectData" element is generated for each unique subject.

- manage the OIDs

Especially for fields from the original files, one should take care that fields that contain the same information in different files, are assigned the same OID. For example, if more than one of the source files has a field with the meaning of "Site identifier", one should assign the same OID to that field in each of the mappings, e.g. "IT.SITE". Keeping e.g. "IT.F7" when that is the seventh field in one of the source files is then not a good idea, as in another of the files, the "Site identifier" can e.g. be the tenth field.

Similarly, different OIDs should be assigned to fields that have different meanings. When generating the ODM from a "flat" files, OIDs like "IT.F1", "IT.F2", etc. are generated. When the files later need to be merged, one should assign an OID that is unique for representing the item. For example, if "IT.F7" means "birthdate" in one file, and "IT.F7" means "systolic blood pressure" in another, this will lead to severe problems. In such a case, one can assign OIDs like "IT.DM.F7" and "IT.VS.F7" (can be quickly done using the "Edit OIDs button"), or even better e.g. "IT.BIRTHDATE" and "IT.SYSTOLIC_BP".

- Identifiers for visits

Very often, one of the fields in the source files represent the visit identifier or visit number. It is important that then, for each file, when using the "ODM Generator", a and the correct field is assigned to "StudyEvent". The software will then automatically take care that correct and consistent OIDs are generated from the unique values in the provided fields. For example, if the fields for "visit" have values "SCREENING", "VISIT1", "VISIT2", ... and "EOS" (end of study), the OIDs generated will be "SE.SCREENING", "SE.VISIT1", "SE.VISIT2", ... and "SE.EOS". In the merged file, the data will then automatically be organized "per subject, per visit, per form, ...".

And last but not least ...

When having merged the by the "ODM Generator" generated ODM files into a single one (or one with metadata and one with clinical data), it is always a good idea to validate the correctness and consistency of the result file(s). This can best be done using the "ODM Checker", which is freely available for CDISC members and academics, and comes at a small cost for others.