Tutorial:
Smart Dataset-XML Viewer

VERSION 2017-06-20

FH-Prof. Dr. Jozef Aerts, Marie-Christin Eisner, Julie Sajan Pattery
UNIVERSITY OF APPLIED SCIENCES FH JOANNEUM | GRAZ, AUSTRIA

Table of Contents

List of Abbreviations

ADaM	Analysis Data Model
ADSL	Subject-Level Analysis Dataset
--CAT	Category
CDA	Clinical Document Architecture
CDISC	Clinical Data Interchange Standards Consortium
CM	Concomitant Medications
CO	Comments
--DTC	Date/Time of Collection
--DY	Study Day Variable
FDA	Food and Drug Administration
HL7	Health Level 7
LB	Laboratory
LOINC	Logical Observation Identifiers Names and Codes
RELREC	Related Record
RFENDTC	Subject Reference End Date/Time
SDTM	Study Data Tabulation Model
SEND	Standard for Exchange of Non-clinical Data
SEQ	Sequence Number
SNOMED	Systemized Nomenclature of Medicine
--STNRHI	Normal Range Upper-Limit in Standard Units
--STRESN	Numeric Result/Finding in Standard Units
--TESTCD	Test Code
XML	Extensible Markup Language

1. Introduction

The "Smart Dataset-XML Viewer" is a Java-based viewer for inspecting SDTM, SEND or ADaM submissions in the new Dataset -XML format.

It allows the following:

- Filtering of data
- Basic validation of the datasets (against the metadata in the corresponding define.xml file)
- Sorting by multiple variables
- Creation of subsets of data
- Find parent records of supplemental qualifier records and of comment records quickly
- Find related records as listed in the RELREC dataset

The "Smart Dataset-XML Viewer" is continuously developed and new features are added regularly, so when working with the viewer, you might encounter features that are not described here yet.

For further information and regular updates, please visit:

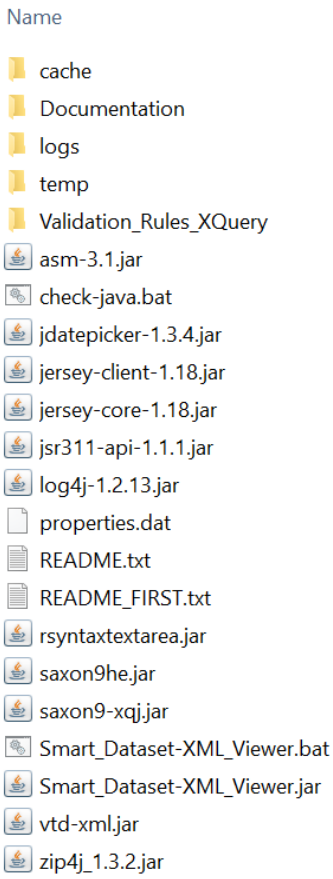
http://cdiscguru.blogspot.com/2016/08/why-lobxfl-should-not-be-in-sdtm.html [1]

To download the “Smart Dataset-XML Viewer”, please visit:

https://sourceforge.net/projects/smart-sds-xml-viewer/files/ [2]

2. Starting Up the Viewer

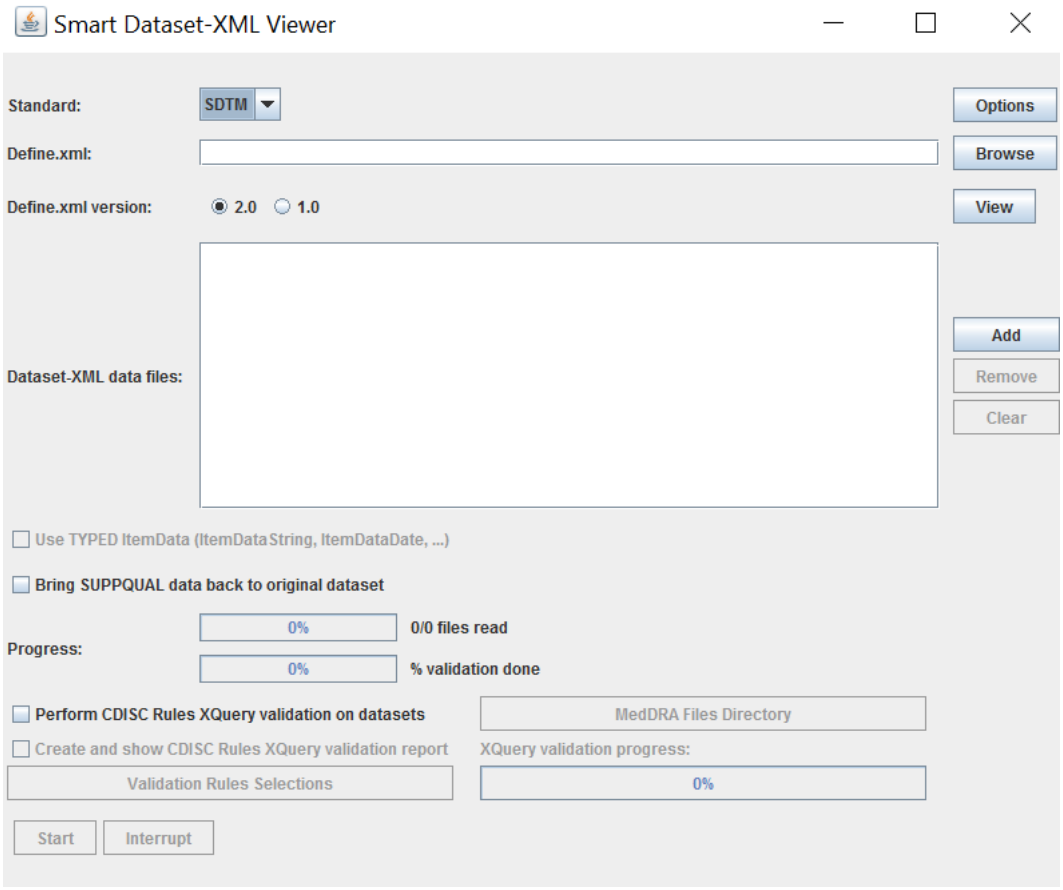
For installation instructions, please refer to the separate "Installation manual". To start the viewer, navigate to the folder where you installed the software. You will find the following files:



Before starting the “Smart Dataset-XML Viewer” for the first time, double click the file “check-java.bat” in order to find out if you have installed the needed java version.

If you see java version “1.7” or above you can continue with the next instructions, otherwise please install the latest version of the Java Runtime Environment (JRE), which you can find on http://www.oracle.com/. [3]

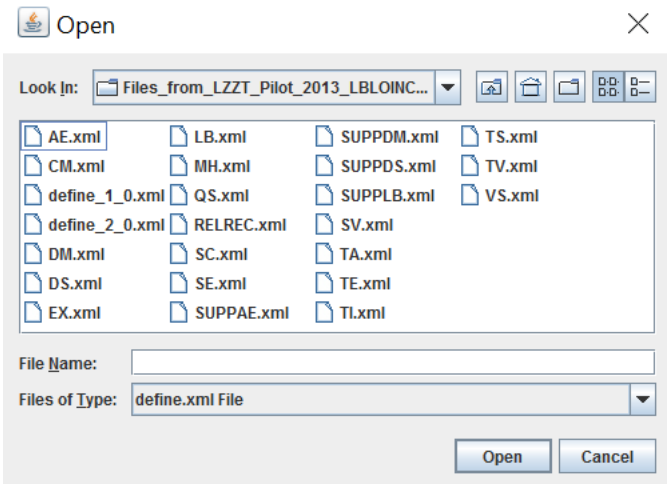
Double click the file "Smart_Dataset-XML_Viewer.bat". Then the following window is displayed:



At the top left, you will find a drop-down selector for selecting which standard (SDTM, SEND or ADaM) you would like to work with. This is necessary as the validation rules for SDTM, SEND and ADaM are different.

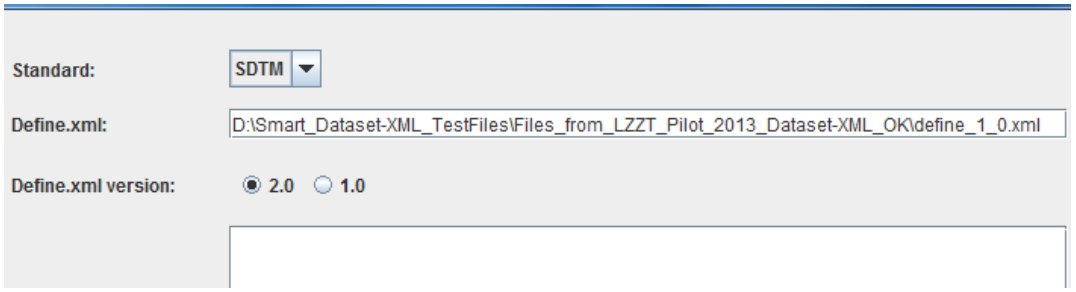
3. Adding the define.xml File

The second row allows you to select the "define.xml" file containing the metadata for your submission. Using the "Browse" button, you can navigate through the available folders and files and pick the "define.xml" file that you need. For example:



The file selector only displays the XML files in the chosen folder. If you would also like to see all other available files in a folder, use the "Files of Type" selector and select "All Files". Then select the file that is the define.xml file, and click "Open".

The text field for the "define.xml" is then filled:



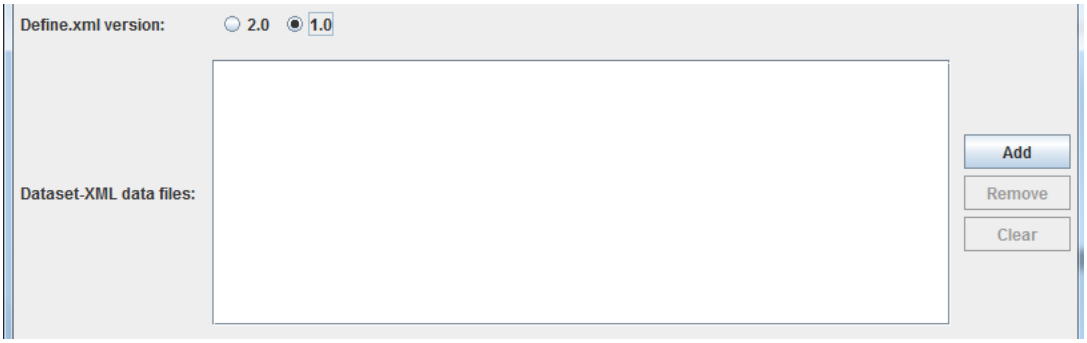
The next important step is to select the correct version of the define.xml standard that you are working with. If you select the wrong version, the viewer will later (after clicking the "Start" button) give an error message. You can then still change the version, and retry.

Please remark that you always need to have a "define.xml" file. As long as no "define.xml" file has been selected, the "Start" button will be disabled, and you cannot start loading data files.

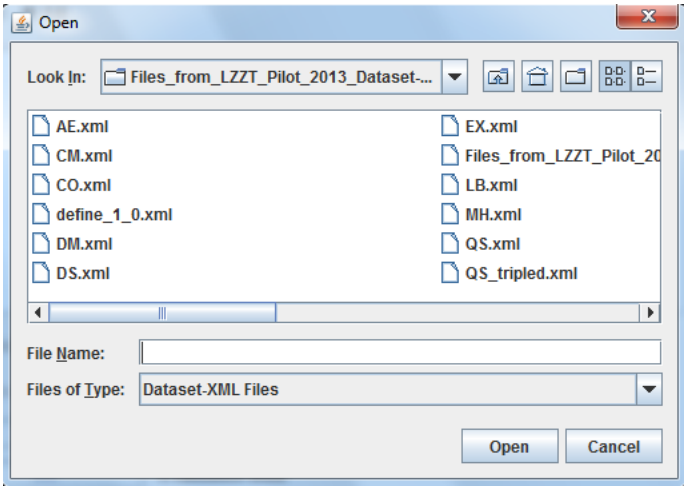
In our case, the define.xml file is still a version 1.0 file, so we need to select the radio button "1.0".

4. Adding Dataset-XML Files

Once a define.xml file and the define.xml version have been selected, you can start adding data files in Dataset-XML files for which the define.xml file applies. This can be done using the "Add" button in the main window:

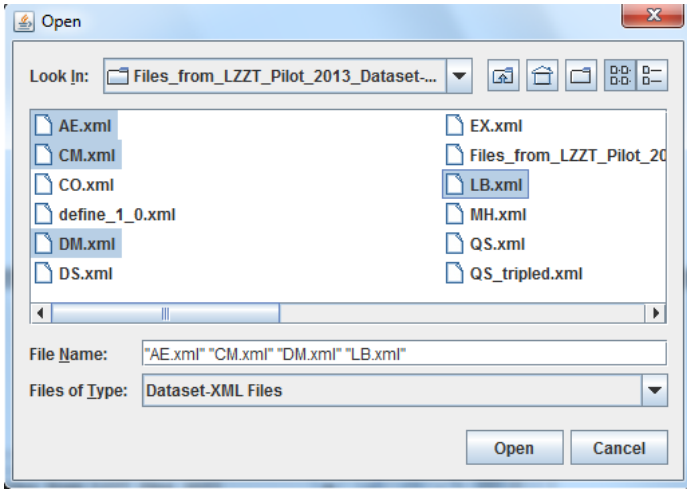


By default, when clicking the "Add" button, you will get a file selector that displays the XML files that are in the same directory. For example:



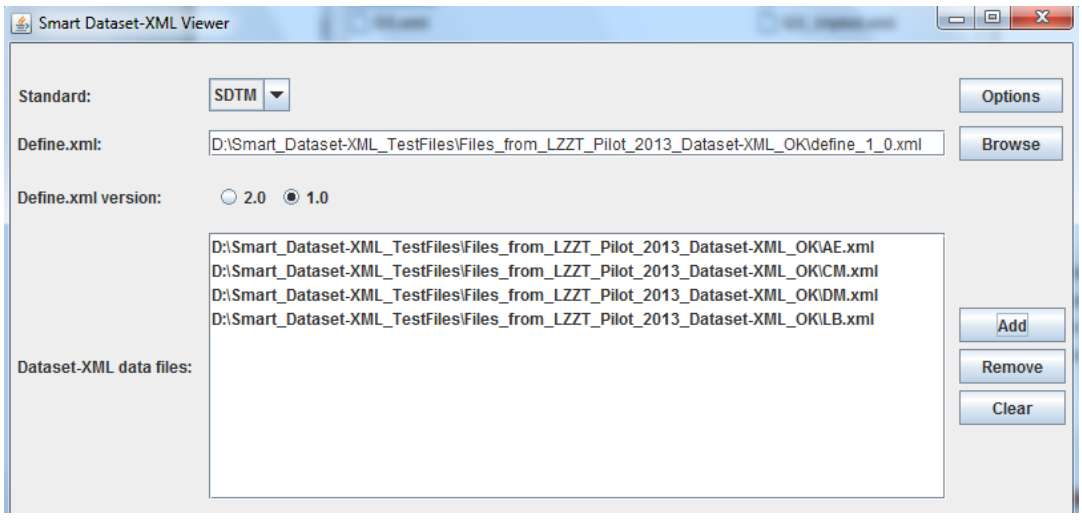
You can now select one or more Dataset-XML files (.xml extension). The files can also be zipped Dataset-XML file (.zip file extension). Remark that you should not add the "define.xml" file here again.

To select multiple files, hold the "CTRL" button and pick several files that you want to load. For example, we here select the "AE.xml", the "CM.xml" and the "DM.xml" and "LB.xml" file:



It usually does not make sense to try to select all files of the submission, as this will lead to millions of data points. Later we will learn how to make selections of subjects, and then load additional files for these selected subjects only. We do however advise to always at least select the dataset for the "DM" (ADSL in case of the ADaM standard) domain, as several of the validation checks need the information from this dataset. We will very soon also learn how to filter on --TESTCD or --CAT values, so that we can only load those LB (laboratory) records that e.g. are about albumin tests or are about urine analysis.

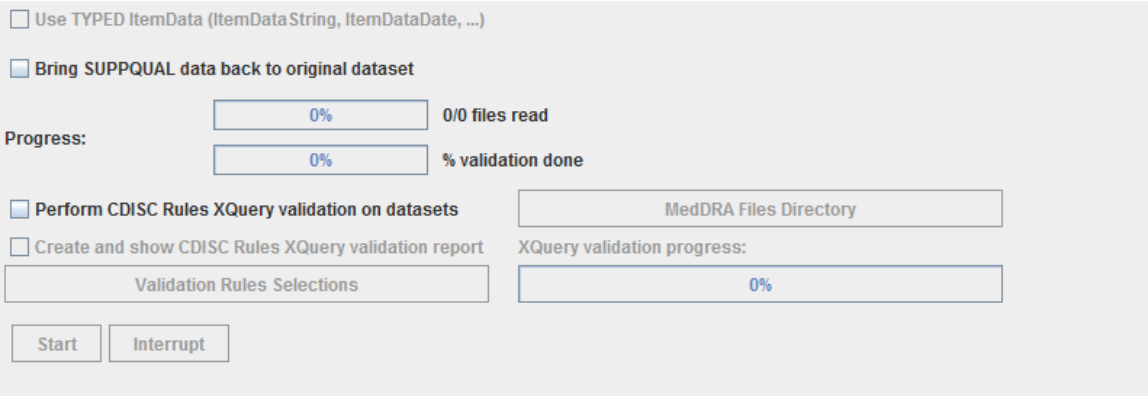
Now click the "Open" button to finalize the selection of the datasets that you want to load. This leads to:



You can now always use the "Add" button to add additional datasets to the list, use the "Remove" button to remove selected ones, or to clear the list using the "Clear" button.

We will not use the "Options" button for now but we will come back to its usage soon.

Once everything that is necessary is provided, we can start loading!



5. Displaying the Datasets

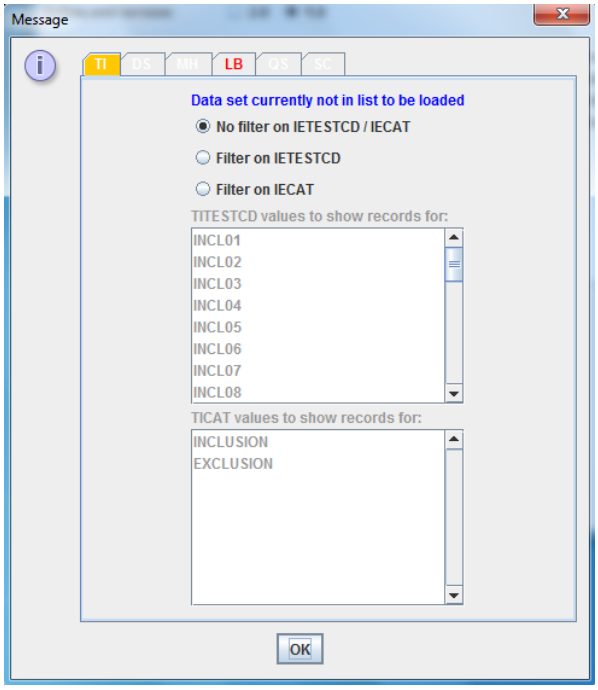
At the bottom of the screen, just above the "Start" button, you can find two "progress bars". The first one allows you to follow the progress of the dataset loading, the second one the progress of validation of the data files[1].

Now click the "Start" button to start generating the tables for the selected datasets.

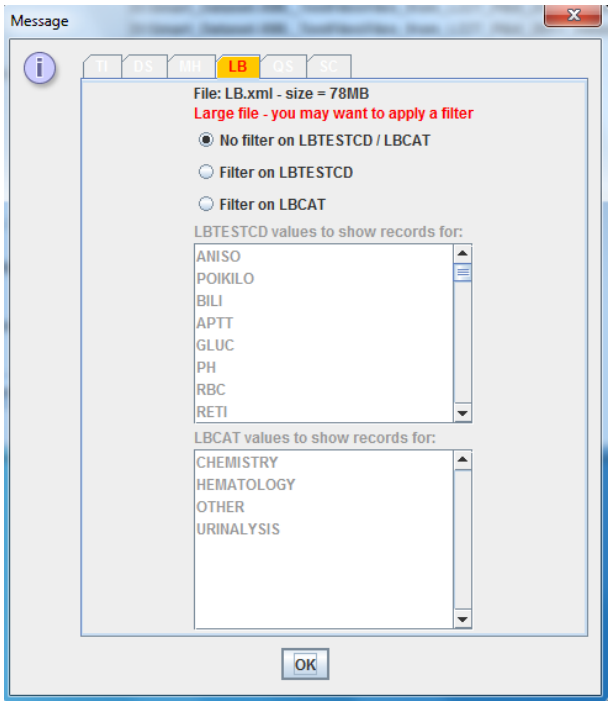
Remark that we have an LB.xml file for which there is controlled terminology for LBTESTCD, and that this file has results for all kinds and types of laboratory tests, often leading to hundreds of thousands or more records. Of course, it does not make sense to load all of these, and apply filtering afterwards. It is much better to apply a filter on LBTESTCD at the time of the loading itself.

The system will now look into the loaded define.xml file, and in our case, find out that there is a code list attached for IETESTCD and for LBTESTCD. The following dialogue is then presented:

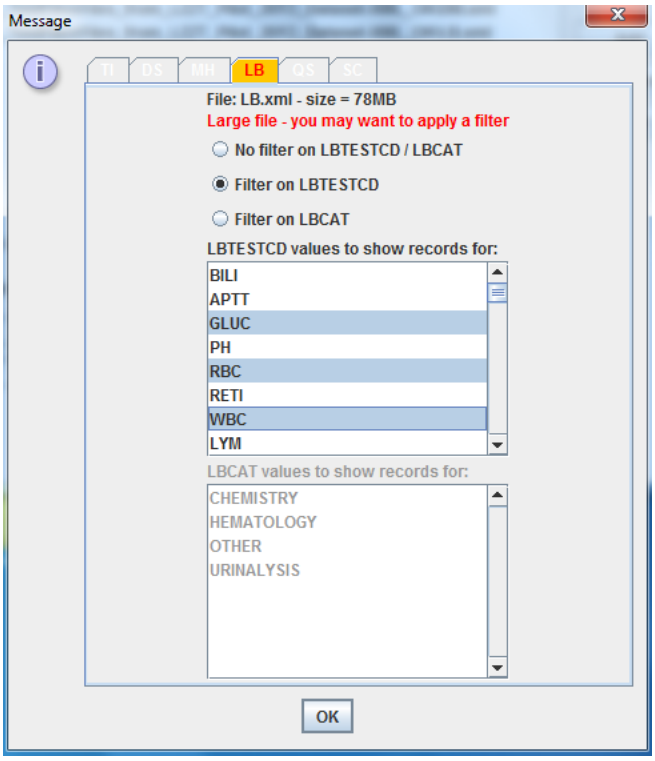
[1] Remark that some basic validation is always performed, even when all validation options using the "Options" button have been switched off.



This dialogue shows us all filtering options based on either --TESTCD or --CAT. Those tabs that have a white tab label (like "TI" in this case) are about data sets that will currently not be loaded, so you do not need to apply any filtering for now. Those that have a red label (like "LB" in our case) are about large files for which it is recommended to apply a filter. Later we will learn how to set the criteria about what a "large" file is and how to change it (default is 20MB). As LB is a very large data set, we want to apply a filter on LBTESTCD, so we click the "LB" tab:



As the file is pretty large, the system suggests that it may be useful to apply a filter during loading. If we do not want to apply any filtering at all on LBTESTCD or on LBCAT, we leave the radiobutton "No filter on LBTESTCD / LBCAT" checked. In case we only want to see the records for specific tests (recommended) we select the radiobutton "Filter on LBTESTCD" which then allows us to select one or more tests (use the CTRL-key to select multiple entries). For example:

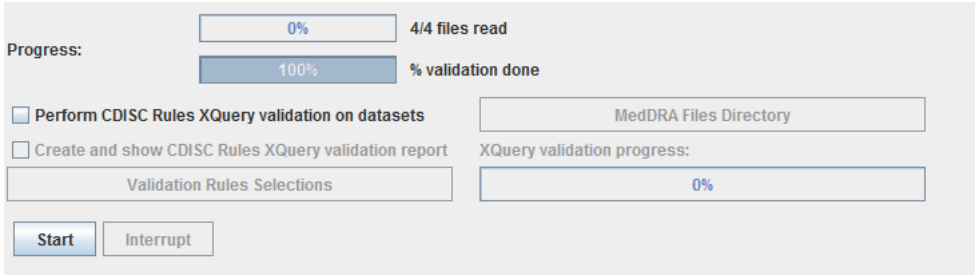


Here, we have selected to only load the records for the lab tests "GLUC" (glucose), "RBC" (red blood count) and "WBC" (white blood count).

If our define.xml file contains controlled terminology for more tests (as defined by --TESTCD) or categories (as defined by --CAT), there will of course be more tabs in this dialogue, so that we can define --TESTCD/--CAT filters for each dataset that has controlled terminology for --TESTCD and/or --CAT, or decide to not do any filtering at all.

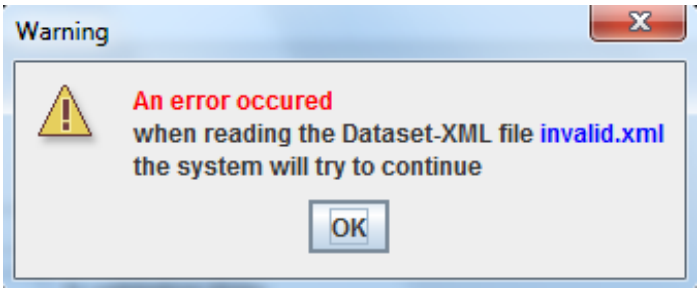
Later we will see that we can switch this filtering completely off using the "Options" - the default is to present the user the possibility to apply filtering each time the "Start" button has been clicked.

After clicking "OK", the loading process starts.

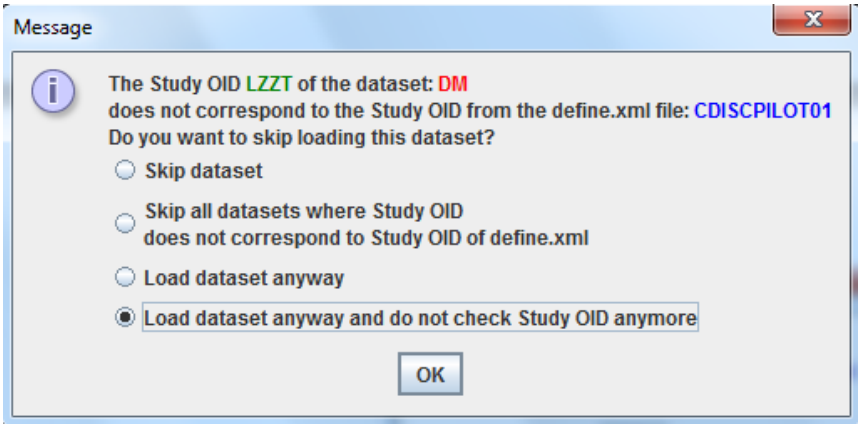


The snapshot above shows that all four files have completely been loaded

In case you try to load a file that is not a valid Dataset-XML file, you will get an error message, and the file will be skipped. For example:



When starting to load the first file, the system will check whether the Study-OID of the Dataset -XML file corresponds to that given in the define.xml file. If it does not, the following dialogue is displayed:



This is being done to avoid that one loads data that does not belong to the metadata given in the define.xml file. One can then either skip this Dataset-XML file, skip all Dataset-XML files in the series for which there is no correspondence, load the file anyway, or load the file anyway and stop checking on the Study-OID.

If all files are loaded correctly, a new window is displayed, in this case:

File Tools Search Options

DM AE CM LB

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC	RFICDTC	RFPENDTC	DTHDTC	DTHFL	SITEID	AGE	AGEU
CDISCPIL...	DM	01-701-1015	1015	2014-01-02	2014-07-02	2014-01-02	2014-07-02		2014-07-02...			701	61	YEARS
CDISCPIL...	DM	01-701-1023	1023	2012-08-05	2012-09-0a	2012-08-02	2012-09-01		2013-02-18			701	64	YEARS
CDISCPIL...	DM	01-701-1028	1028	2013-07-19	2014-01-14	2013-07-19	2014-01-14		2014-01-14...			701	71	YEARS
CDISCPIL...	DM	01-701-1033	1033	2014-03-18	2014-04-14	2014-03-18	2014-03-31		2014-09-15			701	74	YEARS
CDISCPIL...	DM	01-701-1034	1034	2014-07-01	2014-12-30	2014-07-01	2014-12-30		2014-12-30...			701	77	YEARS
CDISCPIL...	DM	01-701-1047	1047	2013-02-12	2013-03-32	2013-02-12	2013-03-09		2013-07-28			701	85	YEARS
CDISCPIL...	DM	01-701-1057	1057						2013-12-27			701	59	YEARS
CDISCPIL...	DM	01-701-1097	1097	2014-01-01	2014-07-09	2014-01-01	2014-07-09		2014-07-09...			701	68	YEARS
CDISCPIL...	DM	01-701-1111	1111	2012-09-07	2012-09-17	2012-09-07	2012-09-16		2013-02-22			701	81	YEARS
CDISCPIL...	DM	01-701-1115	1115	2012-11-30	2013-01-23	2012-11-30	2013-01-23		2013-05-20			701	84	YEARS
CDISCPIL...	DM	01-701-1118	1118	2014-03-12	2014-09-09	2014-03-12	2014-09-09		2014-09-09...			701	52	YEARS
CDISCPIL...	DM	01-701-1130	1130	2014-02-15	2014-08-16	2014-02-15	2014-08-16		2014-08-16...			701	84	YEARS
CDISCPIL...	DM	01-701-1133	1133	2012-10-28	2013-04-29	2012-10-28	2013-04-28		2013-04-29...			701	81	YEARS
CDISCPIL...	DM	01-701-1145	1145						2013-09-19			701	57	YEARS
CDISCPIL...	DM	01-701-1146	1146	2013-05-20	2013-06-30	2013-05-20	2013-06-26		2013-07-15			701	75	YEARS
CDISCPIL...	DM	01-701-1148	1148	2013-08-23	2014-02-20	2013-08-23	2014-02-20		2014-02-20...			701	57	YEARS
CDISCPIL...	DM	01-701-1153	1153	2013-09-23	2014-04-01	2013-09-23	2014-03-16		2014-04-01...			701	79	YEARS
CDISCPIL...	DM	01-701-1162	1162						2013-04-18			701	82	YEARS
CDISCPIL...	DM	01-701-1176	1176						2012-09-30			701	62	YEARS
CDISCPIL...	DM	01-701-1180	1180	2013-02-12	2013-03-23	2013-02-12	2013-03-18		2013-04-07			701	56	YEARS
CDISCPIL...	DM	01-701-1181	1181	2013-12-05	2013-12-12	2013-12-05	2013-12-09		2014-05-23			701	79	YEARS
CDISCPIL...	DM	01-701-1188	1188	2013-02-15	2013-03-25	2013-02-15	2013-03-24		2013-08-04			701	71	YEARS
CDISCPIL...	DM	01-701-1192	1192	2012-07-22	2013-01-20	2012-07-22	2013-01-20		2013-01-20...			701	80	YEARS
CDISCPIL...	DM	01-701-1203	1203	2013-02-02	2013-08-03	2013-02-02	2013-08-03		2013-08-03...			701	81	YEARS
CDISCPIL...	DM	01-701-1211	1211	2012-11-15	2013-01-14	2012-11-15	2013-01-12		2013-01-14	2013-01-14	Y	701	76	YEARS
CDISCPIL...	DM	01-701-1234	1234	2013-03-30	2013-09-22	2013-03-30	2013-09-22		2013-09-22...			701	69	YEARS
CDISCPIL...	DM	01-701-1239	1239	2014-01-11	2014-07-11	2014-01-11	2014-07-10		2014-07-11...			701	56	YEARS
CDISCPIL...	DM	01-701-1240	1240						2013-09-24			701	57	YEARS
CDISCPIL...	DM	01-701-1275	1275	2014-02-07	2014-06-14	2014-02-07	2014-05-31		2014-06-14...			701	61	YEARS
CDISCPIL...	DM	01-701-1287	1287	2014-01-25	2014-07-26	2014-01-25	2014-07-26		2014-07-26...			701	56	YEARS
CDISCPIL...	DM	01-701-1294	1294	2013-03-24	2013-06-14	2013-03-24	2013-06-14		2013-10-08			701	67	YEARS
CDISCPIL...	DM	01-701-1302	1302	2013-08-29	2013-11-05	2013-08-29	2013-11-05		2014-02-13			701	61	YEARS
CDISCPIL...	DM	01-701-1307	1307						2014-01-02			701	80	YEARS
CDISCPIL...	DM	01-701-1317	1317	2014-05-22	2014-11-20	2014-05-22	2014-11-20		2014-11-20...			701	68	YEARS
CDISCPIL...	DM	01-701-1324	1324	2012-10-02	2013-04-02	2012-10-02	2013-04-02		2013-04-02...			701	79	YEARS
CDISCPIL...	DM	01-701-1341	1341	2013-01-05	2013-02-07	2013-01-05	2013-01-26		2013-02-21			701	51	YEARS
CDISCPIL...	DM	01-701-1345	1345	2013-10-08	2014-03-18	2013-10-08	2014-03-18		2014-03-18...			701	63	YEARS
CDISCPIL...	DM	01-701-1356	1356						2014-05-28			701	54	YEARS
CDISCPIL...	DM	01-701-1360	1360	2013-07-31	2013-08-14	2013-07-31	2013-08-05		2014-02-11			701	67	YEARS
CDISCPIL...	DM	01-701-1363	1363	2013-05-30	2013-11-27	2013-05-30	2013-11-27		2013-11-27...			701	81	YEARS
CDISCPIL...	DM	01-701-1369	1369						2013-09-30			701	74	YEARS
CDISCPIL...	DM	01-701-1383	1383	2013-02-04	2013-08-06	2013-02-04	2013-08-06		2013-08-06...			701	72	YEARS
CDISCPIL...	DM	01-701-1386	1386						2014-01-07			701	71	YEARS
CDISCPIL...	DM	01-701-1387	1387	2014-03-12	2014-03-25	2014-03-12	2014-03-25		2014-08-27			701	87	YEARS
CDISCPIL...	DM	01-701-1392	1392	2012-10-28	2013-04-28	2012-10-28	2013-04-28		2013-04-28...			701	78	YEARS
CDISCPIL...	DM	01-701-1411	1411						2012-11-24			701	76	YEARS
CDISCPIL...	DM	01-701-1415	1415	2013-09-23	2014-03-24	2013-09-23	2014-03-24		2014-03-24...			701	85	YEARS
CDISCPIL...	DM	01-701-1429	1429	2013-03-19	2013-04-30	2013-03-19	2013-04-30		2013-04-30...			701	84	YEARS
CDISCPIL...	DM	01-701-1440	1440	2013-08-08	2014-02-05	2013-08-08	2014-02-05		2014-02-05...			701	70	YEARS
CDISCPIL...	DM	01-701-1442	1442	2013-10-26	2014-04-26	2013-10-26	2014-04-26		2014-04-26...			701	57	YEARS
CDISCPIL...	DM	01-701-1444	1444	2013-01-05	2013-02-13	2013-01-05	2013-02-12		2013-06-20			701	63	YEARS
CDISCPIL...	DM	01-702-1082	1082	2013-07-26	2013-11-17	2013-07-26	2013-10-13		2013-11-17...			702	84	YEARS
CDISCPIL...	DM	01-703-1042	1042	2013-03-02	2013-08-31	2013-03-02	2013-08-31		2013-08-31...			703	64	YEARS
CDISCPIL...	DM	01-703-1076	1076	2013-10-25	2013-12-24	2013-10-25	2013-12-24		2014-01-21			703	69	YEARS

Showing the SDTM tables for the datasets "DM", "AE", “CM" and "LB":

File Tools Search Options

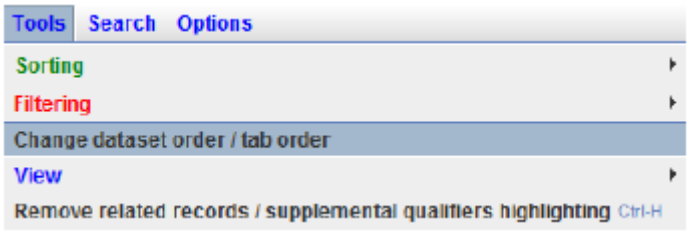
DM AE CM LB

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI	LBSTRESC	LBSTRESN
CDISCPIL...	LB	01-701-1015	17	GLUC	Glucose	CHEMISTRY	85	mg/dL	50	250	4.71835	4.71835
CDISCPIL...	LB	01-701-1015	54	GLUC	Glucose	CHEMISTRY	84	mg/dL	50	250	4.66284	4.66284
CDISCPIL...	LB	01-701-1015	88	GLUC	Glucose	CHEMISTRY	79	mg/dL	50	250	4.38529	4.38529
CDISCPIL...	LB	01-701-1015	118	GLUC	Glucose	CHEMISTRY	92	mg/dL	50	250	5.10692	5.10692
CDISCPIL...	LB	01-701-1015	148	GLUC	Glucose	CHEMISTRY	82	mg/dL	50	250	4.55182	4.55182
CDISCPIL...	LB	01-701-1015	179	GLUC	Glucose	CHEMISTRY	87	mg/dL	50	250	4.82937	4.82937
CDISCPIL...	LB	01-701-1015	213	GLUC	Glucose	CHEMISTRY	86	mg/dL	50	250	4.77386	4.77386
CDISCPIL...	LB	01-701-1015	243	GLUC	Glucose	CHEMISTRY	88	mg/dL	50	250	4.88488	4.88488
CDISCPIL...	LB	01-701-1015	274	GLUC	Glucose	CHEMISTRY	81	mg/dL	50	250	4.49631	4.49631
CDISCPIL...	LB	01-701-1015	308	GLUC	Glucose	CHEMISTRY	92	mg/dL	50	250	5.10692	5.10692
CDISCPIL...	LB	01-701-1015	32	RBC	Erythrocytes	HEMATOL...	5.30	MILL/uL	3.9	5.5	5.3	5.3
CDISCPIL...	LB	01-701-1015	69	RBC	Erythrocytes	HEMATOL...	5.30	MILL/uL	3.9	5.5	5.3	5.3
CDISCPIL...	LB	01-701-1015	101	RBC	Erythrocytes	HEMATOL...	5.00	MILL/uL	3.9	5.5	5	5
CDISCPIL...	LB	01-701-1015	131	RBC	Erythrocytes	HEMATOL...	5.10	MILL/uL	3.9	5.5	5.1	5.1
CDISCPIL...	LB	01-701-1015	161	RBC	Erythrocytes	HEMATOL...	5.10	MILL/uL	3.9	5.5	5.1	5.1
CDISCPIL...	LB	01-701-1015	194	RBC	Erythrocytes	HEMATOL...	5.00	MILL/uL	3.9	5.5	5	5
CDISCPIL...	LB	01-701-1015	226	RBC	Erythrocytes	HEMATOL...	5.20	MILL/uL	3.9	5.5	5.2	5.2
CDISCPIL...	LB	01-701-1015	256	RBC	Erythrocytes	HEMATOL...	4.90	MILL/uL	3.9	5.5	4.9	4.9
CDISCPIL...	LB	01-701-1015	289	RBC	Erythrocytes	HEMATOL...	5.30	MILL/uL	3.9	5.5	5.3	5.3
CDISCPIL...	LB	01-701-1015	321	RBC	Erythrocytes	HEMATOL...	5.40	MILL/uL	3.9	5.5	5.4	5.4
CDISCPIL...	LB	01-701-1015	38	WBC	Leukocytes	HEMATOL...	5.78	THOU/uL	3.8	10.7	5.78	5.78
CDISCPIL...	LB	01-701-1015	73	WBC	Leukocytes	HEMATOL...	5.33	THOU/uL	3.8	10.7	5.33	5.33
CDISCPIL...	LB	01-701-1015	103	WBC	Leukocytes	HEMATOL...	4.54	THOU/uL	3.8	10.7	4.54	4.54
CDISCPIL...	LB	01-701-1015	133	WBC	Leukocytes	HEMATOL...	4.25	THOU/uL	3.8	10.7	4.25	4.25
CDISCPIL...	LB	01-701-1015	163	WBC	Leukocytes	HEMATOL...	5.35	THOU/uL	3.8	10.7	5.35	5.35
CDISCPIL...	LB	01-701-1015	198	WBC	Leukocytes	HEMATOL...	5.24	THOU/uL	3.8	10.7	5.24	5.24
CDISCPIL...	LB	01-701-1015	228	WBC	Leukocytes	HEMATOL...	5.62	THOU/uL	3.8	10.7	5.62	5.62

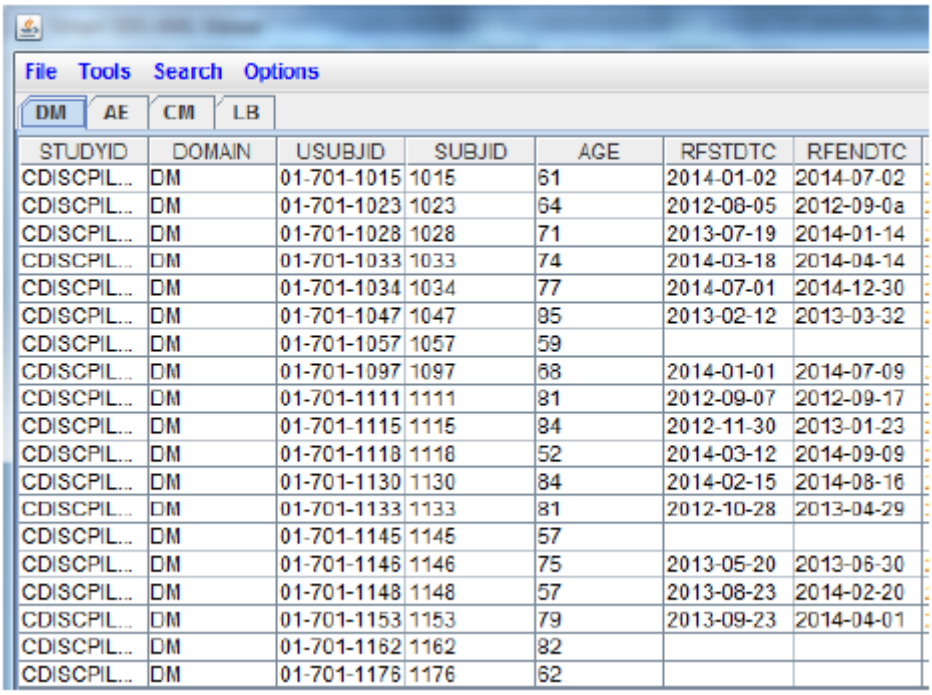
Remark that the table for the "DM" dataset is displayed as the first table, even when it is not the first file in the list of files. For the LB dataset, only those records for LBTESTCD "GLUC", "RBC" and "WBC" have been loaded.

6. Basic Features for Working with the Tables

One can easily switch between the tables by selecting (clicking) the desired tab, or using the menu "Tools - View - Show last selected table" to view the last (previously) selected table. The latter is also possible using "CTRL-B" thus allowing to quickly toggle between two tables. One can also change the order of the tabs by using the menu "Tools - Change dataset order / tab order":

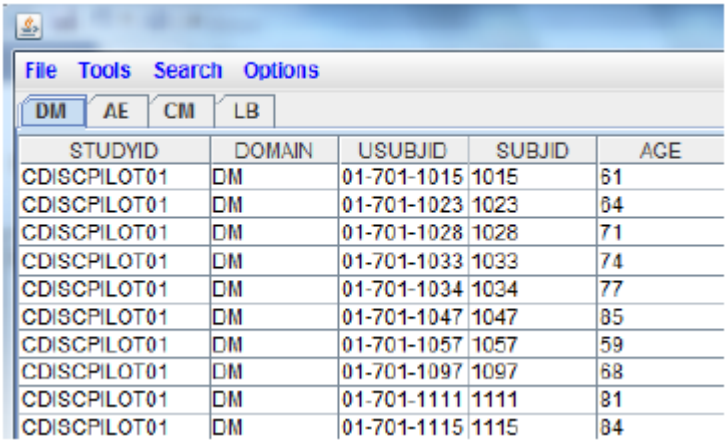


Within each table, one can change the order of the columns by using drag-and-drop. For example, if one would like to see the column "AGE" just after the column "USUBJID" in the DM table, simply put the mouse pointer on the column header of "AGE" and then, drag that column using the left mouse button pressed, and then release the left mouse button when it has arrived immediately right of the "USUBJID" column. This will result in:



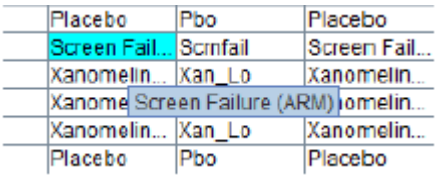
STUDYID	DOMAIN	USUBJID	SUBJID	AGE	RFSTDTC	RFENDTC
CDISCPIL...	DM	01-701-1015	1015	61	2014-01-02	2014-07-02
CDISCPIL...	DM	01-701-1023	1023	64	2012-08-05	2012-09-0a
CDISCPIL...	DM	01-701-1028	1028	71	2013-07-19	2014-01-14
CDISCPIL...	DM	01-701-1033	1033	74	2014-03-18	2014-04-14
CDISCPIL...	DM	01-701-1034	1034	77	2014-07-01	2014-12-30
CDISCPIL...	DM	01-701-1047	1047	85	2013-02-12	2013-03-32
CDISCPIL...	DM	01-701-1057	1057	59		
CDISCPIL...	DM	01-701-1097	1097	68	2014-01-01	2014-07-09
CDISCPIL...	DM	01-701-1111	1111	81	2012-09-07	2012-09-17
CDISCPIL...	DM	01-701-1115	1115	84	2012-11-30	2013-01-23
CDISCPIL...	DM	01-701-1118	1118	52	2014-03-12	2014-09-09
CDISCPIL...	DM	01-701-1130	1130	84	2014-02-15	2014-08-16
CDISCPIL...	DM	01-701-1133	1133	81	2012-10-28	2013-04-29
CDISCPIL...	DM	01-701-1145	1145	57		
CDISCPIL...	DM	01-701-1146	1146	75	2013-05-20	2013-06-30
CDISCPIL...	DM	01-701-1148	1148	57	2013-08-23	2014-02-20
CDISCPIL...	DM	01-701-1153	1153	79	2013-09-23	2014-04-01
CDISCPIL...	DM	01-701-1162	1162	82		
CDISCPIL...	DM	01-701-1176	1176	62		

One can also easily resize the display width of each column by clicking on the separator between two column headers, and then dragging it to the left or the right. For example, clicking the separator on the column header between "STUDYID" and "DOMAIN" and then dragging it to the right results in:

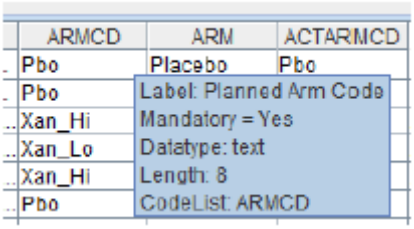


STUDYID	DOMAIN	USUBJID	SUBJID	AGE
CDISCPILOT01	DM	01-701-1015	1015	61
CDISCPILOT01	DM	01-701-1023	1023	64
CDISCPILOT01	DM	01-701-1028	1028	71
CDISCPILOT01	DM	01-701-1033	1033	74
CDISCPILOT01	DM	01-701-1034	1034	77
CDISCPILOT01	DM	01-701-1047	1047	85
CDISCPILOT01	DM	01-701-1057	1057	59
CDISCPILOT01	DM	01-701-1097	1097	68
CDISCPILOT01	DM	01-701-1111	1111	81
CDISCPILOT01	DM	01-701-1115	1115	84

When navigating the mouse over a cell without clicking, a tooltip will show up displaying the full contents of the cell (also when only a part of the cell is visible in the cell itself), with additionally (in brackets) the variable name of that data point[2]. For example:



Additional information about a variable can be obtained by moving the mouse pointer over a column header. A tooltip then pops up, displaying the variable label ("def:Label" in define.xml 1.0 or "Description" in define.xml 2.0), the data type, whether the field is mandatory, the length for the field, and the name of the code list if applicable.



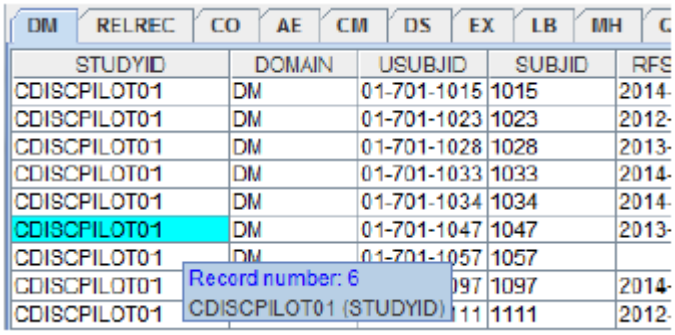
To select a single cell, just click the cell with the left mouse button. Later we will see how this can be used for creating subsets of data to find related records in other datasets/domains easily and quickly.

[2] This feature has great potential. For example, when a coded value is given such as in LBLOINC (lab test LOINC code), a lookup in the LOINC database can be performed and the details / explanation of the code can be displayed.

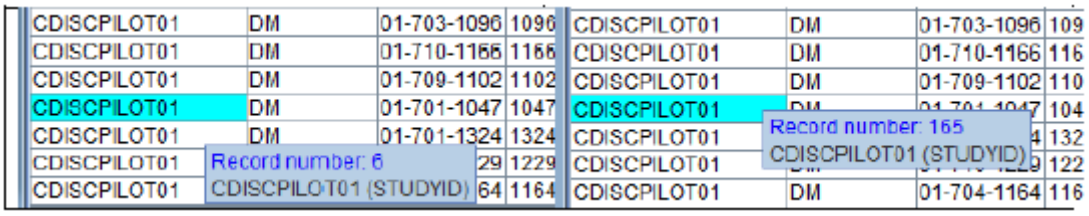
DM | AE | CM

STUDYID	DOMAIN	USUBJID	AGE	SUBJID
CDISCPILOT01	DM	01-701-1015	61	1015
CDISCPILOT01	DM	01-701-1023	64	1023
CDISCPILOT01	DM	01-701-1028	71	1028
CDISCPILOT01	DM	01-701-1033	74	1033
CDISCPILOT01	DM	01-701-1034	77	1034
CDISCPILOT01	DM	01-701-1047	85	1047

In Dataset-XML, each record has a record number (given by the "data:ItemGroupDataSeq" attribute on the "ItemGroupData" element in the Dataset-XML file). This record number can be made visible by hovering the mouse over the "STUDYID" cell:



The record number is really a property of the record itself, not just a sequence number in the view (like in the SASViewer). So, when sorting or filtering the data, the record number for a specific record remains the same:



This makes it easy for reviewers to reference a specific record in discussions with the sponsor or with colleagues or other parties.

7. Basic Sorting within the Table

Basic sorting on the contents of a single column can be done by clicking the column header using the left mouse button. For example, to sort the subjects in the DM table by age, one clicks the column header of the "AGE" column, resulting in:

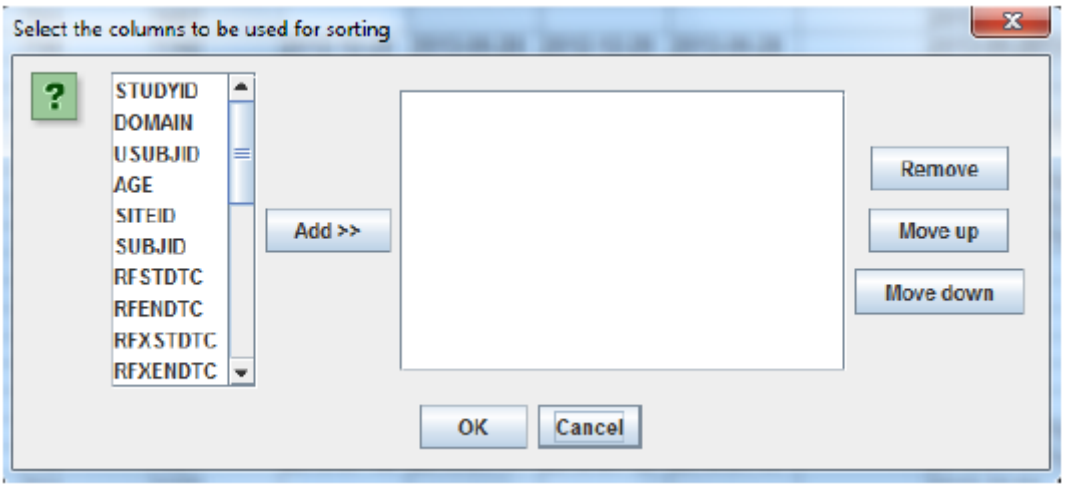
STUDYID	DOMAIN	USUBJID	AGE ▲	SUBJID
CDISCPILOT01	DM	01-715-1134	50	1134
CDISCPILOT01	DM	01-701-1341	51	1341
CDISCPILOT01	DM	01-701-1118	52	1118
CDISCPILOT01	DM	01-701-1356	54	1356
CDISCPILOT01	DM	01-709-1007	54	1007
CDISCPILOT01	DM	01-701-1180	56	1180
CDISCPILOT01	DM	01-701-1239	56	1239
CDISCPILOT01	DM	01-701-1287	56	1287
CDISCPILOT01	DM	01-705-1031	56	1031
CDISCPILOT01	DM	01-705-1280	56	1280
CDISCPILOT01	DM	01-710-1235	56	1235
CDISCPILOT01	DM	01-701-1145	57	1145
CDISCPILOT01	DM	01-701-1148	57	1148
CDISCPILOT01	DM	01-701-1240	57	1240
CDISCPILOT01	DM	01-701-1442	57	1442
CDISCPILOT01	DM	01-708-1296	57	1296
CDISCPILOT01	DM	01-701-1057	59	1057
CDISCPILOT01	DM	01-708-1342	59	1342
CDISCPILOT01	DM	01-715-1155	59	1155
CDISCPILOT01	DM	01-716-1331	59	1331
CDISCPILOT01	DM	01-705-1292	60	1292
CDISCPILOT01	DM	01-706-1049	60	1049
CDISCPILOT01	DM	01-709-1306	60	1306

The subjects are displayed in ascending order of age (youngest subject first). A second click on the same column header then sorts the subjects in descending order of age (oldest subject first):

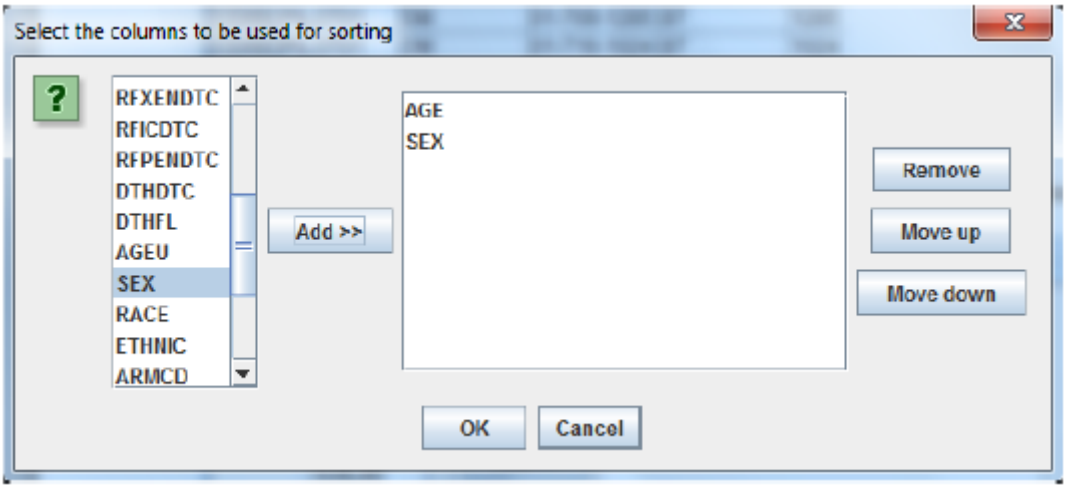
STUDYID	DOMAIN	USUBJID	AGE ▼	SUBJID
CDISCPILOT01	DM	01-705-1058	89	1058
CDISCPILOT01	DM	01-710-1083	89	1083
CDISCPILOT01	DM	01-710-1376	89	1376
CDISCPILOT01	DM	01-703-1295	88	1295
CDISCPILOT01	DM	01-708-1067	88	1067
CDISCPILOT01	DM	01-709-1237	88	1237
CDISCPILOT01	DM	01-710-1002	88	1002
CDISCPILOT01	DM	01-710-1368	88	1368
CDISCPILOT01	DM	01-710-1443	88	1443
CDISCPILOT01	DM	01-714-1035	88	1035
CDISCPILOT01	DM	01-701-1387	87	1387
CDISCPILOT01	DM	01-704-1233	87	1233
CDISCPILOT01	DM	01-705-1199	87	1199
CDISCPILOT01	DM	01-705-1421	87	1421
CDISCPILOT01	DM	01-708-1352	87	1352
CDISCPILOT01	DM	01-708-1353	87	1353
CDISCPILOT01	DM	01-709-1285	87	1285
CDISCPILOT01	DM	01-716-1024	87	1024
CDISCPILOT01	DM	01-704-1241	86	1241
CDISCPILOT01	DM	01-705-1349	86	1349
CDISCPILOT01	DM	01-708-1236	86	1236
CDISCPILOT01	DM	01-709-1081	86	1081
CDISCPILOT01	DM	01-710-1271	86	1271
CDISCPILOT01	DM	01-711-1022	86	1022

8. Multiple Column Sorting within the Table

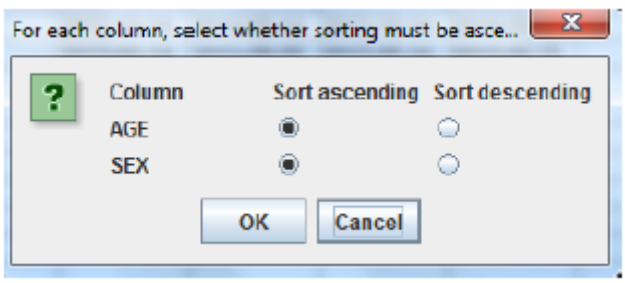
To sort multiple columns, use the menu "Tools – Sorting – Sort table", or use the keyboard shortcut CTRL-T. The following dialogue is displayed in the case of the DM table:



One can now select columns from the list on the left side, and add them to the "sorting list" (on the right side) by clicking the "Add" button. For example, if we want to sort the subjects by age (primary sort) and then by sex, we select "AGE" and "SEX" and then click the "Add" button. This results in:

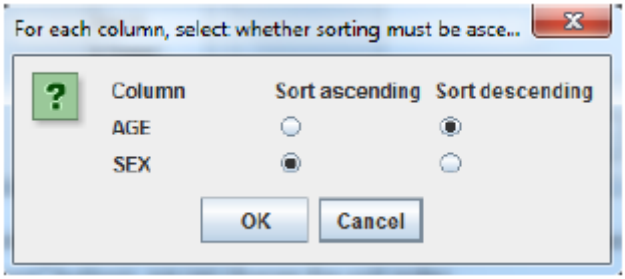


Using the "Remove" button, we can remove columns from the list. Using the "Move up" and "Move down" buttons, we can change the sort order. Essentially, we can add as many columns as we want to the list for sorting. When now clicking the "OK" button, the following dialogue is displayed:



A window shows up, which asks whether we would like to sort in ascending or descending order. In case the variable is of data type "text", the sorting will be done alphabetically.

If, for example, we would like to have the female subjects to come before the male subjects (secondary sort), and the oldest subjects first (primary sort), we need:



Clicking "OK" results in:

STUDYID	DOMAIN	USUBJID	AGE	SEX	SUBJI
CDISCPILOT01	DM	01-710-1083	89	F	1083
CDISCPILOT01	DM	01-710-1376	89	F	1376
CDISCPILOT01	DM	01-705-1058	89	M	1058
CDISCPILOT01	DM	01-703-1295	88	F	1295
CDISCPILOT01	DM	01-709-1237	88	F	1237
CDISCPILOT01	DM	01-710-1368	88	F	1368
CDISCPILOT01	DM	01-710-1443	88	F	1443
CDISCPILOT01	DM	01-714-1035	88	F	1035
CDISCPILOT01	DM	01-708-1067	88	M	1067
CDISCPILOT01	DM	01-710-1002	88	M	1002
CDISCPILOT01	DM	01-701-1387	87	F	1387
CDISCPILOT01	DM	01-704-1233	87	F	1233
CDISCPILOT01	DM	01-708-1352	87	F	1352
CDISCPILOT01	DM	01-708-1353	87	F	1353
CDISCPILOT01	DM	01-716-1024	87	F	1024
CDISCPILOT01	DM	01-705-1199	87	M	1199
CDISCPILOT01	DM	01-705-1421	87	M	1421
CDISCPILOT01	DM	01-709-1285	87	M	1285
CDISCPILOT01	DM	01-705-1349	86	F	1349
CDISCPILOT01	DM	01-708-1236	86	F	1236
CDISCPILOT01	DM	01-709-1081	86	F	1081

Remark again that we can use as much columns as desired for sorting.

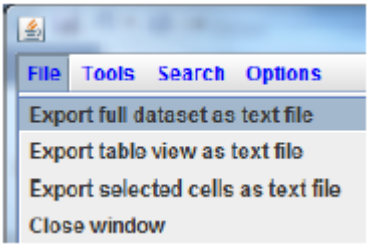
9. Removing the Sorting

To remove all sorting, and return to the original sequence (i.e. the sequence order as in the dataset), use the menu "Tools – Sorting – Unsort table", or simply use the keyboard shortcut CTRL-U.

10. Exporting Data Tables as Text Files

In some cases, it may be useful to be able to import Dataset-XML data in other software tools, e.g. in software tools that do not understand XML yet. In such a case, one can always export the data, or part of the data, as a text file.

In order to export data as text, use the menu "File – Export full dataset as text file" or "File – Export table view as text file" or "File – Export selected cells as text file":



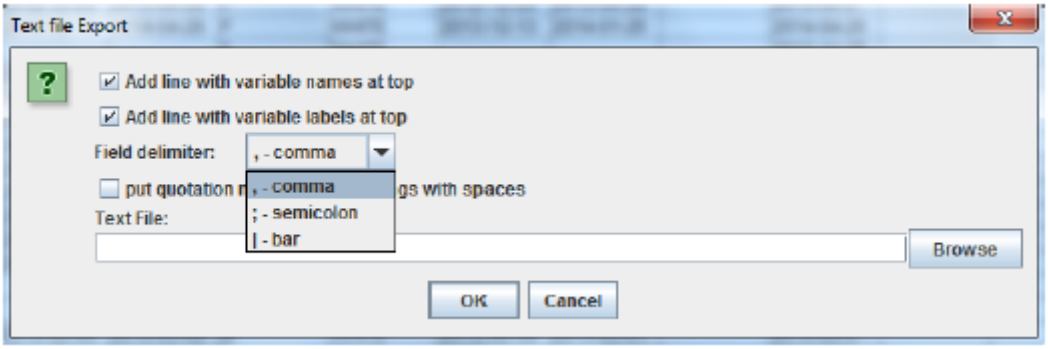
The difference between these three options is the following:

- If you choose "Export full dataset as text file", the whole dataset will be exported, even if you first applied some filters (exception: when you implemented filters during loading of the dataset) and with the data in the original order.
- If you choose "Export table view as text file", only the data that is currently displayed in the table is exported, i.e. data that was "filtered out" will not be exported. The data is exported in the order that is currently displayed. So for example, if you sorted the table on a lab value, the data will be exported in the order of that value.
- If you choose "Export selected cells as text file", only those cells that are currently selected will be exported. For example, if your current selection is:

SUBJID	AGE	RFSTDTC	RFENDTC	SEX	RACE	RFXST
1134	50			F	WHITE	
1341	51	2013-01-05	2013-02-07	M	WHITE	2013-0
1118	52	2014-03-12	2014-09-09	M	WHITE	2014-0
1356	54			F	WHITE	
1007	54	2012-07-31	2012-09-01	F	WHITE	2012-0
1180	56	2013-02-12	2013-03-23	M	WHITE	2013-0
1239	56	2014-01-11	2014-07-11	M	WHITE	2014-0
1297	56	2014-01-25	2014-07-26	F	WHITE	2014-0

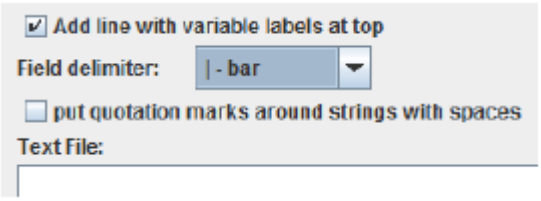
Then only data for subjects 1118, 1356, 1007, and 1239, and for the columns "SUBJID", "AGE", "RFSTDTC", "SEX" and "RACE" will be exported.

The following dialogue is then displayed:



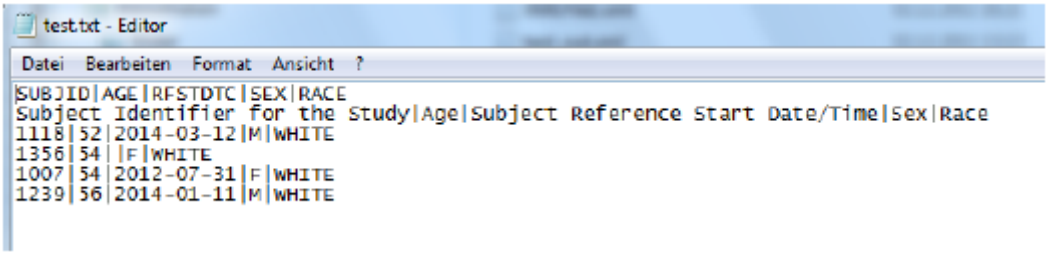
The first checkbox allows you to add a line with the variable names at the top of the export file. The second checkbox allows you to add a line with the variable labels at the top of the export file (or after the line with the variable names). For the separation of the fields, you have the choice between the comma (i.e. file will be a CSV file), the semicolon, or the vertical bar (as in HL7-v2 messages or the CDISC Lab Standard).

There is also a checkbox "put quotation marks around strings with spaces":



If this box is checked, the system will check whether a data value (or a variable label) contains one or more spaces, and if so, will add quotation marks at the beginning and end of the string.

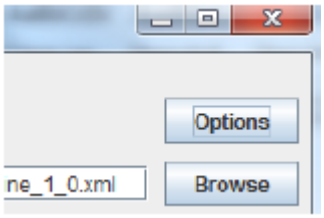
We then only need to select a file to export to, and click the "OK" button. An example result is:



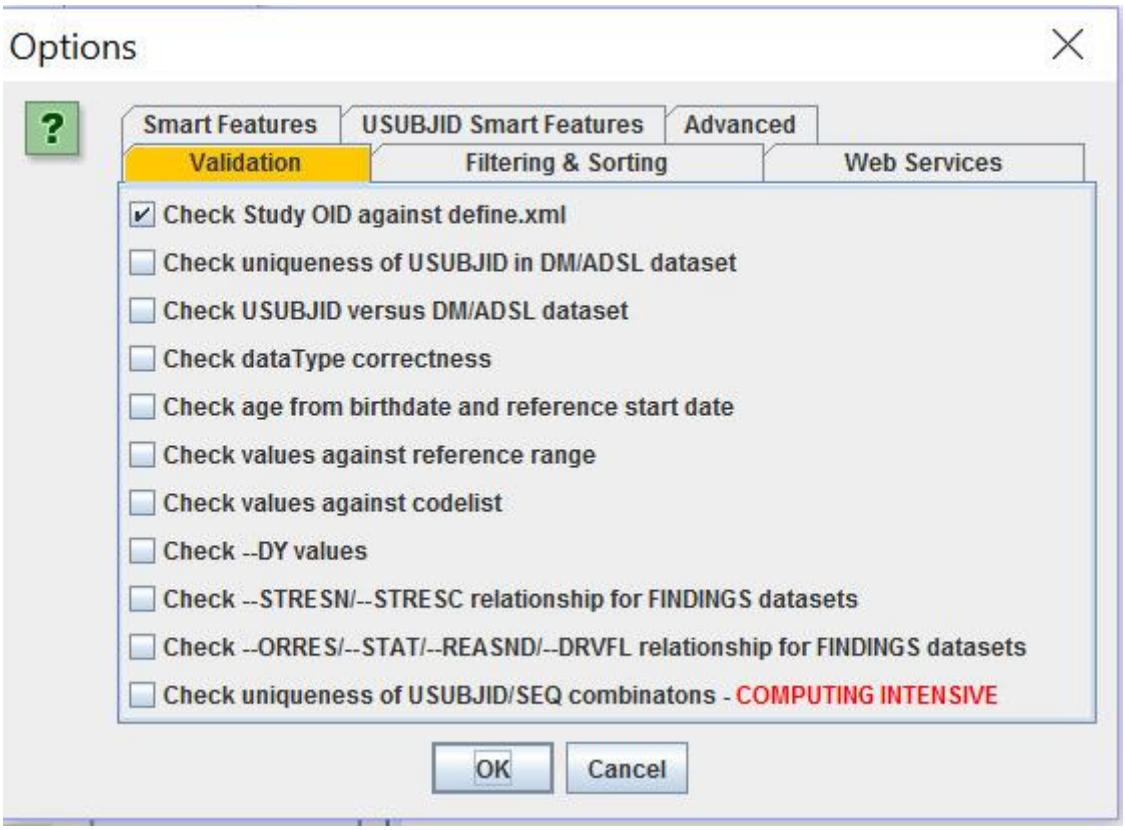
11. Simple Validation of the Datasets

Now close the window containing the tables by either clicking the cross in the upper right corner or by using the "File – Close window" menu.

In the main window, now click the button "Options" on the upper right side of the window:



This results in a new dialogue:



Remark that the contents of this dialogue may differ upon which version you are using. As the software is open source, organizations/companies/developers may have added additional features and options.

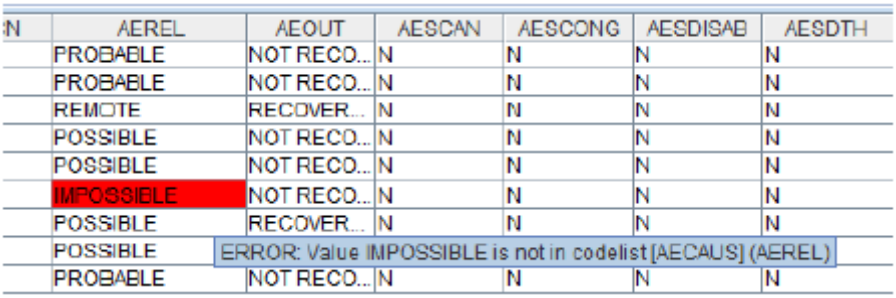
The options are divided into several categories. We will first discuss the "validation" options.

The following validation features have currently been implemented:

- Check Study OID of the dataset against the Study OID in the define.xml. This option is switched on by default – see later.
- Check uniqueness of USUBJID in DM/ADSL dataset. A basic rule of SDTM/SEND and of ADSL datasets is that the value of "USUBJID" is unique within the DM or ADSL dataset. When switching on this option, the uniqueness of "USUBJID" will be checked during loading of the datasets, and cells violating this rule will be marked (i.e. obtain a red background).
- Check USUBJID versus DM/ADSL dataset. Each value of "USUBJID" in any dataset must also occur in the DM dataset (for SDTM and SEND) or in the ADSL dataset (for ADaM). When checked, each occurrence of USUBJID will be checked against the values in the DM or ADSL dataset. If a violation is found, the cell is marked.
- Check data type correctness. When checked, the system will check each value upon loading whether it is correct with respect to the data type defined in the define.xml file.
- Check age from birthdate and reference start date. The age of each subject (in the DM or ADSL dataset) is calculated and compared to the value given in the "AGE" column. If a discrepancy is found, the "AGE" cell is marked.
- Check values against reference range. Values of the variables --STRESN ("Numeric Result/Finding in Standard Units") in the "Findings" data sets are compared to the values given in the variables --STNRLO ("Reference Range Lower Limit") and --STNRHI (("Reference Range Higher Limit"). If the value in --STRESN is outside the reference range, the --STRESN cell is coloured yellow. E.g.:

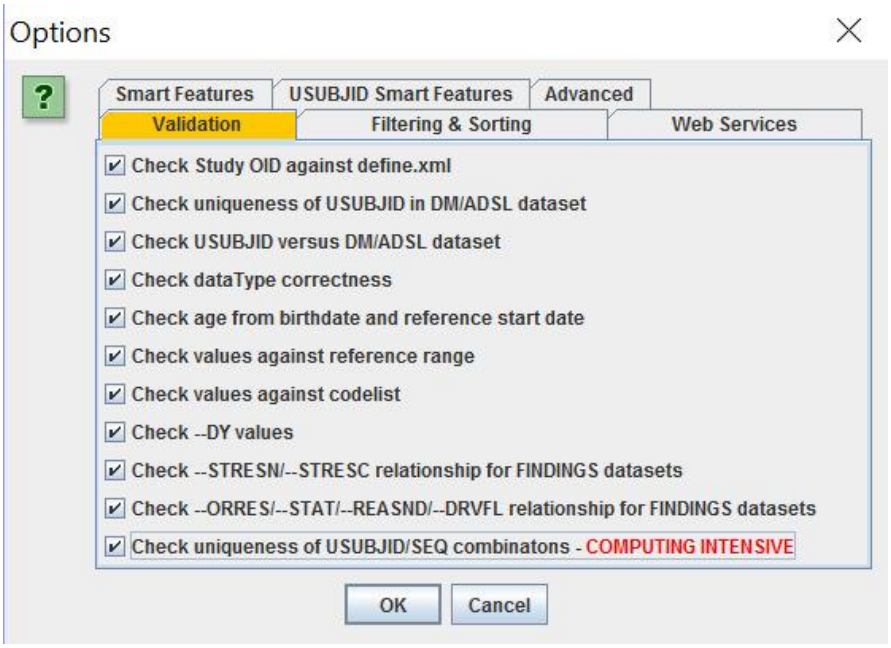
LBSTRESC	LBSTRESN	LBSTRESU	LBSTNRLO	LBSTNRHI
38	38	g/L	33	49
34	34	U/L	35	115
50	50	U/L	35	115
41	41	U/L	35	115
43	43	U/L	35	115
47	47	U/L	35	115
53	53	U/L	35	115
41	41	U/L	35	115
45	45	U/L	35	115
46	46	U/L	35	115
44	44	U/L	35	115
27	27	U/L	6	34
41	41	U/L	6	34

- Check values against code list. In case the variable is governed by a code list (except for an external code list[3]), the value is checked against that code list (as given in the define.xml file). If the given value is not a value from that code list, the cell is marked as having an error (i.e. coloured red). For example:



- Check uniqueness of USUBJID / SEQ combinations. This option allows to check the uniqueness of the combination of USUBJID and --SEQ in the datasets where both occur. If a combination is found not to be unique within that dataset, both cells are marked as having an error. Remark that currently, the system does not check uniqueness across split domains. Validation of the uniqueness of the combination of USUBJID and --SEQ is computing intensive and will usually considerably slow down the loading and validation of the data sets, especially in the case of large data sets. In many cases, it is advised to do this validation using OpenCDISC.

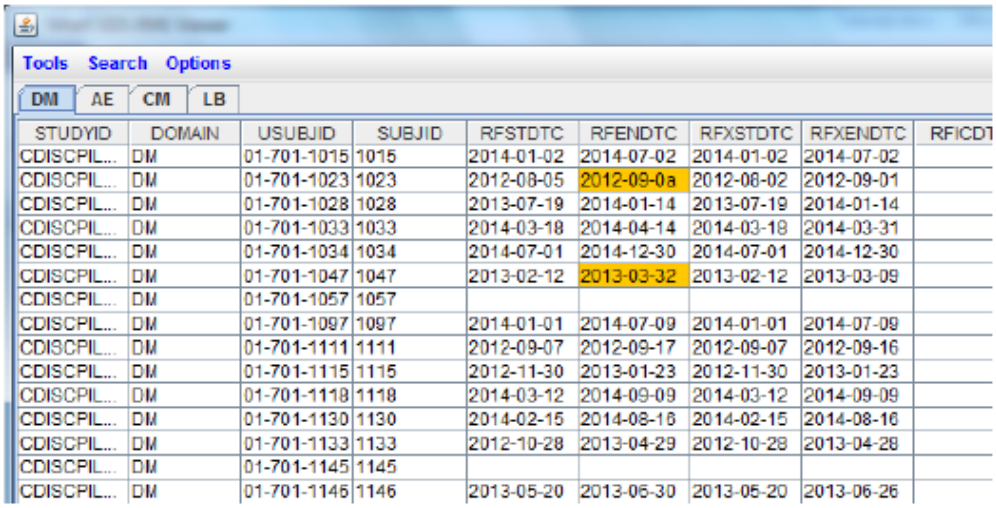
Let us now switch on a number of these options:



[3] This may become a very useful feature in the future using web services

Remark that for this tutorial, we intentionally introduced some errors and discrepancies in the data sets.

After clicking OK in the "Options" dialogue, clicking the "Start" button now reloads all the files. As we have added extra validation and an extra data set, loading of the files will take somewhat more time. The result is:



Tools Search Options

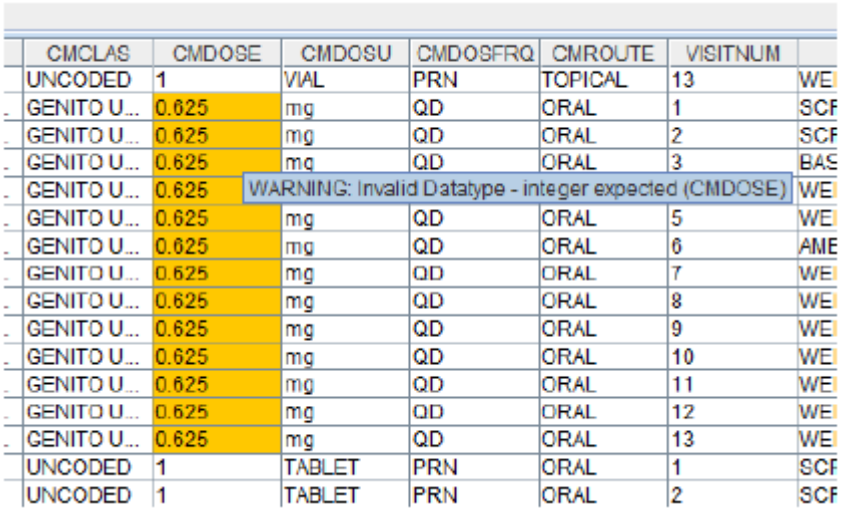
DM AE CM LB

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC	RFICDT
CDISCPIL...	DM	01-701-1015	1015	2014-01-02	2014-07-02	2014-01-02	2014-07-02	
CDISCPIL...	DM	01-701-1023	1023	2012-08-05	2012-09-0a	2012-08-02	2012-09-01	
CDISCPIL...	DM	01-701-1028	1028	2013-07-19	2014-01-14	2013-07-19	2014-01-14	
CDISCPIL...	DM	01-701-1033	1033	2014-03-18	2014-04-14	2014-03-18	2014-03-31	
CDISCPIL...	DM	01-701-1034	1034	2014-07-01	2014-12-30	2014-07-01	2014-12-30	
CDISCPIL...	DM	01-701-1047	1047	2013-02-12	2013-03-32	2013-02-12	2013-03-09	
CDISCPIL...	DM	01-701-1057	1057					
CDISCPIL...	DM	01-701-1097	1097	2014-01-01	2014-07-09	2014-01-01	2014-07-09	
CDISCPIL...	DM	01-701-1111	1111	2012-09-07	2012-09-17	2012-09-07	2012-09-16	
CDISCPIL...	DM	01-701-1115	1115	2012-11-30	2013-01-23	2012-11-30	2013-01-23	
CDISCPIL...	DM	01-701-1118	1118	2014-03-12	2014-09-09	2014-03-12	2014-09-09	
CDISCPIL...	DM	01-701-1130	1130	2014-02-15	2014-08-16	2014-02-15	2014-08-16	
CDISCPIL...	DM	01-701-1133	1133	2012-10-28	2013-04-29	2012-10-28	2013-04-28	
CDISCPIL...	DM	01-701-1145	1145					
CDISCPIL...	DM	01-701-1146	1146	2013-05-20	2013-06-30	2013-05-20	2013-06-26	

For the DM data set, we immediately see that there are two "warnings" (cells with an orange background): the cell for RFENDTC ("Subject Reference End Date/Time") containing the value "2012-09-0a) contains an invalid value for the data type "datetime" (as defined in the define.xml file). Also the cell with the value "2013-03-32" is marked, as there is no date "February 32, 2013"[4].

Let us now switch to the CM ("Concomitant Medications") table.

After scrolling down a bit, we can e.g. find the following:



CMCLAS	CMDOSE	CMDOSU	CMDOSFRQ	CMROUTE	VISITNUM	
UNCODED	1	VIAL	PRN	TOPICAL	13	WE
GENITO U...	0.625	mg	QD	ORAL	1	SCF
GENITO U...	0.625	mg	QD	ORAL	2	SCF
GENITO U...	0.625	mg	QD	ORAL	3	BAS
GENITO U...	0.625	WARNING: Invalid Datatype - integer expected (CMDOSE)				WE
GENITO U...	0.625	mg	QD	ORAL	5	WE
GENITO U...	0.625	mg	QD	ORAL	6	AME
GENITO U...	0.625	mg	QD	ORAL	7	WE
GENITO U...	0.625	mg	QD	ORAL	8	WE
GENITO U...	0.625	mg	QD	ORAL	9	WE
GENITO U...	0.625	mg	QD	ORAL	10	WE
GENITO U...	0.625	mg	QD	ORAL	11	WE
GENITO U...	0.625	mg	QD	ORAL	12	WE
GENITO U...	0.625	mg	QD	ORAL	13	WE
UNCODED	1	TABLET	PRN	ORAL	1	SCF
UNCODED	1	TABLET	PRN	ORAL	2	SCF

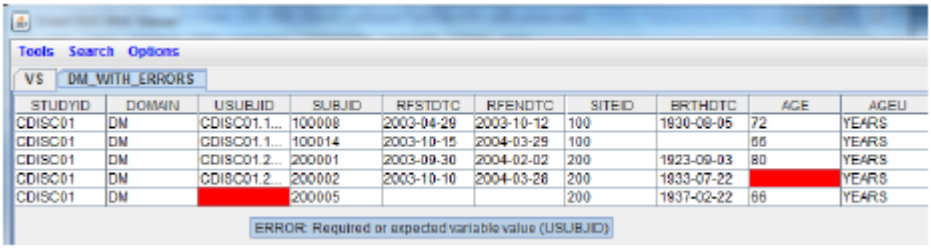
[4] Remark that also the date "2013-02-29" would be marked as being incorrect, as there is no February 29th in 2013.

Some of the cells are coloured orange (i.e. warning), and when holding the mouse over such a cell a tooltip is shown giving more information. In this case, it states that the data type is invalid, as "integer" was expected. The reason for this is probably an error in the define.xml where "integer" was defined as being the data type for CMDOSE, whereas "float" would have been more appropriate.

12. Basic Validation That Is Always Done

As already stated, even when none of the checkboxes from the "Options" dialogue is checked, a basic validation will always be performed. This basic validation is currently limited to a check whether a value is present in case the variable was defined as being mandatory (i.e. "required"[5] in SDTM/SEND/ADaM).

An example of a view that can be obtained in such a case[6] is:

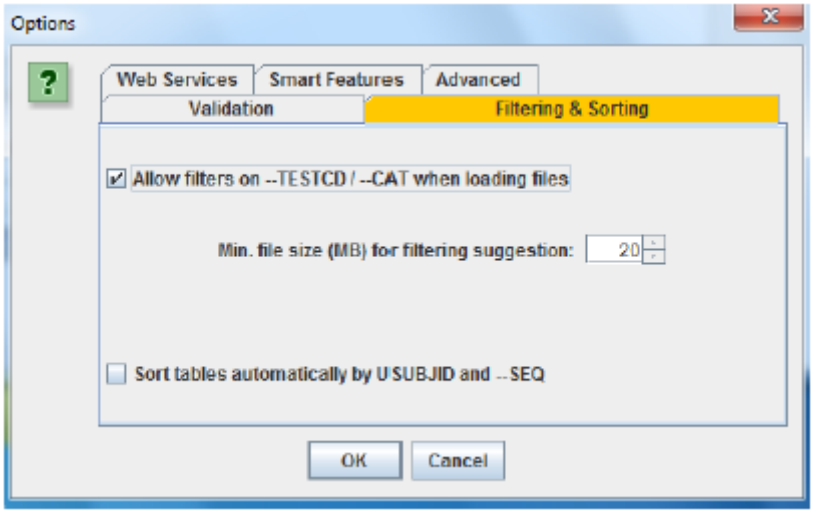


Giving an error for empty "USUBJID" and "AGE" cells, as these are "required" or "expected".

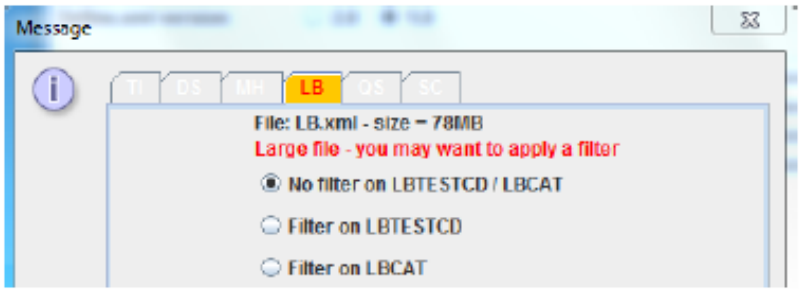
13. Filtering Options

The second tab of the "Options" panel ("Filtering and Sorting") allows us to set some parameters for filtering before loading, based on --TESTCD and/or --CAT, or to completely switch this off:

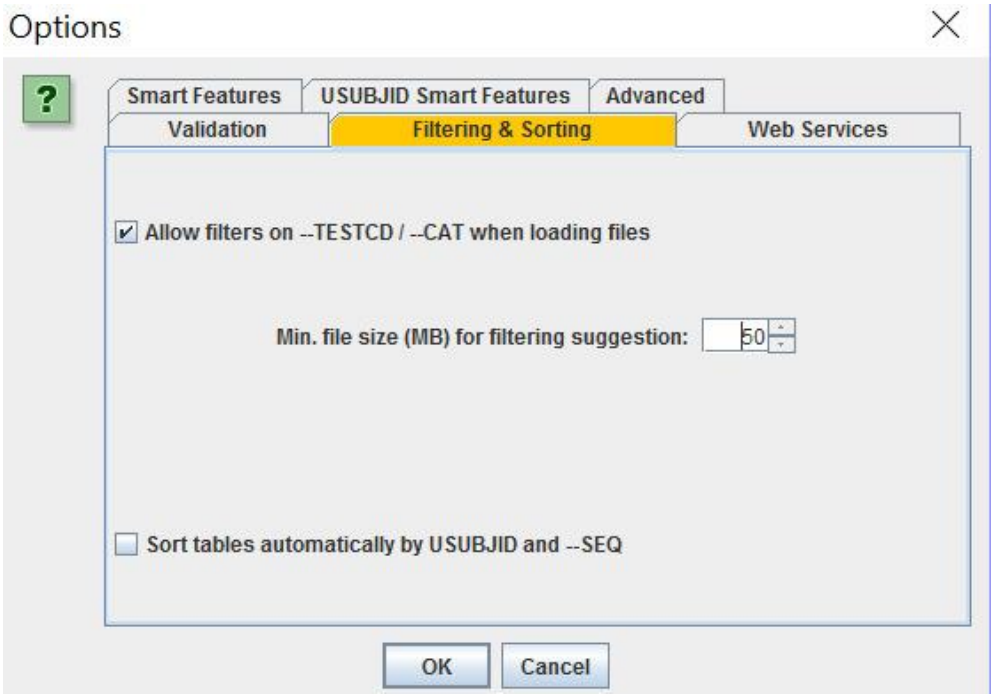
[5] No check is being done for "expected" variables, as these can have null values, depending on the value of other variables. In order to do "expected variable" validation, please use OpenCDISC.
[6] For demonstrating this feature, we deliberately removed some data points from the data file.



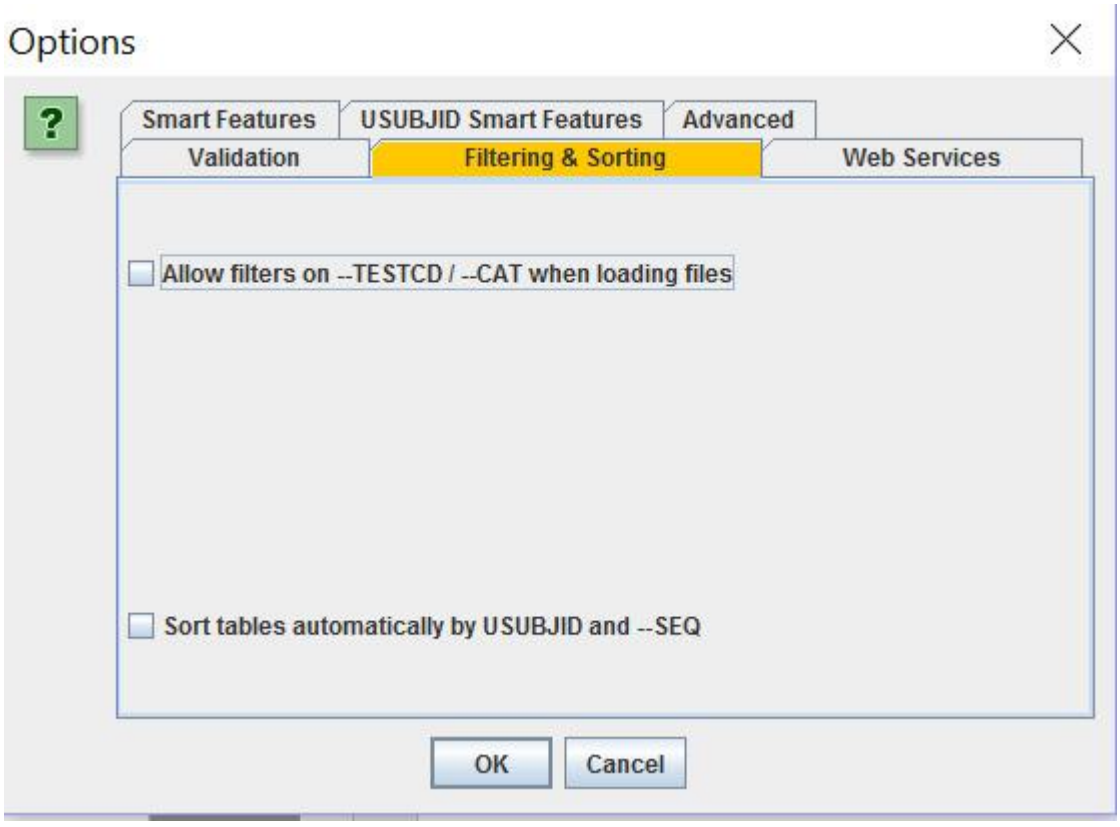
When the checkbox "Allow filters on --TESTCD / --CAT when loading files" is unchecked, the dialogue allowing to set filters for each data set will not be displayed, and no filtering on --TESTCD or --CAT will be performed at all. As we have seen, if a data set is large and has controlled terminology on --TESTCD and/or --CAT, a suggestion is displayed in the filtering panel to apply a filter. For example, for a large LB file:



The file size is displayed (78MB in this case) and a suggestion (red text) is displayed that it might be worth applying a filter due to the file size. The file size threshold for displaying this suggestion can be altered using the "Min. file size (MB) for filtering suggestions". For example, if one wants to set it to 50MB, the value can easily be changed using the little "arrows":



The value can be set between 1MB and 100MB. In case the "Allow filters" checkbox is unchecked, the line for setting the threshold will disappear from the panel:



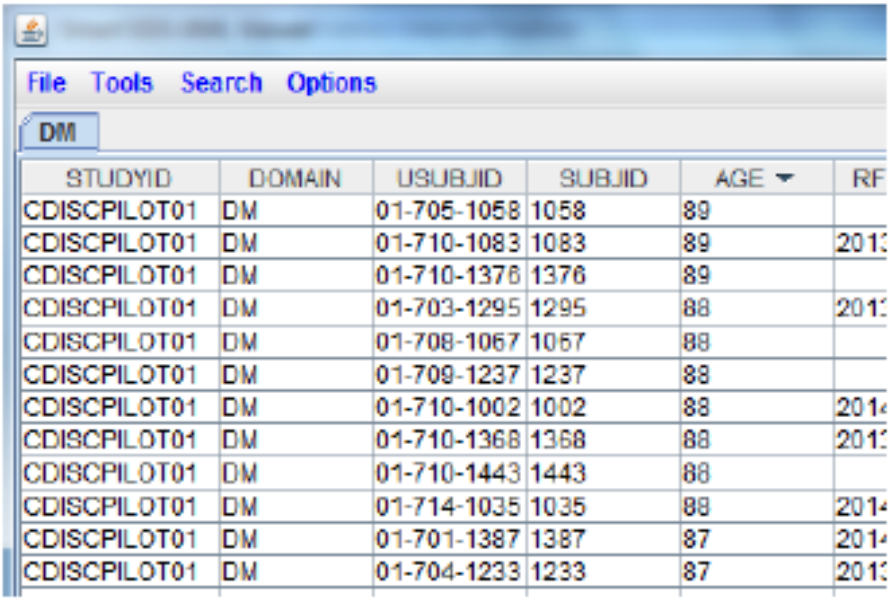
Last but not least, you can set whether the tables should automatically be sorted by USUBJID and --SEQ after loading. This might be useful in the case the original data set is not already sorted.

14. Filtering Subjects

In many cases, users will not want to see all data of all subjects. Instead, they would like to inspect data in detail of subpopulations of subjects, based e.g. on age, sex, site, lab values that are out of the normal range, etc..

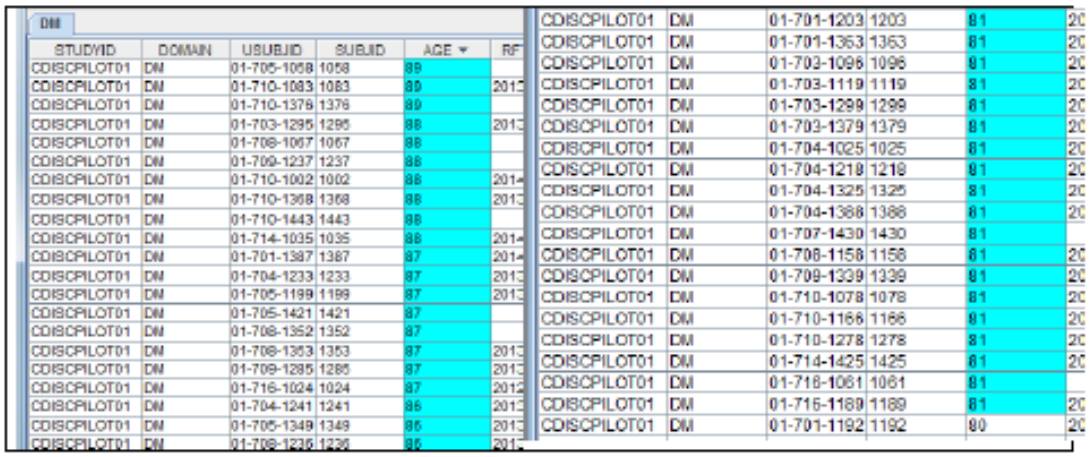
The "Smart Dataset-XML Viewer" has a good number of filtering capabilities to do so. These capabilities come on top of the filtering based on --TESTCD when loading the data files.

First, let us load the "DM" dataset (DM.xml) only (we do not load any other datasets yet). We can now sort on e.g. the age of the subjects, either by clicking the header of the "AGE" column, or by using the "Tools – Sorting". As the study is an Alzheimer study, the age of the subjects ranges between 50 and 89 years. Suppose we sort by age in a descending way:



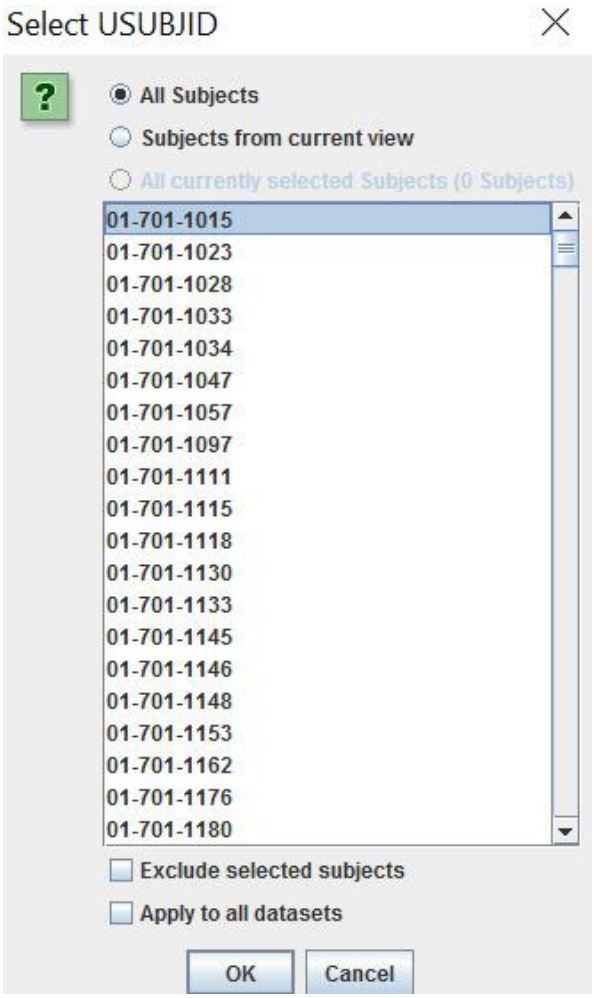
STUDYID	DOMAIN	USUBJID	SUBJID	AGE	RF
CDISCPILOT01	DM	01-705-1058	1058	89	
CDISCPILOT01	DM	01-710-1083	1083	89	201
CDISCPILOT01	DM	01-710-1376	1376	89	
CDISCPILOT01	DM	01-703-1295	1295	88	201
CDISCPILOT01	DM	01-708-1067	1067	88	
CDISCPILOT01	DM	01-709-1237	1237	88	
CDISCPILOT01	DM	01-710-1002	1002	88	201
CDISCPILOT01	DM	01-710-1368	1368	88	201
CDISCPILOT01	DM	01-710-1443	1443	88	
CDISCPILOT01	DM	01-714-1035	1035	88	201
CDISCPILOT01	DM	01-701-1387	1387	87	201
CDISCPILOT01	DM	01-704-1233	1233	87	201

We have already moved the "AGE" column to the left for better clarity. We are now interested in the lab values of all subjects of 81 years and older. We can simply select these subjects using the mouse. The selected cells are then coloured cyan:

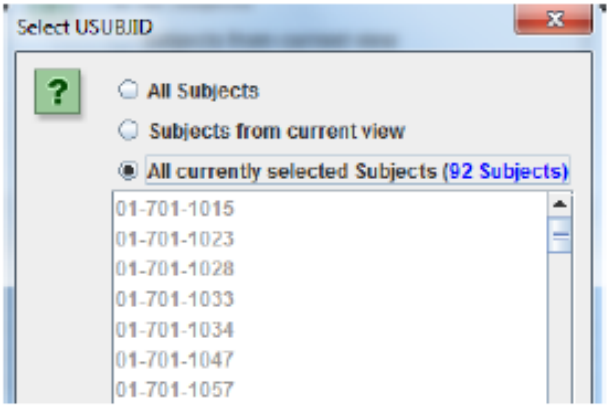


We now use the menu "Tools – Filtering – Filter on USUBJID".

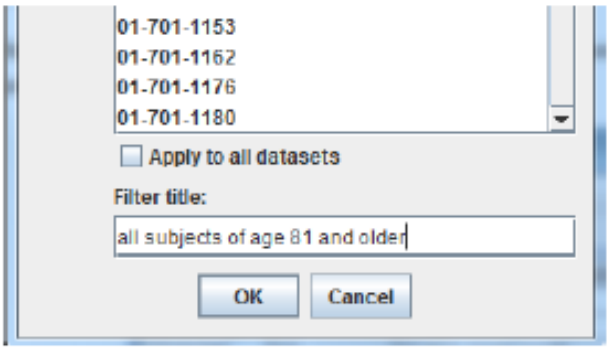
The following dialogue is displayed:



We can either do a manual selection of subjects (option "All Subjects") and pick the desired ones from the list, or only pick subjects from the current subpopulation (explained later), or filter all currently selected (i.e. those selected using the mouse) subjects. In the latter case, the dialogue is:

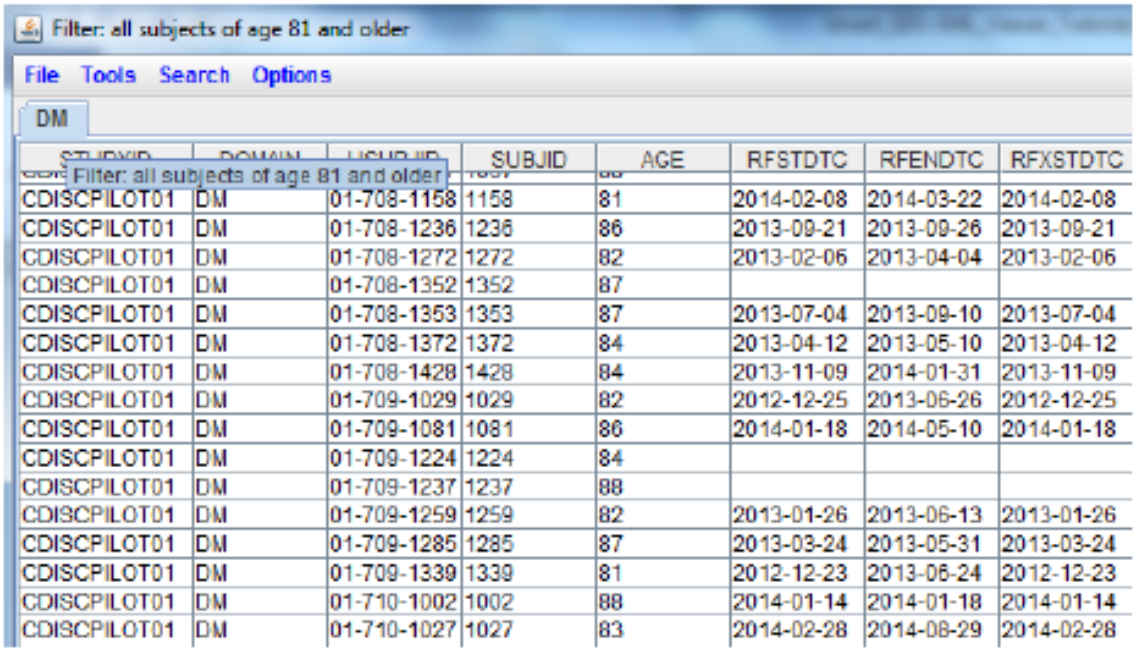


i.e. the list with subjects to pick from is grayed out - we will filter the 92 subjects that were selected using the mouse, i.e. all subjects older than 80 years.
Near the bottom, we also find the checkbox "Apply to all datasets":



If this box is checked, the filtering will also be applied to all datasets that have been loaded and for which a table exists. As we currently only have the DM table present, it does for now not matter whether this box is checked or not. We can also give the filter a title. It will then appear as a tooltip on the "DM" tab. If "apply to all datasets" was checked, the filter title is also displayed on the top of the window.

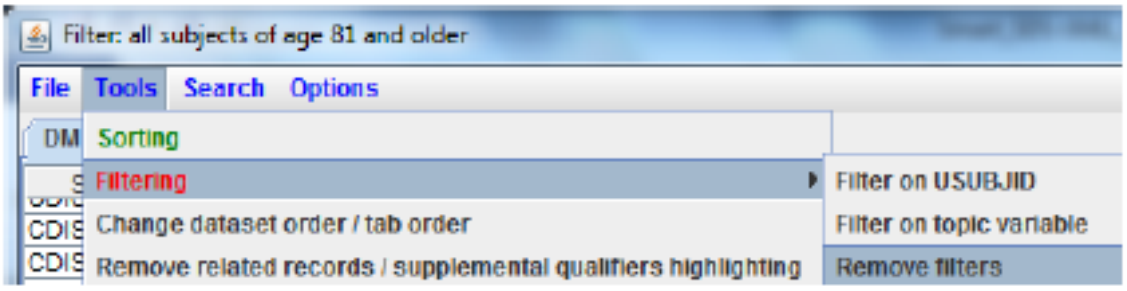
Now click the "OK" button. The DM table is refreshed and we now only see the subject of 81 years and older:



STUDYID	DOMAIN	USUBJID	SUBJID	AGE	RFSTDTC	RFENDTC	RFXSTDTC
CDISCPILOT01	DM	01-708-1158	1158	81	2014-02-08	2014-03-22	2014-02-08
CDISCPILOT01	DM	01-708-1236	1236	86	2013-09-21	2013-09-26	2013-09-21
CDISCPILOT01	DM	01-708-1272	1272	82	2013-02-06	2013-04-04	2013-02-06
CDISCPILOT01	DM	01-708-1352	1352	87			
CDISCPILOT01	DM	01-708-1353	1353	87	2013-07-04	2013-09-10	2013-07-04
CDISCPILOT01	DM	01-708-1372	1372	84	2013-04-12	2013-05-10	2013-04-12
CDISCPILOT01	DM	01-708-1428	1428	84	2013-11-09	2014-01-31	2013-11-09
CDISCPILOT01	DM	01-709-1029	1029	82	2012-12-25	2013-06-26	2012-12-25
CDISCPILOT01	DM	01-709-1081	1081	86	2014-01-18	2014-05-10	2014-01-18
CDISCPILOT01	DM	01-709-1224	1224	84			
CDISCPILOT01	DM	01-709-1237	1237	88			
CDISCPILOT01	DM	01-709-1259	1259	82	2013-01-26	2013-06-13	2013-01-26
CDISCPILOT01	DM	01-709-1285	1285	87	2013-03-24	2013-05-31	2013-03-24
CDISCPILOT01	DM	01-709-1339	1339	81	2012-12-23	2013-06-24	2012-12-23
CDISCPILOT01	DM	01-710-1002	1002	88	2014-01-14	2014-01-18	2014-01-14
CDISCPILOT01	DM	01-710-1027	1027	83	2014-02-28	2014-08-29	2014-02-28

The display order is identical to the original order (as in the original dataset). Remark that when selecting subjects, it does not matter which column is used, so we just can select the "AGE" cells, and then do a selection on subjects anyway - we do not first need to do the selection using the "USUBJID" cells.

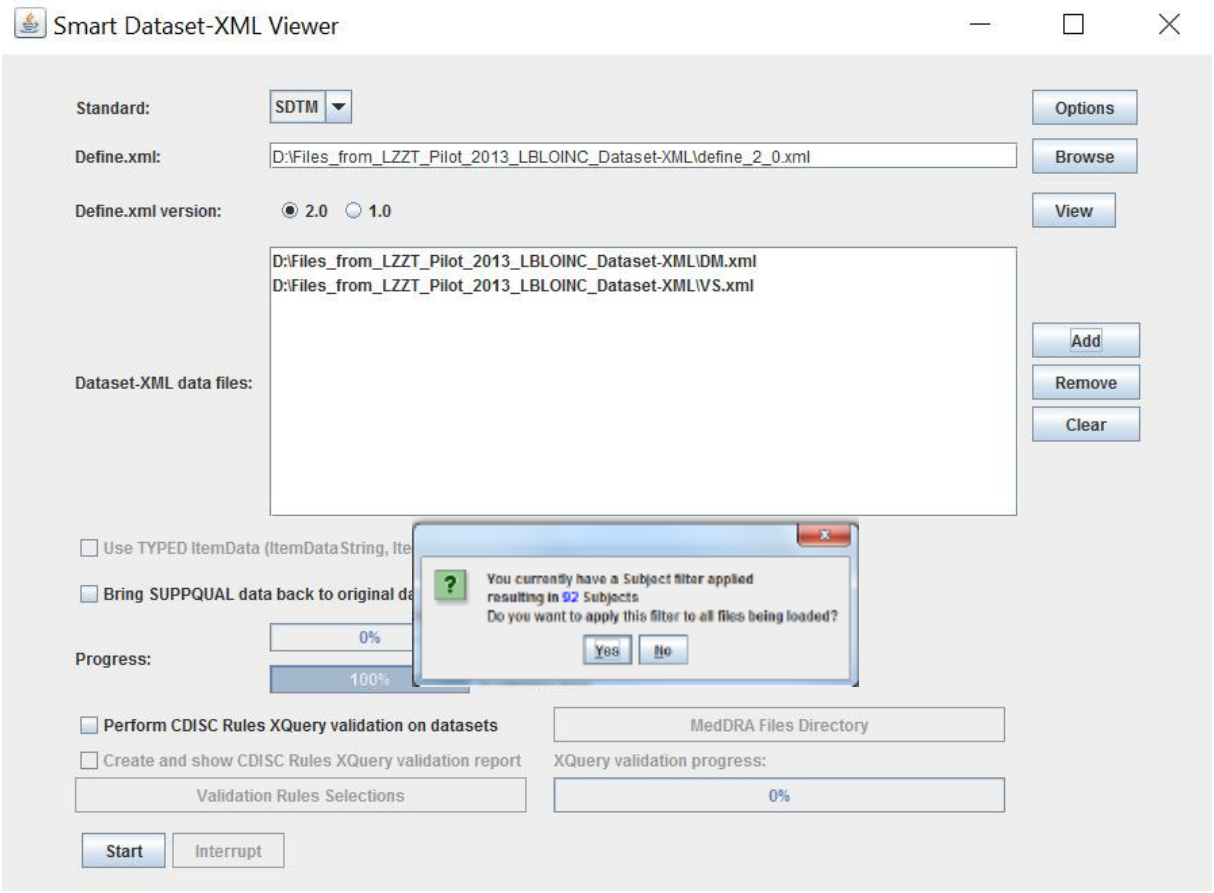
To remove the filtering, use the menu "Tools – Filtering – Remove filters":



However, we will keep working with the filtered table with subjects of 81 years and older.

We now would like to inspect the vital signs of this population of subjects. As the VS table is a pretty large one, we only want to load the VS table so that it only contains the vital data of our population (i.e. the subjects of 81 years and older). To do so, close the window with the DM table.

Then add the VS.xml dataset to the list with files that need to be loaded, and click the "Start" button. The following dialogue appears:



It states that a filter has been applied in the previous step (resulting in 92 subjects of 81 years and older), and now asks you whether you also want that filter being applied to the VS table. If you click "Yes", only these 92 subjects will be loaded for as well the DM as the VS dataset, all others will be ignored, i.e. not the complete dataset will be loaded[7]. If "No" is clicked, the complete "DM" as well as the complete "VS" dataset will be loaded.

If you click "Yes", the following result is obtained:

[7] This does not only make review easier, it also saves memory, which is important in the case of very large datasets with millions of data points.

DM | VS

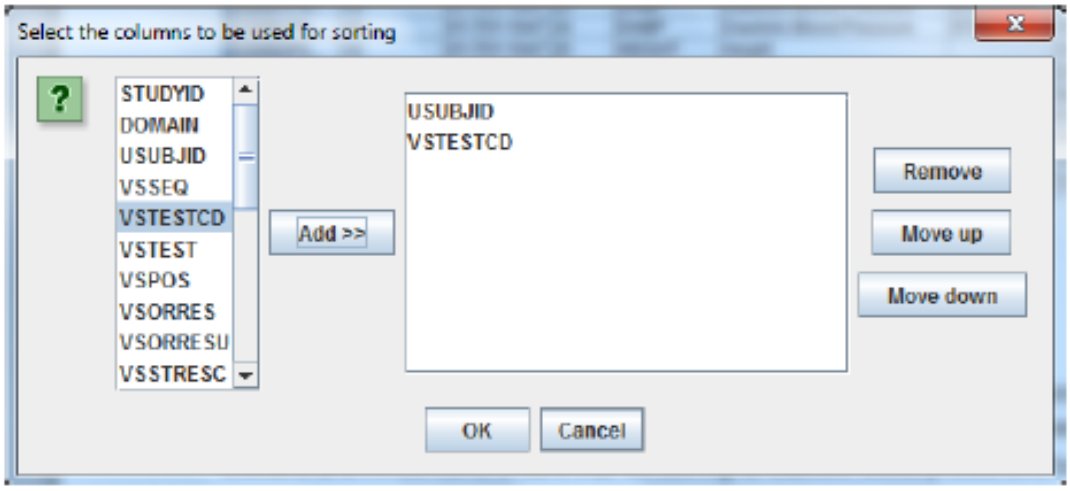
STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU
CDISCPIL...	VS	01-701-1047	1	DIABP	Diastolic Bl...	SUPINE	68	mmHg
CDISCPIL...	VS	01-701-1047	2	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	3	DIABP	Diastolic Bl...	STANDING	83	mmHg
CDISCPIL...	VS	01-701-1047	4	DIABP	Diastolic Bl...	SUPINE	65	mmHg
CDISCPIL...	VS	01-701-1047	5	DIABP	Diastolic Bl...	STANDING	86	mmHg
CDISCPIL...	VS	01-701-1047	6	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	7	DIABP	Diastolic Bl...	SUPINE	61	mmHg
CDISCPIL...	VS	01-701-1047	8	DIABP	Diastolic Bl...	STANDING	68	mmHg
CDISCPIL...	VS	01-701-1047	9	DIABP	Diastolic Bl...	STANDING	78	mmHg
CDISCPIL...	VS	01-701-1047	10	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	11	DIABP	Diastolic Bl...	STANDING	56	mmHg
CDISCPIL...	VS	01-701-1047	12	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1047	13	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	14	DIABP	Diastolic Bl...	STANDING	84	mmHg
CDISCPIL...	VS	01-701-1047	15	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1047	16	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	17	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	18	DIABP	Diastolic Bl...	STANDING	78	mmHg
CDISCPIL...	VS	01-701-1047	19	DIABP	Diastolic Bl...	SUPINE	71	mmHg
CDISCPIL...	VS	01-701-1047	20	DIABP	Diastolic Bl...	STANDING	63	mmHg
CDISCPIL...	VS	01-701-1047	21	DIABP	Diastolic Bl...	STANDING	57	mmHg
CDISCPIL...	VS	01-701-1047	22	DIABP	Diastolic Bl...	SUPINE	76	mmHg
CDISCPIL...	VS	01-701-1047	23	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	24	DIABP	Diastolic Bl...	STANDING	65	mmHg
CDISCPIL...	VS	01-701-1047	25	HEIGHT	Height		58.5	IN
CDISCPIL...	VS	01-701-1047	26	PULSE	Pulse Rate	SUPINE	53	BEATS/MIN
CDISCPIL...	VS	01-701-1047	27	PULSE	Pulse Rate	STANDING	64	BEATS/MIN
CDISCPIL...	VS	01-701-1047	28	PULSE	Pulse Rate	STANDING	61	BEATS/MIN
CDISCPIL...	VS	01-701-1047	29	PULSE	Pulse Rate	SUPINE	58	BEATS/MIN
CDISCPIL...	VS	01-701-1047	30	PULSE	Pulse Rate	STANDING	74	BEATS/MIN

Only displaying the vital signs of the subjects of 81 years and older. Remark that this filter can be combined with a filter upon loading based on VSTESTCD (when the latter has a code list attached).

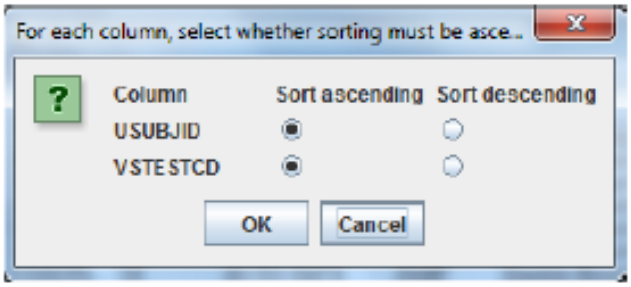
15. Sorting and Filtering on Topic Variable

This brings us to a second possibility for sorting and filtering. Most of the "Findings" tables are hyper vertical tables according to the "Entity - Attribute - Value" (EAV) model. This means that there is a subject ID, a test code, and a test result. We can now sort on USUBJID and VSTESTCD to ensure that we see e.g. all diastolic blood pressures of each single subject together[8]. So we can use "Tools – Sorting" and then add "USUBJID" and "VSTESTCD" to the list on which sorting must be applied:

[8] In many cases, the tables are already organized per subject per test code, but this is not a requirement.



We are asked whether we should sort ascending or descending for both variables:

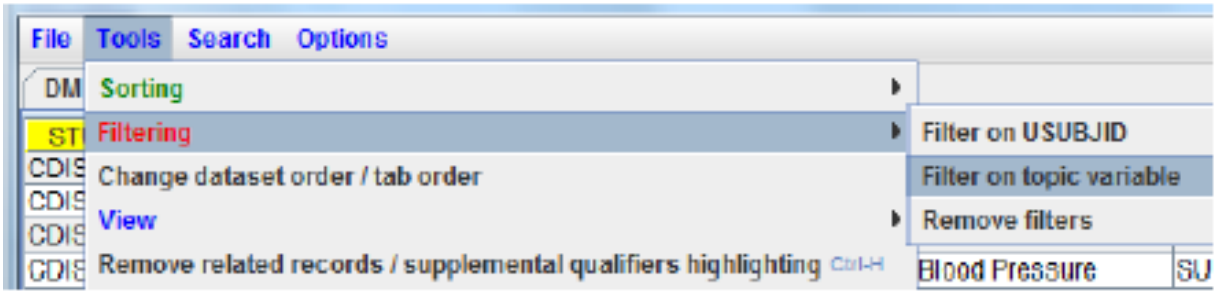


If we choose for "ascending" sorting for both variables, the result is:

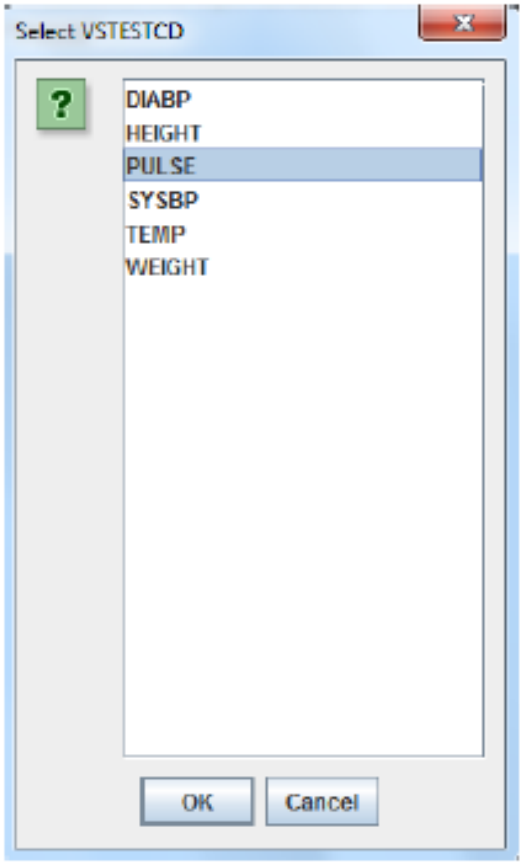
DM | VS

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU
CDISCPIL...	VS	01-701-1047	1	DIABP	Diastolic Bl...	SUPINE	68	mmHg
CDISCPIL...	VS	01-701-1047	2	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	3	DIABP	Diastolic Bl...	STANDING	83	mmHg
CDISCPIL...	VS	01-701-1047	4	DIABP	Diastolic Bl...	SUPINE	65	mmHg
CDISCPIL...	VS	01-701-1047	5	DIABP	Diastolic Bl...	STANDING	86	mmHg
CDISCPIL...	VS	01-701-1047	6	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	7	DIABP	Diastolic Bl...	SUPINE	61	mmHg
CDISCPIL...	VS	01-701-1047	8	DIABP	Diastolic Bl...	STANDING	68	mmHg
CDISCPIL...	VS	01-701-1047	9	DIABP	Diastolic Bl...	STANDING	78	mmHg
CDISCPIL...	VS	01-701-1047	10	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	11	DIABP	Diastolic Bl...	STANDING	56	mmHg
CDISCPIL...	VS	01-701-1047	12	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1047	13	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	14	DIABP	Diastolic Bl...	STANDING	84	mmHg
CDISCPIL...	VS	01-701-1047	15	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1047	16	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1047	17	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	18	DIABP	Diastolic Bl...	STANDING	78	mmHg
CDISCPIL...	VS	01-701-1047	19	DIABP	Diastolic Bl...	SUPINE	71	mmHg
CDISCPIL...	VS	01-701-1047	20	DIABP	Diastolic Bl...	STANDING	63	mmHg
CDISCPIL...	VS	01-701-1047	21	DIABP	Diastolic Bl...	STANDING	57	mmHg
CDISCPIL...	VS	01-701-1047	22	DIABP	Diastolic Bl...	SUPINE	76	mmHg
CDISCPIL...	VS	01-701-1047	23	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1047	24	DIABP	Diastolic Bl...	STANDING	65	mmHg
CDISCPIL...	VS	01-701-1047	25	HEIGHT	Height		58.5	IN
CDISCPIL...	VS	01-701-1047	26	PULSE	Pulse Rate	SUPINE	53	BEATS/MIN
CDISCPIL...	VS	01-701-1047	27	PULSE	Pulse Rate	STANDING	64	BEATS/MIN
CDISCPIL...	VS	01-701-1047	28	PULSE	Pulse Rate	STANDING	61	BEATS/MIN

Suppose we are only interested in the pulse rate for now (VSTESTCD=PULSE). We can then filter out all other test results, by using the menu "Tools – Filter on topic variable":



A list of all values of the topic variable[9] (in this case VSTESTCD) is then displayed, from which we choose "PULSE":



Remark that multiple selection is possible using the CTRL-key.

After clicking "OK", only the records for which VSTESTCD=PULSE are displayed, allowing us to closely inspect the pulse rates of all subjects of 81 years and older:

[9] In the case of a SUPP-- dataset, the topic variable is QNAM.

DM VS

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU
CDISCPIL...	VS	01-701-1047	26	PULSE	Pulse Rate	SUPINE	53	BEATS/MIN
CDISCPIL...	VS	01-701-1047	27	PULSE	Pulse Rate	STANDING	64	BEATS/MIN
CDISCPIL...	VS	01-701-1047	28	PULSE	Pulse Rate	STANDING	61	BEATS/MIN
CDISCPIL...	VS	01-701-1047	29	PULSE	Pulse Rate	SUPINE	58	BEATS/MIN
CDISCPIL...	VS	01-701-1047	30	PULSE	Pulse Rate	STANDING	74	BEATS/MIN
CDISCPIL...	VS	01-701-1047	31	PULSE	Pulse Rate	STANDING	73	BEATS/MIN
CDISCPIL...	VS	01-701-1047	32	PULSE	Pulse Rate	SUPINE	58	BEATS/MIN
CDISCPIL...	VS	01-701-1047	33	PULSE	Pulse Rate	STANDING	67	BEATS/MIN
CDISCPIL...	VS	01-701-1047	34	PULSE	Pulse Rate	STANDING	67	BEATS/MIN
CDISCPIL...	VS	01-701-1047	35	PULSE	Pulse Rate	SUPINE	62	BEATS/MIN
CDISCPIL...	VS	01-701-1047	36	PULSE	Pulse Rate	STANDING	74	BEATS/MIN
CDISCPIL...	VS	01-701-1047	37	PULSE	Pulse Rate	STANDING	72	BEATS/MIN
CDISCPIL...	VS	01-701-1047	38	PULSE	Pulse Rate	SUPINE	66	BEATS/MIN
CDISCPIL...	VS	01-701-1047	39	PULSE	Pulse Rate	STANDING	75	BEATS/MIN
CDISCPIL...	VS	01-701-1047	40	PULSE	Pulse Rate	STANDING	78	BEATS/MIN
CDISCPIL...	VS	01-701-1047	41	PULSE	Pulse Rate	SUPINE	69	BEATS/MIN
CDISCPIL...	VS	01-701-1047	42	PULSE	Pulse Rate	STANDING	64	BEATS/MIN
CDISCPIL...	VS	01-701-1047	43	PULSE	Pulse Rate	STANDING	70	BEATS/MIN
CDISCPIL...	VS	01-701-1047	44	PULSE	Pulse Rate	SUPINE	78	BEATS/MIN
CDISCPIL...	VS	01-701-1047	45	PULSE	Pulse Rate	STANDING	90	BEATS/MIN
CDISCPIL...	VS	01-701-1047	46	PULSE	Pulse Rate	STANDING	87	BEATS/MIN
CDISCPIL...	VS	01-701-1047	47	PULSE	Pulse Rate	SUPINE	58	BEATS/MIN
CDISCPIL...	VS	01-701-1047	48	PULSE	Pulse Rate	STANDING	72	BEATS/MIN
CDISCPIL...	VS	01-701-1047	49	PULSE	Pulse Rate	STANDING	72	BEATS/MIN
CDISCPIL...	VS	01-701-1111	17	PULSE	Pulse Rate	SUPINE	80	BEATS/MIN
CDISCPIL...	VS	01-701-1111	18	PULSE	Pulse Rate	STANDING	85	BEATS/MIN
CDISCPIL...	VS	01-701-1111	19	PULSE	Pulse Rate	STANDING	84	BEATS/MIN

Dragging the "VISITDY" column to the left allows a better interpretation of the data. E.g.:

DM VS

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VISITDY	VSORRESU
CDISCPIL...	VS	01-701-1047	26	PULSE	Pulse Rate	SUPINE	53	-7	BEATS/MIN
CDISCPIL...	VS	01-701-1047	27	PULSE	Pulse Rate	STANDING	64	-7	BEATS/MIN
CDISCPIL...	VS	01-701-1047	28	PULSE	Pulse Rate	STANDING	61	-7	BEATS/MIN
CDISCPIL...	VS	01-701-1047	29	PULSE	Pulse Rate	SUPINE	58	-1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	30	PULSE	Pulse Rate	STANDING	74	-1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	31	PULSE	Pulse Rate	STANDING	73	-1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	32	PULSE	Pulse Rate	SUPINE	58	1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	33	PULSE	Pulse Rate	STANDING	67	1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	34	PULSE	Pulse Rate	STANDING	67	1	BEATS/MIN
CDISCPIL...	VS	01-701-1047	35	PULSE	Pulse Rate	SUPINE	62	13	BEATS/MIN
CDISCPIL...	VS	01-701-1047	36	PULSE	Pulse Rate	STANDING	74	13	BEATS/MIN
CDISCPIL...	VS	01-701-1047	37	PULSE	Pulse Rate	STANDING	72	13	BEATS/MIN
CDISCPIL...	VS	01-701-1047	38	PULSE	Pulse Rate	SUPINE	66	14	BEATS/MIN
CDISCPIL...	VS	01-701-1047	39	PULSE	Pulse Rate	STANDING	75	14	BEATS/MIN
CDISCPIL...	VS	01-701-1047	40	PULSE	Pulse Rate	STANDING	78	14	BEATS/MIN
CDISCPIL...	VS	01-701-1047	41	PULSE	Pulse Rate	SUPINE	69	28	BEATS/MIN
CDISCPIL...	VS	01-701-1047	42	PULSE	Pulse Rate	STANDING	64	29	BEATS/MIN
CDISCPIL...	VS	01-701-1047	43	PULSE	Pulse Rate	STANDING	70	28	BEATS/MIN
CDISCPIL...	VS	01-701-1047	44	PULSE	Pulse Rate	SUPINE	78	30	BEATS/MIN
CDISCPIL...	VS	01-701-1047	45	PULSE	Pulse Rate	STANDING	90	30	BEATS/MIN
CDISCPIL...	VS	01-701-1047	46	PULSE	Pulse Rate	STANDING	87	30	BEATS/MIN
CDISCPIL...	VS	01-701-1047	47	PULSE	Pulse Rate	SUPINE	58	168	BEATS/MIN
CDISCPIL...	VS	01-701-1047	48	PULSE	Pulse Rate	STANDING	72	168	BEATS/MIN
CDISCPIL...	VS	01-701-1047	49	PULSE	Pulse Rate	STANDING	72	168	BEATS/MIN
CDISCPIL...	VS	01-701-1111	17	PULSE	Pulse Rate	SUPINE	80	-7	BEATS/MIN
CDISCPIL...	VS	01-701-1111	18	PULSE	Pulse Rate	STANDING	85	-7	BEATS/MIN
CDISCPIL...	VS	01-701-1111	19	PULSE	Pulse Rate	STANDING	84	-7	BEATS/MIN

If the data is not already sorted on VISITDY (within each subject) we can again use "Tools - Sorting" and sort on USUBJID as primary variable and VISITDY as secondary variable. Of course, we additionally can do all kind of sorting, e.g. on VSPOS.

To remove the filter that we applied (using VSTESTCD=PULSE) use the menu "Tools – Filtering – Remove filters". The following dialogue is displayed:

If we select "Yes", the un-filtering will be applied on all currently loaded datasets. If we select "No", only the currently displayed table will be un-filtered. In this case, it does not matter, as the filter was on the topic variable "VSTESTCD" and this variable does not occur in the DM table that was also loaded. The result is:

DM | VS

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES
CDISCPILOT01	VS	01-701-1047	1	DIABP	Diastolic Bl...	SUPINE	68
CDISCPILOT01	VS	01-701-1047	2	DIABP	Diastolic Bl...	STANDING	71
CDISCPILOT01	VS	01-701-1047	3	DIABP	Diastolic Bl...	STANDING	83
CDISCPILOT01	VS	01-701-1047	4	DIABP	Diastolic Bl...	SUPINE	65
CDISCPILOT01	VS	01-701-1047	5	DIABP	Diastolic Bl...	STANDING	86
CDISCPILOT01	VS	01-701-1047	6	DIABP	Diastolic Bl...	STANDING	71
CDISCPILOT01	VS	01-701-1047	7	DIABP	Diastolic Bl...	SUPINE	61
CDISCPILOT01	VS	01-701-1047	8	DIABP	Diastolic Bl...	STANDING	68
CDISCPILOT01	VS	01-701-1047	9	DIABP	Diastolic Bl...	STANDING	78
CDISCPILOT01	VS	01-701-1047	10	DIABP	Diastolic Bl...	SUPINE	72
CDISCPILOT01	VS	01-701-1047	11	DIABP	Diastolic Bl...	STANDING	56
CDISCPILOT01	VS	01-701-1047	12	DIABP	Diastolic Bl...	STANDING	73
CDISCPILOT01	VS	01-701-1047	13	DIABP	Diastolic Bl...	SUPINE	72
CDISCPILOT01	VS	01-701-1047	14	DIABP	Diastolic Bl...	STANDING	84
CDISCPILOT01	VS	01-701-1047	15	DIABP	Diastolic Bl...	STANDING	73
CDISCPILOT01	VS	01-701-1047	16	DIABP	Diastolic Bl...	SUPINE	72
CDISCPILOT01	VS	01-701-1047	17	DIABP	Diastolic Bl...	STANDING	71
CDISCPILOT01	VS	01-701-1047	18	DIABP	Diastolic Bl...	STANDING	78
CDISCPILOT01	VS	01-701-1047	19	DIABP	Diastolic Bl...	SUPINE	71
CDISCPILOT01	VS	01-701-1047	20	DIABP	Diastolic Bl...	STANDING	63
CDISCPILOT01	VS	01-701-1047	21	DIABP	Diastolic Bl...	STANDING	57
CDISCPILOT01	VS	01-701-1047	22	DIABP	Diastolic Bl...	SUPINE	76
CDISCPILOT01	VS	01-701-1047	23	DIABP	Diastolic Bl...	STANDING	71
CDISCPILOT01	VS	01-701-1047	24	DIABP	Diastolic Bl...	STANDING	65
CDISCPILOT01	VS	01-701-1047	25	HEIGHT	Height		58.5
CDISCPILOT01	VS	01-701-1047	26	PULSE	Pulse Rate	SUPINE	53
CDISCPILOT01	VS	01-701-1047	27	PULSE	Pulse Rate	STANDING	64
CDISCPILOT01	VS	01-701-1047	28	PULSE	Pulse Rate	STANDING	61
CDISCPILOT01	VS	01-701-1047	29	PULSE	Pulse Rate	SUPINE	58

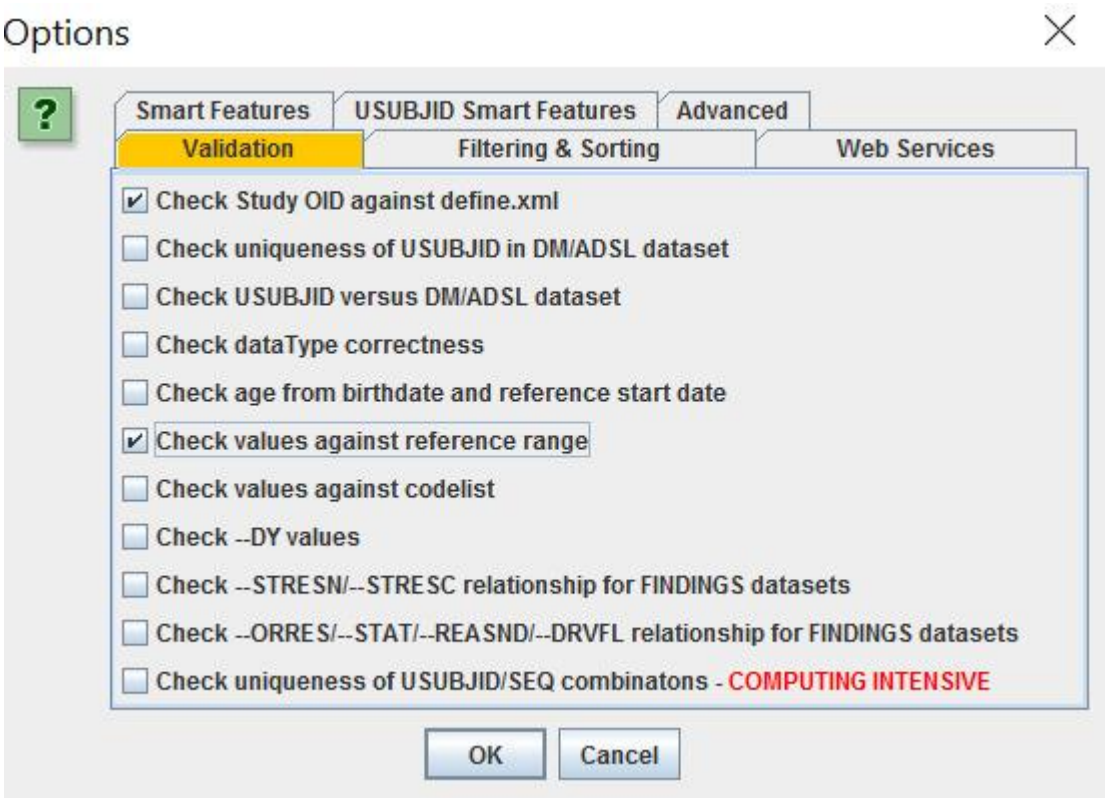
Remark that this table is still only containing the information for the subjects of 81 years and older, as the filter on age was applied during loading of the datasets. In order to see all subjects again, we will have to reload the tables.

16. Advanced Filtering

Suppose that we want to inspect all the laboratory data of a few of the subjects of 81 years and older, e.g. those senior subjects that have at least one abnormal haemoglobin value.

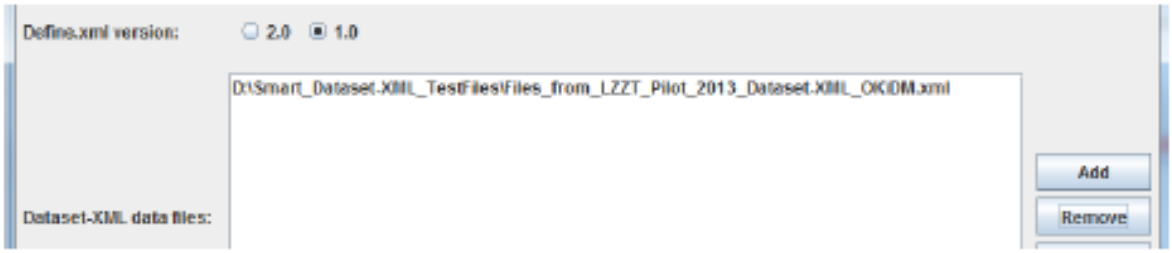
The best strategy for this is to first select the subjects of interest (e.g. based on age), apply the filter, and then load the laboratory data for these subjects.

Before starting loading, we click the "Options" button, and check the option "Check values against reference range":



As we want to see all laboratory data for our senior subjects, we can already switch off "Allow filters on --TESTCD / --CAT". It is however usually better to do this on the dataset level. So we leave the box checked for the moment.

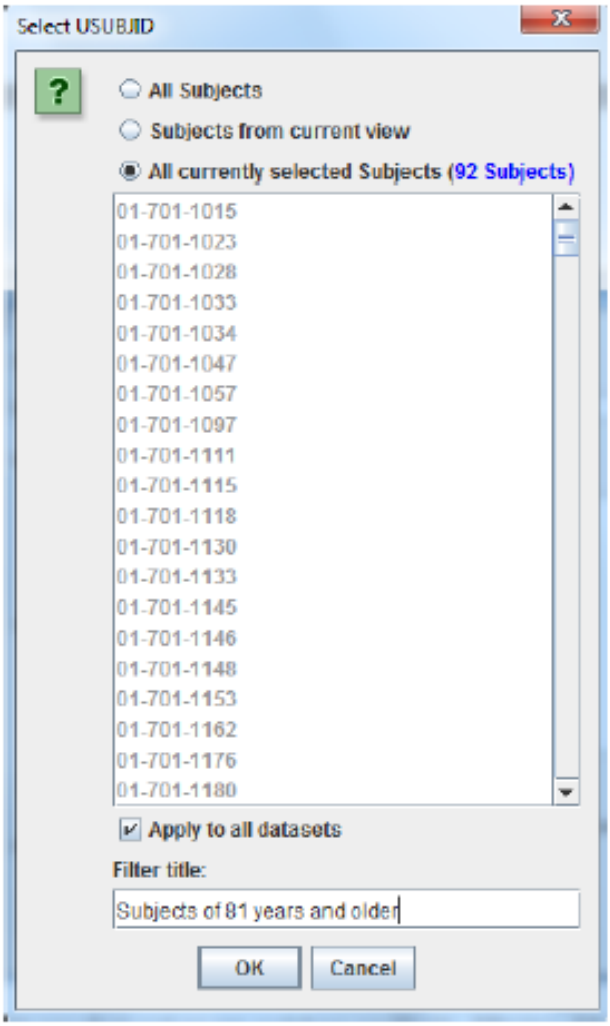
We now only set DM.xml file to be loaded:



And then start loading:

As the define.xml file contains controlled terminology for LBTESTCD and IETESTCD, the filtering dialogue is presented (later we will learn how to switch it off), but we can just go with the option "No filter on LBTESTCD" for LB on and load the DM dataset.

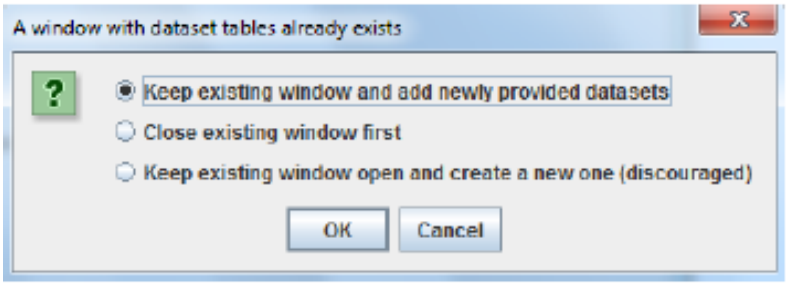
We now first sort on age in the DM table and then use "Tools – Filtering – Filter on USUBJID" to filter on subjects of 81 years and older:



We also add a title in "Filter title" where we add a short description of about what our filter is doing. We type in "Subjects of 81 years and older", and then click "OK". This again results in a DM table only containing the subjects that are 81 years and older. Remark that in this case all other subjects are kept in memory, so that un-filtering will lead to a view containing all subjects again, at least when we use "Yes, remove on all datasets" when un-filtering.

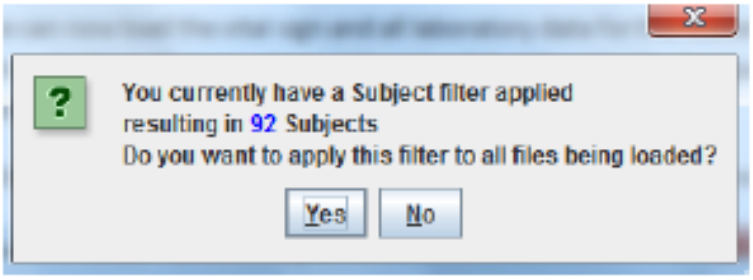
In a second step, we can now load the vital sign and all laboratory data for this subset of subjects (subjects of 81 years and older). First we add the VS.xml and the LB.xml to the list (just leave the window with the DM table open for the moment).

Clicking the "Start" button results in the following dialog, as we still have a window with tables open:

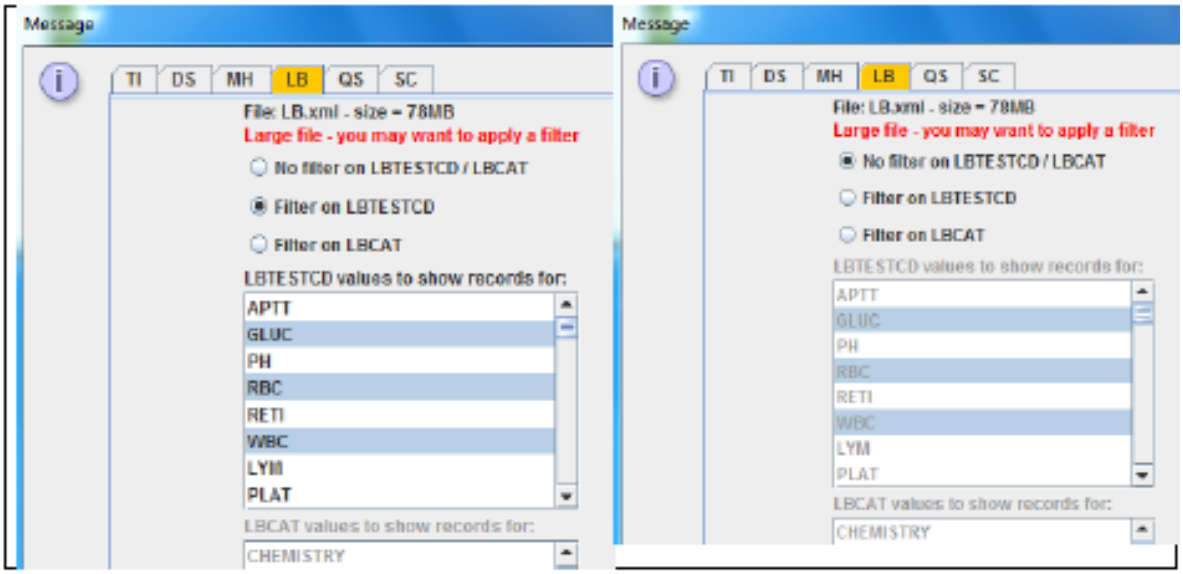


In most cases, the best choice will be "keep existing window and add newly provided datasets". Keeping several windows open at the same time is often not such a good idea, due to memory consumption.

We are now asked whether we want to apply our subject filter:



And if we click on "Yes":



We might still have had the filter for LBTESTCD applied (left image), but as we now want to see all lab results for our senior subjects, we select "No filter on LBTESTCD" (right image)

Our LB table then looks like:

DM VS LB

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU
CDISCPIL...	LB	01-701-1047	1	ALB	Albumin	CHEMISTRY	4.2	g/dL
CDISCPIL...	LB	01-701-1047	38	ALB	Albumin	CHEMISTRY	3.8	g/dL
CDISCPIL...	LB	01-701-1047	73	ALB	Albumin	CHEMISTRY	3.8	g/dL
CDISCPIL...	LB	01-701-1047	120	ALB	Albumin	CHEMISTRY	3.8	g/dL
CDISCPIL...	LB	01-701-1047	2	ALP	Alkaline Ph...	CHEMISTRY	78	U/L
CDISCPIL...	LB	01-701-1047	39	ALP	Alkaline Ph...	CHEMISTRY	67	U/L
CDISCPIL...	LB	01-701-1047	74	ALP	Alkaline Ph...	CHEMISTRY	71	U/L
CDISCPIL...	LB	01-701-1047	121	ALP	Alkaline Ph...	CHEMISTRY	75	U/L
CDISCPIL...	LB	01-701-1047	3	ALT	Alanine Am...	CHEMISTRY	22	U/L
CDISCPIL...	LB	01-701-1047	40	ALT	Alanine Am...	CHEMISTRY	16	U/L
CDISCPIL...	LB	01-701-1047	75	ALT	Alanine Am...	CHEMISTRY	20	U/L
CDISCPIL...	LB	01-701-1047	122	ALT	Alanine Am...	CHEMISTRY	17	U/L
CDISCPIL...	LB	01-701-1047	4	AST	Aspartate A...	CHEMISTRY	25	U/L
CDISCPIL...	LB	01-701-1047	41	AST	Aspartate A...	CHEMISTRY	21	U/L
CDISCPIL...	LB	01-701-1047	76	AST	Aspartate A...	CHEMISTRY	24	U/L
CDISCPIL...	LB	01-701-1047	123	AST	Aspartate A...	CHEMISTRY	21	U/L
CDISCPIL...	LB	01-701-1047	5	BASO	Basophils	HEMATOL...	0.05	THOU/uL
CDISCPIL...	LB	01-701-1047	42	BASO	Basophils	HEMATOL...	0.05	THOU/uL
CDISCPIL...	LB	01-701-1047	77	BASO	Basophils	HEMATOL...	0.04	THOU/uL
CDISCPIL...	LB	01-701-1047	103	BASO	Basophils	HEMATOL...	0.03	THOU/uL
CDISCPIL...	LB	01-701-1047	6	BILI	Bilirubin	CHEMISTRY	0.4	mg/dL
CDISCPIL...	LB	01-701-1047	43	BILI	Bilirubin	CHEMISTRY	0.5	mg/dL

Remark that the filter title "Subjects of 81 years and older" is displayed at the top:

If we now scroll down, we will find a few lab values that are outside the reference range:

CDISCPIL...	LB	01-701-1130	242	GLUC	Glucose	CHEMISTRY	88	mg/dL	50	250	4.88488	4.88488	mmol/L	2.8	13.9
CDISCPIL...	LB	01-701-1130	273	GLUC	Glucose	CHEMISTRY	88	mg/dL	50	250	4.88488	4.88488	mmol/L	2.8	13.9
CDISCPIL...	LB	01-701-1130	307	GLUC	Glucose	CHEMISTRY	94	mg/dL	50	250	5.21794	5.21794	mmol/L	2.8	13.9
CDISCPIL...	LB	01-701-1130	17	HCT	Hematocrit	HEMATOL...	36.0	%	37	51	0.36	0.36	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	54	HCT	Hematocrit	HEMATOL...	37.0	%	37	51	0.37	0.37	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	88	HCT	Hematocrit	HEMATOL...	31.0	%	37	51	0.31	0.31	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	118	HCT	Hematocrit	HEMATOL...	35.0	%	37	51	0.35	0.35	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	148	HCT	Hematocrit	HEMATOL...	33.0	%	37	51	0.33	0.33	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	179	HCT	Hematocrit	HEMATOL...	33.0	%	37	51	0.33	0.33	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	213	HCT	Hematocrit	HEMATOL...	38.0	%	37	51	0.38	0.38	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	243	HCT	Hematocrit	HEMATOL...	33.0	%	37	51	0.33	0.33	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	274	HCT	Hematocrit	HEMATOL...	37.0	%	37	51	0.37	0.37	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	308	HCT	Hematocrit	HEMATOL...	34.0	%	37	51	0.34	0.34	1	0.37	0.51
CDISCPIL...	LB	01-701-1130	18	HGB	Hemoglobin	HEMATOL...	12.1	g/dL	12.5	17	7.50926	7.50926	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	55	HGB	Hemoglobin	HEMATOL...	12.6	g/dL	12.5	17	7.81956	7.81956	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	89	HGB	Hemoglobin	HEMATOL...	10.5	g/dL	12.5	17	6.5163	6.5163	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	119	HGB	Hemoglobin	HEMATOL...	11.8	g/dL	12.5	17	7.32308	7.32308	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	149	HGB	Hemoglobin	HEMATOL...	11.2	g/dL	12.5	17	6.95072	6.95072	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	180	HGB	Hemoglobin	HEMATOL...	11.5	g/dL	12.5	17	7.1369	7.1369	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	214	HGB	Hemoglobin	HEMATOL...	11.9	g/dL	12.5	17	7.38514	7.38514	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	244	HGB	Hemoglobin	HEMATOL...	11.1	g/dL	12.5	17	6.88866	6.88866	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	275	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	309	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L	7.76	10.55
CDISCPIL...	LB	01-701-1130	19	K	Potassium	CHEMISTRY	4.2	mEq/L	3.4	5.4	4.2	4.2	mmol/L	3.4	5.4

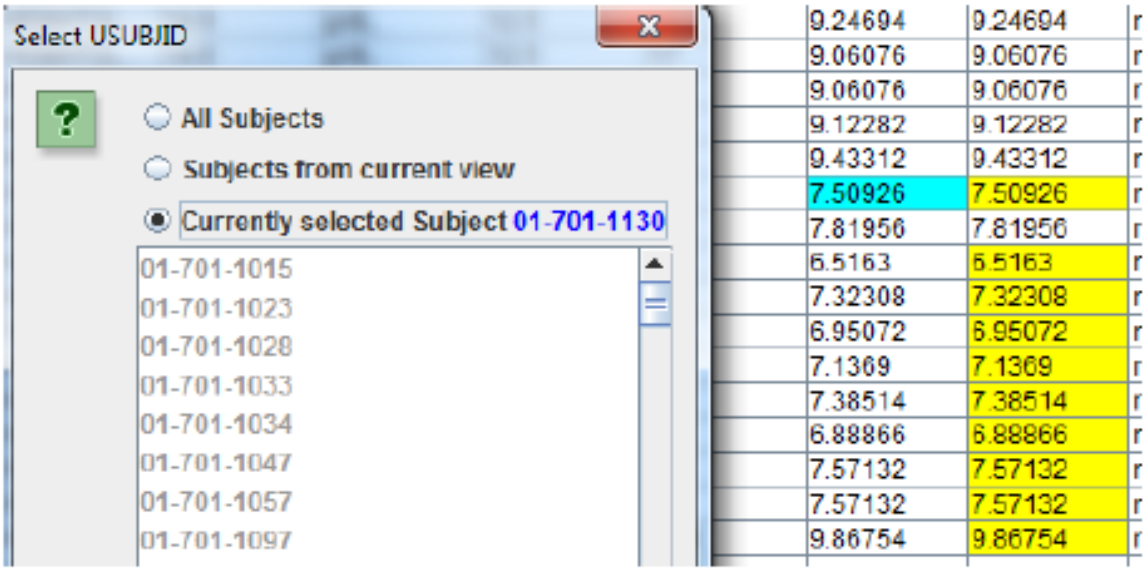
And we would like to concentrate on those subjects (with age > 80) that showed an abnormal Haemoglobin value. To do so, we apply a "topic variable" filtering on LBTESTCD=HGB using "Tools – Filtering – Filter on topic variable" and then pick "HGB", resulting in:

DM | VS | LB

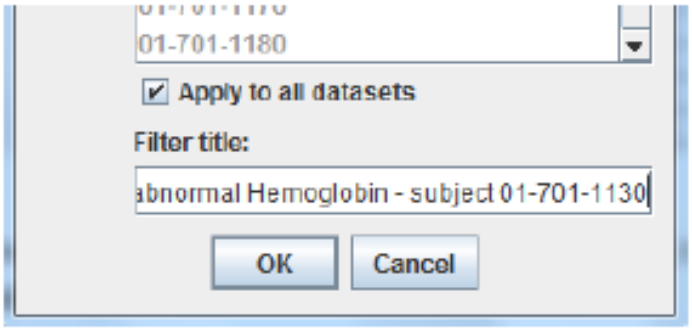
STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI	LBSTRESC	LBSTRESN	LBSTRESU
CDISCPIL...	LB	01-701-1047	18	HGB	Hemoglobin	HEMATOL...	15.3	g/dL	11.5	15.8	9.49518	9.49518	mmol/L
CDISCPIL...	LB	01-701-1047	55	HGB	Hemoglobin	HEMATOL...	14.0	g/dL	11.5	15.8	8.6884	8.6884	mmol/L
CDISCPIL...	LB	01-701-1047	89	HGB	Hemoglobin	HEMATOL...	13.5	g/dL	11.5	15.8	8.3781	8.3781	mmol/L
CDISCPIL...	LB	01-701-1047	107	HGB	Hemoglobin	HEMATOL...	13.3	g/dL	11.5	15.8	8.25398	8.25398	mmol/L
CDISCPIL...	LB	01-701-1111	18	HGB	Hemoglobin	HEMATOL...	12.0	g/dL	11.5	15.8	7.4472	7.4472	mmol/L
CDISCPIL...	LB	01-701-1111	56	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	11.5	15.8	7.57132	7.57132	mmol/L
CDISCPIL...	LB	01-701-1115	18	HGB	Hemoglobin	HEMATOL...	14.9	g/dL	12.5	17	9.24694	9.24694	mmol/L
CDISCPIL...	LB	01-701-1115	55	HGB	Hemoglobin	HEMATOL...	14.6	g/dL	12.5	17	9.06076	9.06076	mmol/L
CDISCPIL...	LB	01-701-1115	89	HGB	Hemoglobin	HEMATOL...	14.6	g/dL	12.5	17	9.06076	9.06076	mmol/L
CDISCPIL...	LB	01-701-1115	134	HGB	Hemoglobin	HEMATOL...	14.7	g/dL	12.5	17	9.12282	9.12282	mmol/L
CDISCPIL...	LB	01-701-1115	165	HGB	Hemoglobin	HEMATOL...	15.2	g/dL	12.5	17	9.43312	9.43312	mmol/L
CDISCPIL...	LB	01-701-1130	18	HGB	Hemoglobin	HEMATOL...	12.1	g/dL	12.5	17	7.50926	7.50926	mmol/L
CDISCPIL...	LB	01-701-1130	55	HGB	Hemoglobin	HEMATOL...	12.6	g/dL	12.5	17	7.81956	7.81956	mmol/L
CDISCPIL...	LB	01-701-1130	89	HGB	Hemoglobin	HEMATOL...	10.5	g/dL	12.5	17	6.5163	6.5163	mmol/L
CDISCPIL...	LB	01-701-1130	119	HGB	Hemoglobin	HEMATOL...	11.8	g/dL	12.5	17	7.32308	7.32308	mmol/L
CDISCPIL...	LB	01-701-1130	149	HGB	Hemoglobin	HEMATOL...	11.2	g/dL	12.5	17	6.95072	6.95072	mmol/L
CDISCPIL...	LB	01-701-1130	180	HGB	Hemoglobin	HEMATOL...	11.5	g/dL	12.5	17	7.1369	7.1369	mmol/L
CDISCPIL...	LB	01-701-1130	214	HGB	Hemoglobin	HEMATOL...	11.9	g/dL	12.5	17	7.38514	7.38514	mmol/L
CDISCPIL...	LB	01-701-1130	244	HGB	Hemoglobin	HEMATOL...	11.1	g/dL	12.5	17	6.88866	6.88866	mmol/L
CDISCPIL...	LB	01-701-1130	275	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L
CDISCPIL...	LB	01-701-1130	309	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L
CDISCPIL...	LB	01-701-1133	18	HGB	Hemoglobin	HEMATOL...	15.9	g/dL	11.5	15.8	9.86754	9.86754	mmol/L
CDISCPIL...	LB	01-701-1133	42	HGB	Hemoglobin	HEMATOL...	15.4	g/dL	11.5	15.8	9.55724	9.55724	mmol/L
CDISCPIL...	LB	01-701-1133	71	HGB	Hemoglobin	HEMATOL...	15.3	g/dL	11.5	15.8	9.49518	9.49518	mmol/L
CDISCPIL...	LB	01-701-1133	101	HGB	Hemoglobin	HEMATOL...	15.0	g/dL	11.5	15.8	9.309	9.309	mmol/L
CDISCPIL...	LB	01-701-1133	131	HGB	Hemoglobin	HEMATOL...	15.2	g/dL	11.5	15.8	9.43312	9.43312	mmol/L
CDISCPIL...	LB	01-701-1133	162	HGB	Hemoglobin	HEMATOL...	15.2	g/dL	11.5	15.8	9.43312	9.43312	mmol/L
CDISCPIL...	LB	01-701-1133	196	HGB	Hemoglobin	HEMATOL...	15.1	g/dL	11.5	15.8	9.37106	9.37106	mmol/L
CDISCPIL...	LB	01-701-1133	226	HGB	Hemoglobin	HEMATOL...	15.1	g/dL	11.5	15.8	9.37106	9.37106	mmol/L

We see that subject "01-701-1130" has a good number of "out of range" haemoglobin values, on which we would like to concentrate.

We can now either isolate the data for subject "01-701-1130" by either using subject filtering and only pick "01-701-1130", or by first selecting any cell for subject "01-701-1130" and then choose "Tools – Filtering – Filter on USUBJID" and the check "Currently selected subject".



When we then also check the box "Apply to all datasets", then we will also only see the data of this single subject in the DM table and in the VS table.



Remark that we can of course also select 2 or 3 or more subjects that show "out of range" Haemoglobin values.

In case we isolate the data of subject "01-701-1130", the DM table becomes:

Remark the title at the top. The VS table becomes:

DM VS LB

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU
CDISCPIL...	VS	01-701-1130	1	DIABP	Diastolic Bl...	SUPINE	68	mmHg
CDISCPIL...	VS	01-701-1130	2	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1130	3	DIABP	Diastolic Bl...	STANDING	70	mmHg
CDISCPIL...	VS	01-701-1130	4	DIABP	Diastolic Bl...	SUPINE	71	mmHg
CDISCPIL...	VS	01-701-1130	5	DIABP	Diastolic Bl...	STANDING	64	mmHg
CDISCPIL...	VS	01-701-1130	6	DIABP	Diastolic Bl...	STANDING	76	mmHg
CDISCPIL...	VS	01-701-1130	7	DIABP	Diastolic Bl...	SUPINE	74	mmHg
CDISCPIL...	VS	01-701-1130	8	DIABP	Diastolic Bl...	STANDING	67	mmHg
CDISCPIL...	VS	01-701-1130	9	DIABP	Diastolic Bl...	STANDING	68	mmHg
CDISCPIL...	VS	01-701-1130	10	DIABP	Diastolic Bl...	SUPINE	75	mmHg
CDISCPIL...	VS	01-701-1130	11	DIABP	Diastolic Bl...	STANDING	72	mmHg
CDISCPIL...	VS	01-701-1130	12	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1130	13	DIABP	Diastolic Bl...	SUPINE	72	mmHg
CDISCPIL...	VS	01-701-1130	14	DIABP	Diastolic Bl...	STANDING	69	mmHg
CDISCPIL...	VS	01-701-1130	15	DIABP	Diastolic Bl...	STANDING	65	mmHg
CDISCPIL...	VS	01-701-1130	16	DIABP	Diastolic Bl...	SUPINE	74	mmHg
CDISCPIL...	VS	01-701-1130	17	DIABP	Diastolic Bl...	STANDING	80	mmHg
CDISCPIL...	VS	01-701-1130	18	DIABP	Diastolic Bl...	STANDING	72	mmHg
CDISCPIL...	VS	01-701-1130	19	DIABP	Diastolic Bl...	SUPINE	68	mmHg
CDISCPIL...	VS	01-701-1130	20	DIABP	Diastolic Bl...	STANDING	68	mmHg
CDISCPIL...	VS	01-701-1130	21	DIABP	Diastolic Bl...	STANDING	75	mmHg
CDISCPIL...	VS	01-701-1130	22	DIABP	Diastolic Bl...	SUPINE	64	mmHg
CDISCPIL...	VS	01-701-1130	23	DIABP	Diastolic Bl...	STANDING	70	mmHg
CDISCPIL...	VS	01-701-1130	24	DIABP	Diastolic Bl...	STANDING	71	mmHg
CDISCPIL...	VS	01-701-1130	25	DIABP	Diastolic Bl...	SUPINE	75	mmHg
CDISCPIL...	VS	01-701-1130	26	DIABP	Diastolic Bl...	STANDING	68	mmHg
CDISCPIL...	VS	01-701-1130	27	DIABP	Diastolic Bl...	STANDING	76	mmHg
CDISCPIL...	VS	01-701-1130	28	DIABP	Diastolic Bl...	SUPINE	67	mmHg
CDISCPIL...	VS	01-701-1130	29	DIABP	Diastolic Bl...	STANDING	60	mmHg
CDISCPIL...	VS	01-701-1130	30	DIABP	Diastolic Bl...	STANDING	72	mmHg
CDISCPIL...	VS	01-701-1130	31	DIABP	Diastolic Bl...	SUPINE	76	mmHg
CDISCPIL...	VS	01-701-1130	32	DIABP	Diastolic Bl...	STANDING	73	mmHg
CDISCPIL...	VS	01-701-1130	33	DIABP	Diastolic Bl...	STANDING	75	mmHg
CDISCPIL...	VS	01-701-1130	34	DIABP	Diastolic Bl...	SUPINE	63	mmHg
CDISCPIL...	VS	01-701-1130	35	DIABP	Diastolic Bl...	STANDING	64	mmHg
CDISCPIL...	VS	01-701-1130	36	DIABP	Diastolic Bl...	STANDING	56	mmHg
CDISCPIL...	VS	01-701-1130	37	DIABP	Diastolic Bl...	SUPINE	75	mmHg
CDISCPIL...	VS	01-701-1130	38	DIABP	Diastolic Bl...	STANDING	76	mmHg
CDISCPIL...	VS	01-701-1130	39	DIABP	Diastolic Bl...	STANDING	80	mmHg
CDISCPIL...	VS	01-701-1130	40	DIABP	Diastolic Bl...	SUPINE	70	mmHg
CDISCPIL...	VS	01-701-1130	41	DIABP	Diastolic Bl...	STANDING	91	mmHg
CDISCPIL...	VS	01-701-1130	42	DIABP	Diastolic Bl...	STANDING	84	mmHg
CDISCPIL...	VS	01-701-1130	43	HEIGHT	Height		66.8	IN
CDISCPIL...	VS	01-701-1130	44	PULSE	Pulse Rate	SUPINE	71	BEATS/MIN

Also here, we could apply an additional filter, e.g. on "PULSE", but usually just sorting by VSTESTCD and e.g. VISITDY will also do the job.

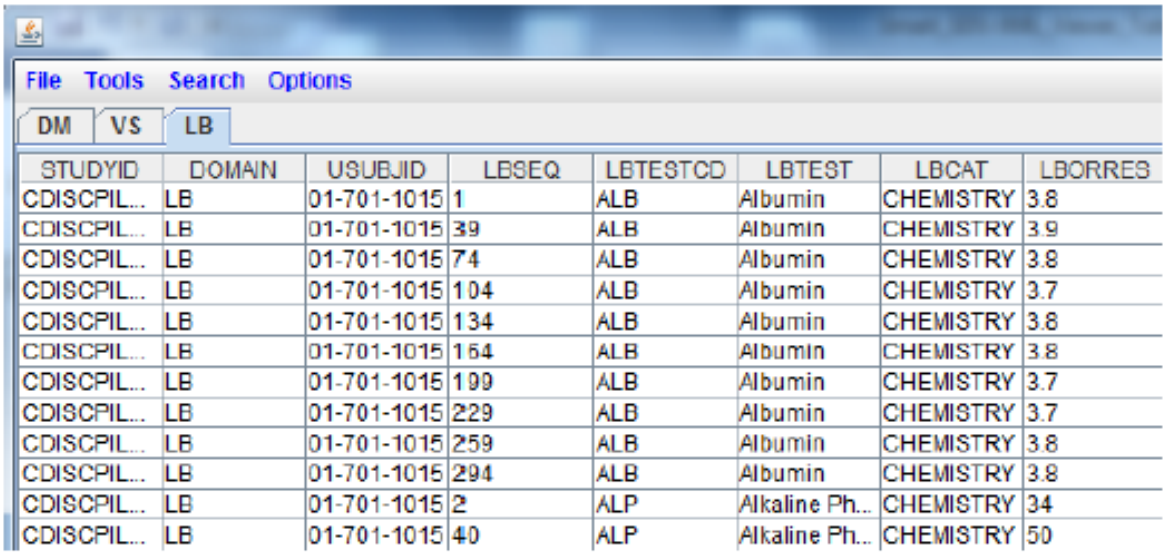
The LB table view now is:

DM VS LB

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI	LBSTRESC	LBSTRESN	LBSTRESU
CDISCPIL...	LB	01-701-1130	18	HGB	Hemoglobin	HEMATOL...	12.1	g/dL	12.5	17	7.50926	7.50926	mmol/L
CDISCPIL...	LB	01-701-1130	55	HGB	Hemoglobin	HEMATOL...	12.6	g/dL	12.5	17	7.81956	7.81956	mmol/L
CDISCPIL...	LB	01-701-1130	89	HGB	Hemoglobin	HEMATOL...	10.5	g/dL	12.5	17	6.5163	6.5163	mmol/L
CDISCPIL...	LB	01-701-1130	119	HGB	Hemoglobin	HEMATOL...	11.8	g/dL	12.5	17	7.32308	7.32308	mmol/L
CDISCPIL...	LB	01-701-1130	149	HGB	Hemoglobin	HEMATOL...	11.2	g/dL	12.5	17	6.95072	6.95072	mmol/L
CDISCPIL...	LB	01-701-1130	180	HGB	Hemoglobin	HEMATOL...	11.5	g/dL	12.5	17	7.1369	7.1369	mmol/L
CDISCPIL...	LB	01-701-1130	214	HGB	Hemoglobin	HEMATOL...	11.9	g/dL	12.5	17	7.38514	7.38514	mmol/L
CDISCPIL...	LB	01-701-1130	244	HGB	Hemoglobin	HEMATOL...	11.1	g/dL	12.5	17	6.88866	6.88866	mmol/L
CDISCPIL...	LB	01-701-1130	275	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L
CDISCPIL...	LB	01-701-1130	309	HGB	Hemoglobin	HEMATOL...	12.2	g/dL	12.5	17	7.57132	7.57132	mmol/L

If we now use "Tools – Filtering – Remove filters" and use "Yes – remove on all datasets", the original views with all the subjects are displayed again.

Resulting in:



STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES
CDISCPIL...	LB	01-701-1015	1	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	39	ALB	Albumin	CHEMISTRY	3.9
CDISCPIL...	LB	01-701-1015	74	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	104	ALB	Albumin	CHEMISTRY	3.7
CDISCPIL...	LB	01-701-1015	134	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	164	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	199	ALB	Albumin	CHEMISTRY	3.7
CDISCPIL...	LB	01-701-1015	229	ALB	Albumin	CHEMISTRY	3.7
CDISCPIL...	LB	01-701-1015	259	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	294	ALB	Albumin	CHEMISTRY	3.8
CDISCPIL...	LB	01-701-1015	2	ALP	Alkaline Ph...	CHEMISTRY	34
CDISCPIL...	LB	01-701-1015	40	ALP	Alkaline Ph...	CHEMISTRY	50

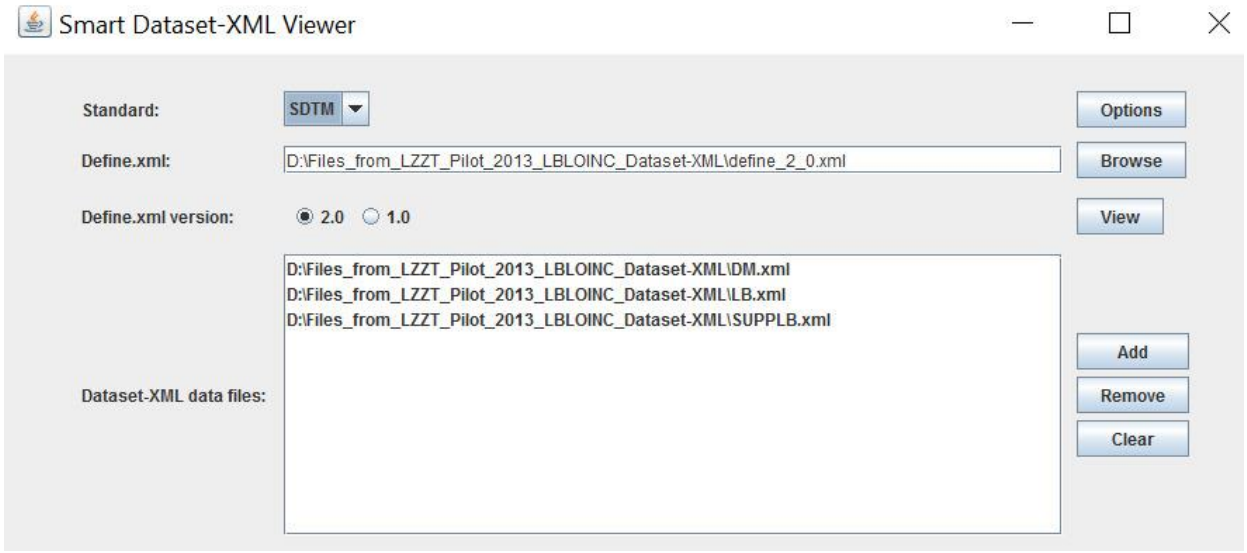
If, however filtering was applied during loading of the datasets, the datasets will be displayed with that filter applied.

17. Working with Supplemental Qualifiers

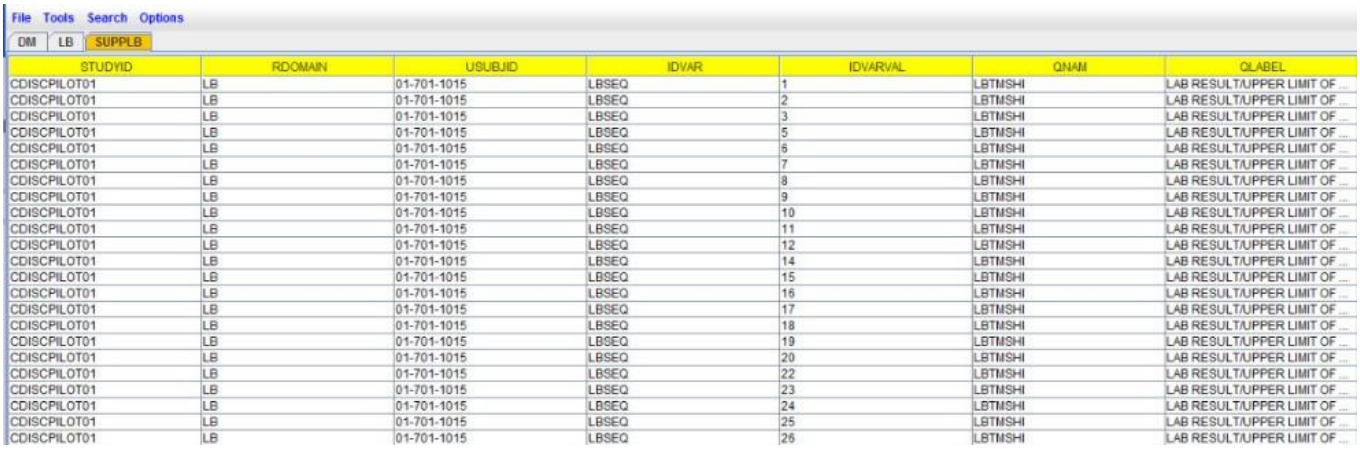
One of the great advantages of the Dataset-XML standard is that supplemental qualifier variables can be kept in the original dataset, i.e. it is not necessary to generate (or "split of") SUPP-- datasets. These variables can come after the timing variables, but need to be defined as such in the "define.xml" structure by the "Role" attribute.

In many cases however, SUPP-- datasets will still be present in the submission. Therefore, the "Smart Dataset-XML Viewer" has a number of features to connect data points in a SUPP-- dataset to the corresponding data point in the parent domain.

As an example, let us load the DM.xml dataset, and both the LB.xml and SUPPLB.xml datasets:



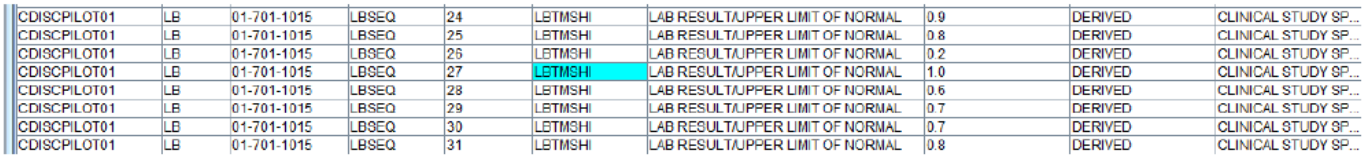
This results in a new window with three tabs, one containing the DM table, one containing the LB table and one containing the SUPPLB table, the latter looking like:



STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL
CDISCPILOT01	LB	01-701-1015	LBSEQ	1	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	2	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	3	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	5	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	6	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	7	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	8	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	9	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	10	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	11	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	12	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	14	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	15	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	16	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	17	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	18	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	19	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	20	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	22	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	23	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	24	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	25	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...
CDISCPILOT01	LB	01-701-1015	LBSEQ	26	LBTMSHI	LAB RESULT/UPPER LIMIT OF ...

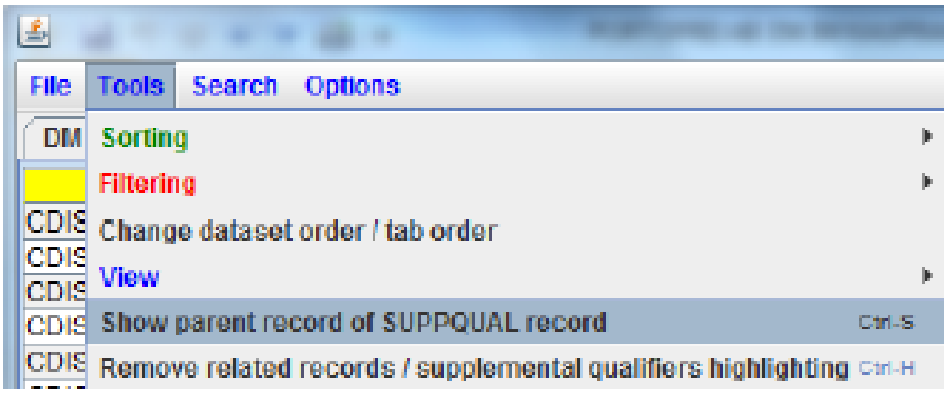
Further inspection of this table reveals that there are two different supplemental qualifiers: LBTBSHI ("Lab Result/Upper Limit of Normal") and ENDPOINT ("Endpoint Flag Value"). Ideally, we would like to see these as two columns in the LB table. This has however not been implemented yet (the box "Bring SUPPQUAL data back to original dataset" is currently also greyed out). We can however inspect the records in the SUPPLB domain and quickly find their parent record(s) back in the LB domain.

In the SUPPLB table, select a single row that is of interest to you. For example:



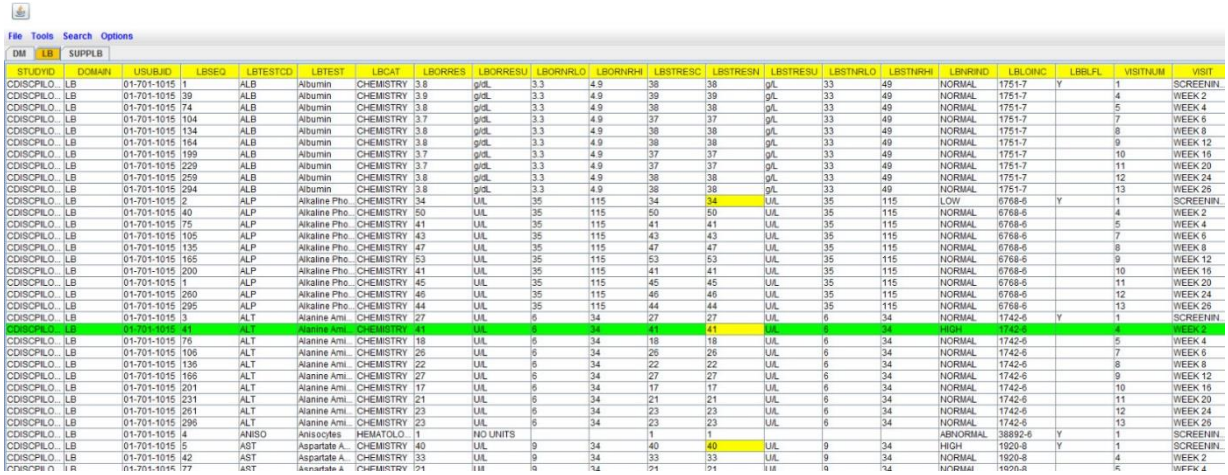
CDISCPILOT01	LB	01-701-1015	LBSEQ	24	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.9	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	25	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.8	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	26	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.2	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	27	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	1.0	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	28	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.6	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	29	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.7	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	30	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.7	DERIVED	CLINICAL STUDY SP...
CDISCPILOT01	LB	01-701-1015	LBSEQ	31	LBTMSHI	LAB RESULT/UPPER LIMIT OF NORMAL	0.8	DERIVED	CLINICAL STUDY SP...

And then use the menu "Tools – Show parent record of SUPPQUAL record":



The software will now start searching in non-suppqual domains for the parent record, based on the values of IDVAR and IDVARVAL. Soon the following dialog is shown:

And when clicking "OK", the system immediately selects the LB table and scrolls to and highlights the related record:

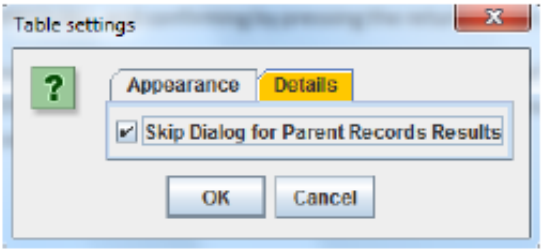


STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI	LBSTRESC	LBSTRESN	LBSTRESU	LBSTNRLO	LBSTNRHI	LBNRIND	LBLOINC	LBBLFL	VISITNUM	VISIT
CDISCPILO...	LB	01-701-1015	1	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7	Y	1	SCREENIN...
CDISCPILO...	LB	01-701-1015	39	ALB	Albumin	CHEMISTRY	3.9	g/dL	3.3	4.9	39	39	g/L	33	49	NORMAL	1751-7		4	WEEK 2
CDISCPILO...	LB	01-701-1015	74	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7		5	WEEK 4
CDISCPILO...	LB	01-701-1015	104	ALB	Albumin	CHEMISTRY	3.7	g/dL	3.3	4.9	37	37	g/L	33	49	NORMAL	1751-7		7	WEEK 6
CDISCPILO...	LB	01-701-1015	134	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7		8	WEEK 8
CDISCPILO...	LB	01-701-1015	164	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7		9	WEEK 12
CDISCPILO...	LB	01-701-1015	199	ALB	Albumin	CHEMISTRY	3.7	g/dL	3.3	4.9	37	37	g/L	33	49	NORMAL	1751-7		10	WEEK 16
CDISCPILO...	LB	01-701-1015	229	ALB	Albumin	CHEMISTRY	3.7	g/dL	3.3	4.9	37	37	g/L	33	49	NORMAL	1751-7		11	WEEK 20
CDISCPILO...	LB	01-701-1015	259	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7		12	WEEK 24
CDISCPILO...	LB	01-701-1015	294	ALB	Albumin	CHEMISTRY	3.8	g/dL	3.3	4.9	38	38	g/L	33	49	NORMAL	1751-7		13	WEEK 26
CDISCPILO...	LB	01-701-1015	2	ALP	Alkaline Pho...	CHEMISTRY	34	U/L	35	115	34	34	U/L	35	115	LOW	6768-6	Y	1	SCREENIN...
CDISCPILO...	LB	01-701-1015	40	ALP	Alkaline Pho...	CHEMISTRY	50	U/L	35	115	50	50	U/L	35	115	NORMAL	6768-6		4	WEEK 2
CDISCPILO...	LB	01-701-1015	75	ALP	Alkaline Pho...	CHEMISTRY	41	U/L	35	115	41	41	U/L	35	115	NORMAL	6768-6		5	WEEK 4
CDISCPILO...	LB	01-701-1015	105	ALP	Alkaline Pho...	CHEMISTRY	43	U/L	35	115	43	43	U/L	35	115	NORMAL	6768-6		7	WEEK 6
CDISCPILO...	LB	01-701-1015	135	ALP	Alkaline Pho...	CHEMISTRY	47	U/L	35	115	47	47	U/L	35	115	NORMAL	6768-6		8	WEEK 8
CDISCPILO...	LB	01-701-1015	165	ALP	Alkaline Pho...	CHEMISTRY	53	U/L	35	115	53	53	U/L	35	115	NORMAL	6768-6		9	WEEK 12
CDISCPILO...	LB	01-701-1015	200	ALP	Alkaline Pho...	CHEMISTRY	41	U/L	35	115	41	41	U/L	35	115	NORMAL	6768-6		10	WEEK 16
CDISCPILO...	LB	01-701-1015	1	ALP	Alkaline Pho...	CHEMISTRY	45	U/L	35	115	45	45	U/L	35	115	NORMAL	6768-6		11	WEEK 20
CDISCPILO...	LB	01-701-1015	260	ALP	Alkaline Pho...	CHEMISTRY	46	U/L	35	115	46	46	U/L	35	115	NORMAL	6768-6		12	WEEK 24
CDISCPILO...	LB	01-701-1015	295	ALP	Alkaline Pho...	CHEMISTRY	44	U/L	35	115	44	44	U/L	35	115	NORMAL	6768-6		13	WEEK 26
CDISCPILO...	LB	01-701-1015	3	ALT	Alanine Ami...	CHEMISTRY	27	U/L	6	34	27	27	U/L	6	34	NORMAL	1742-6	Y	1	SCREENIN...
CDISCPILO...	LB	01-701-1015	41	ALT	Alanine Ami...	CHEMISTRY	41	U/L	6	34	41	41	U/L	6	34	HIGH	1742-6		4	WEEK 2
CDISCPILO...	LB	01-701-1015	76	ALT	Alanine Ami...	CHEMISTRY	18	U/L	6	34	18	18	U/L	6	34	NORMAL	1742-6		5	WEEK 4
CDISCPILO...	LB	01-701-1015	106	ALT	Alanine Ami...	CHEMISTRY	26	U/L	6	34	26	26	U/L	6	34	NORMAL	1742-6		7	WEEK 6
CDISCPILO...	LB	01-701-1015	136	ALT	Alanine Ami...	CHEMISTRY	22	U/L	6	34	22	22	U/L	6	34	NORMAL	1742-6		8	WEEK 8
CDISCPILO...	LB	01-701-1015	166	ALT	Alanine Ami...	CHEMISTRY	27	U/L	6	34	27	27	U/L	6	34	NORMAL	1742-6		9	WEEK 12
CDISCPILO...	LB	01-701-1015	201	ALT	Alanine Ami...	CHEMISTRY	17	U/L	6	34	17	17	U/L	6	34	NORMAL	1742-6		10	WEEK 16
CDISCPILO...	LB	01-701-1015	231	ALT	Alanine Ami...	CHEMISTRY	21	U/L	6	34	21	21	U/L	6	34	NORMAL	1742-6		11	WEEK 20
CDISCPILO...	LB	01-701-1015	261	ALT	Alanine Ami...	CHEMISTRY	23	U/L	6	34	23	23	U/L	6	34	NORMAL	1742-6		12	WEEK 24
CDISCPILO...	LB	01-701-1015	296	ALT	Alanine Ami...	CHEMISTRY	23	U/L	6	34	23	23	U/L	6	34	NORMAL	1742-6		13	WEEK 26
CDISCPILO...	LB	01-701-1015	4	ANISO	Anisocytes	HEMATOLO...	1	NO UNITS			1	1				ABNORMAL	38892-6	Y	1	SCREENIN...
CDISCPILO...	LB	01-701-1015	5	AST	Aspartate A...	CHEMISTRY	40	U/L	9	34	40	40	U/L	9	34	HIGH	1920-8	Y	1	SCREENIN...
CDISCPILO...	LB	01-701-1015	42	AST	Aspartate A...	CHEMISTRY	33	U/L	9	34	33	33	U/L	9	34	NORMAL	1920-8		4	WEEK 2

All the same can also be achieved without needing the mouse by using the shortcut CTRL-S ("S" for SUPPQUAL), and confirming by pressing the return button.

One can now also easily switch back and forth (i.e. "toggle") between the SUPPLB table and LB table (either by clicking the tabs, or use CTRL-B, or use the menu "Tools – View – "Show last selected table" to inspect further details.

The popping up of these messages can be switched off using the menu "Options – Setting" and then selecting the tab "Details" and then selecting "Skip Dialogue for Parent Records Results":



In order to remove the highlighting in the LB table, use the menu "Tools – Remove related records / supplemental qualifiers highlighting".

Remark: in case the supplemental qualifier table has "--CAT" as "IDVAR", then in most cases there are multiple records in the parent dataset that are related to that supplemental qualifier. In this case all these records will be highlighted when the software toggles to the parent dataset.

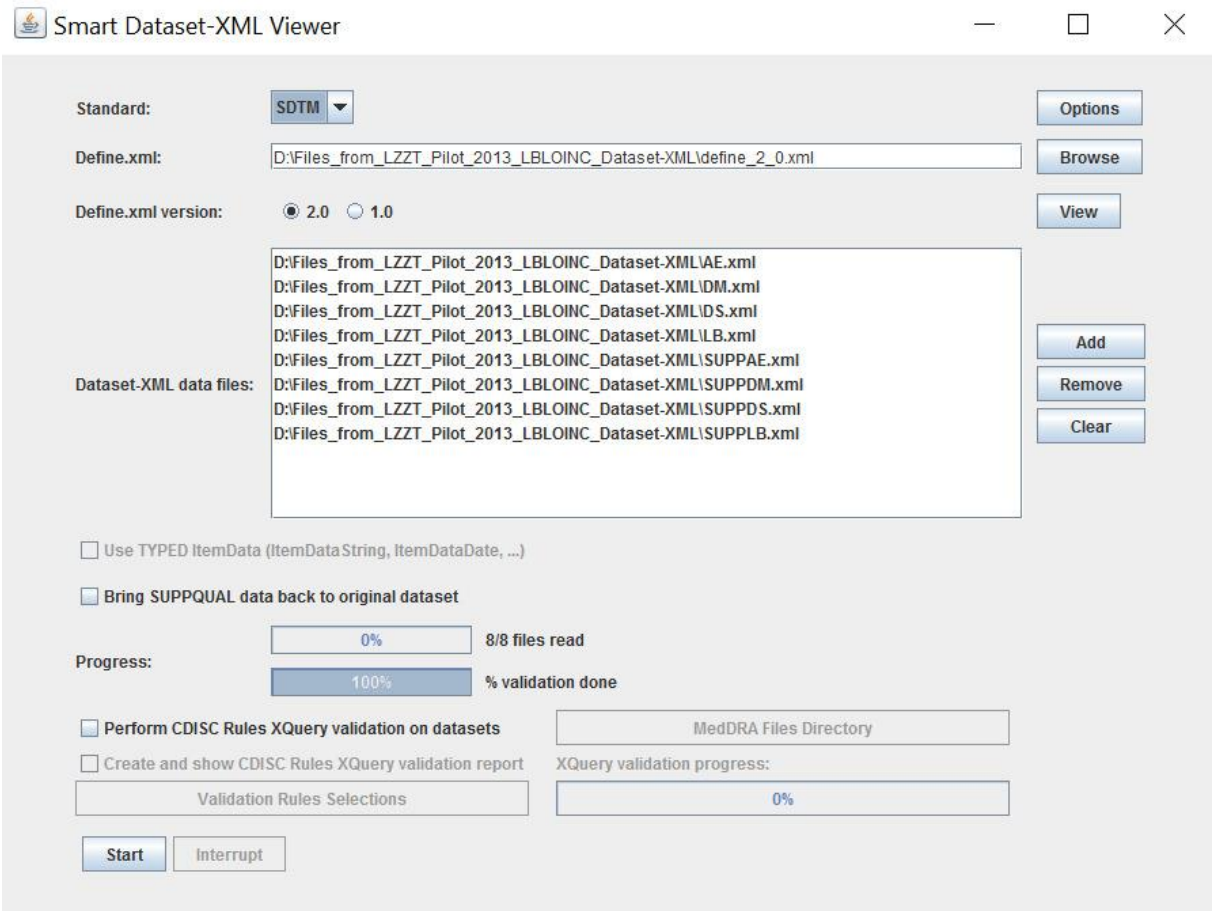
18. Bringing Non-Standard Variables Back to the Parent Dataset

Unfortunately, it is currently not allowed yet by CDISC nor the FDA to keep the non-standard variables in the data set where they really belong. So sponsors usually "ban" these data points to a supplemental qualifier data set in the very last step of preparing the submission.

Using the menu "Tools – Show parent record of SUPPQUAL record" one can already toggle between a record in the SUPPxx data set and its parent record or records in the parent data set. A new feature that recently has been added is to merge non-standard variable values back into their parent domain during loading. This feature is still experimental and can be pretty computing intensive in case both the SUPPxx dataset and the parent data set contain many records (typical example: SUPPLB with LB).

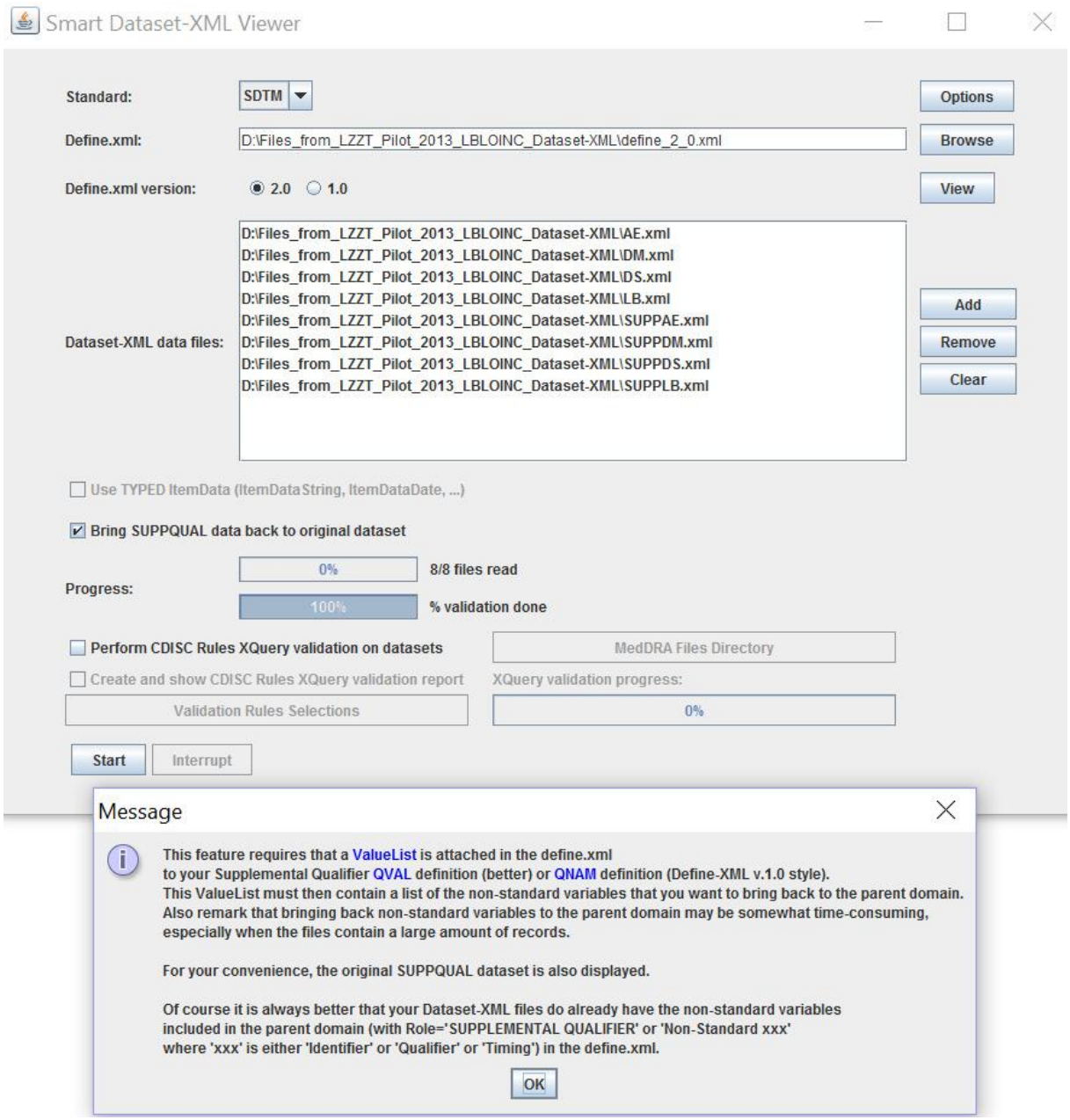
We will demonstrate this feature using the files that can be found in the directory "Files_from_LZZT_Pilot_2013_LBLOINC_Dataset-XML".

After having loaded the define.xml file, we select the following files: "DM.xml", "AE.xml". "DS.xml", and "LB.xml" as well as "SUPPAE.xml", "SUPPDM.xml", "SUPPDS.xml" and "SUPPLB.xml":



In order to bring back non-standard variables to the parent domain during loading, we check the checkbox "Bring SUPPQUAL data back to original dataset".

A message is displayed explaining that this is an experimental feature and may require extra computing time:

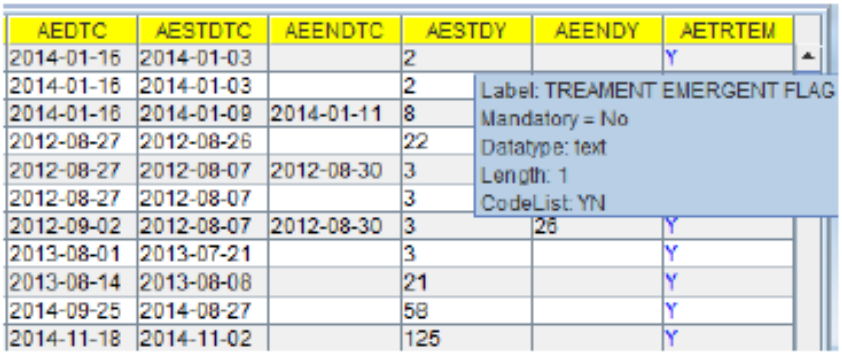


Remark that this feature relies that there is a "ValueList" attached to the QNAM variable of the supplemental qualifier in the define.xml, containing all possible values for QNAM which essentially equals to a list with the non-standard variables and their metadata (name, data type, length, whether a code list is associated, etc.).

Our AE / SUPPAE, and DM / SUPPDM, and DS / SUPPDS data sets are all very small, so merging the non-standard variables back to the parent domain will be very fast. In the case of LB / SUPPLB, the system will need to merge almost 65,000 SUPPLB records back to the parent data

set which contains almost 60,000 records. Although this seems a lot, it is done in less than half a minute[10].

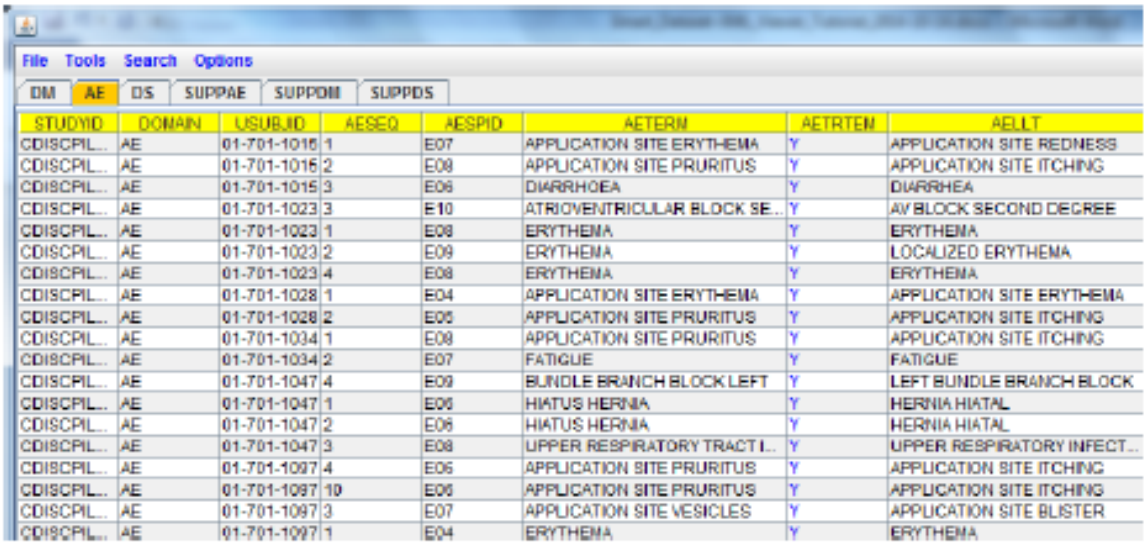
The AE data set as displayed in the viewer is then displayed like:



AEDTC	AESTDTC	AEENDTC	AESTDY	AEENDY	AETRTEM
2014-01-16	2014-01-03		2		Y
2014-01-16	2014-01-03		2		
2014-01-16	2014-01-09	2014-01-11	8		
2012-08-27	2012-08-26		22		
2012-08-27	2012-08-07	2012-08-30	3		
2012-08-27	2012-08-07		3		
2012-09-02	2012-08-07	2012-08-30	3	26	Y
2013-08-01	2013-07-21		3		Y
2013-08-14	2013-08-06		21		Y
2014-09-25	2014-08-27		58		Y
2014-11-18	2014-11-02		125		Y

Label: TREAMENT EMERGENT FLAG
Mandatory = No
Datatype: text
Length: 1
CodeList: YN

An extra column is added to the display, showing the values for "AETRTEM" (label: "treatment emergent flag"). These values are coloured blue, indicating this is a non-standard variable. Of course one can now also move this column to the left, e.g. immediately after "AETERM":

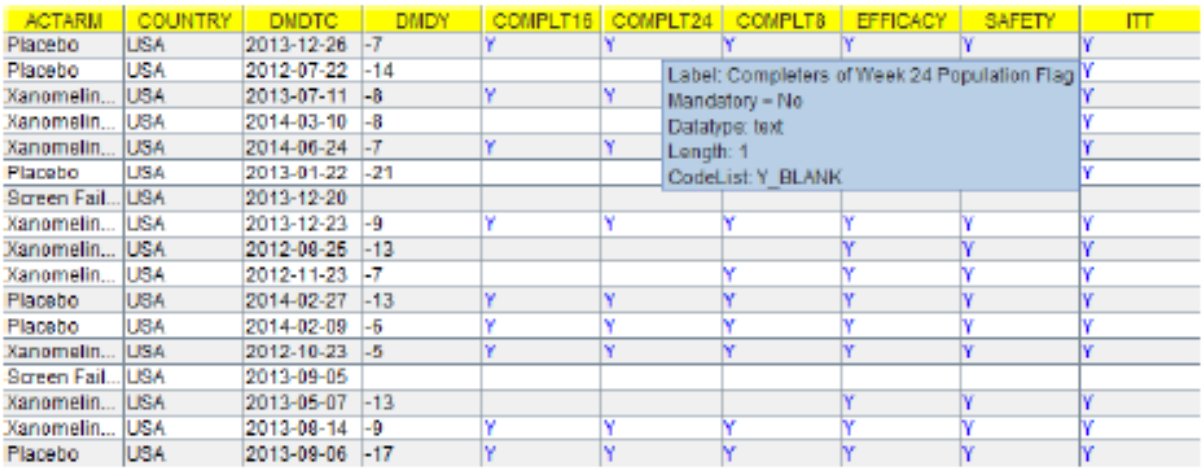


File Tools Search Options

DM AE DS SUPPAE SUPPDM SUPPDS

STUDYID	DOMAIN	USUBJID	AESEQ	AESPID	AETERM	AETRTEM	AELLT
CDISCPIL...	AE	01-701-1015	1	E07	APPLICATION SITE ERYTHEMA	Y	APPLICATION SITE REDNESS
CDISCPIL...	AE	01-701-1015	2	E08	APPLICATION SITE PRURITUS	Y	APPLICATION SITE ITCHING
CDISCPIL...	AE	01-701-1015	3	E06	DIARRHOEA	Y	DIARRHEA
CDISCPIL...	AE	01-701-1023	3	E10	ATRIOVENTRICULAR BLOCK SE...	Y	AV BLOCK SECOND DEGREE
CDISCPIL...	AE	01-701-1023	1	E08	ERYTHEMA	Y	ERYTHEMA
CDISCPIL...	AE	01-701-1023	2	E09	ERYTHEMA	Y	LOCALIZED ERYTHEMA
CDISCPIL...	AE	01-701-1023	4	E08	ERYTHEMA	Y	ERYTHEMA
CDISCPIL...	AE	01-701-1028	1	E04	APPLICATION SITE ERYTHEMA	Y	APPLICATION SITE ERYTHEMA
CDISCPIL...	AE	01-701-1028	2	E05	APPLICATION SITE PRURITUS	Y	APPLICATION SITE ITCHING
CDISCPIL...	AE	01-701-1034	1	E08	APPLICATION SITE PRURITUS	Y	APPLICATION SITE ITCHING
CDISCPIL...	AE	01-701-1034	2	E07	FATIGUE	Y	FATIGUE
CDISCPIL...	AE	01-701-1047	4	E09	BUNDLE BRANCH BLOCK LEFT	Y	LEFT BUNDLE BRANCH BLOCK
CDISCPIL...	AE	01-701-1047	1	E06	HIATUS HERNIA	Y	HERNIA HIATAL
CDISCPIL...	AE	01-701-1047	2	E06	HIATUS HERNIA	Y	HERNIA HIATAL
CDISCPIL...	AE	01-701-1047	3	E08	UPPER RESPIRATORY TRACT I...	Y	UPPER RESPIRATORY INFECT...
CDISCPIL...	AE	01-701-1097	4	E06	APPLICATION SITE PRURITUS	Y	APPLICATION SITE ITCHING
CDISCPIL...	AE	01-701-1097	10	E05	APPLICATION SITE PRURITUS	Y	APPLICATION SITE ITCHING
CDISCPIL...	AE	01-701-1097	3	E07	APPLICATION SITE VESICLES	Y	APPLICATION SITE BLISTER
CDISCPIL...	AE	01-701-1097	1	E04	ERYTHEMA	Y	ERYTHEMA

The result for the DM data set is:



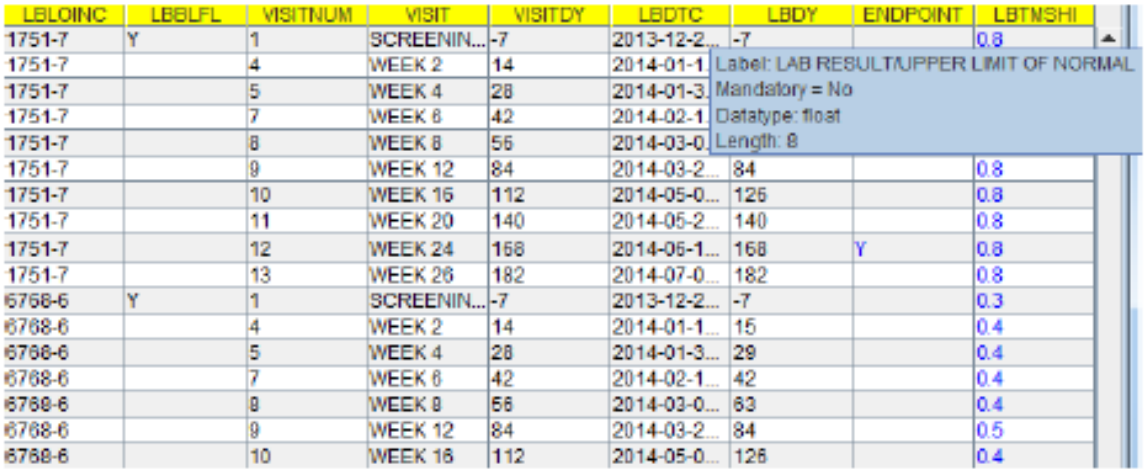
ACTARM	COUNTRY	DMDTC	DMDY	COMPLT16	COMPLT24	COMPLT8	EFFICACY	SAFETY	ITT
Placebo	USA	2013-12-26	-7	Y	Y	Y	Y	Y	Y
Placebo	USA	2012-07-22	-14						Y
Xanomelin...	USA	2013-07-11	-8	Y	Y				Y
Xanomelin...	USA	2014-03-10	-8						Y
Xanomelin...	USA	2014-06-24	-7	Y	Y				Y
Placebo	USA	2013-01-22	-21						Y
Screen Fail...	USA	2013-12-20							
Xanomelin...	USA	2013-12-23	-9	Y	Y	Y	Y	Y	Y
Xanomelin...	USA	2012-09-26	-13				Y	Y	Y
Xanomelin...	USA	2012-11-23	-7			Y	Y	Y	Y
Placebo	USA	2014-02-27	-13	Y	Y	Y	Y	Y	Y
Placebo	USA	2014-02-09	-5	Y	Y	Y	Y	Y	Y
Xanomelin...	USA	2012-10-23	-5	Y	Y	Y	Y	Y	Y
Screen Fail...	USA	2013-09-05							
Xanomelin...	USA	2013-05-07	-13				Y	Y	Y
Xanomelin...	USA	2013-09-14	-9	Y	Y	Y	Y	Y	Y
Placebo	USA	2013-09-06	-17	Y	Y	Y	Y	Y	Y

Label: Completers of Week 24 Population Flag
Mandatory = No
Datatype: text
Length: 1
CodeList: Y_BLANK

[10] We are of course very curious about user experiences with data sets containing millions of data points!

One notices that the metadata for the non-standard variable as displayed as column header tooltips.

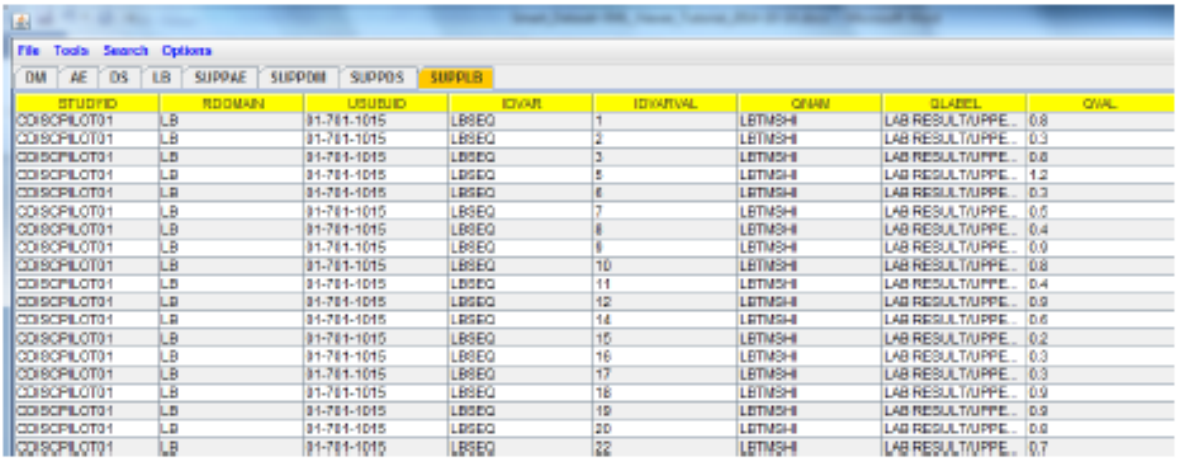
For the LB dataset we obtain:



LBLOINC	LBBLFL	VISITNUM	VISIT	VISITDY	LBDTC	LBDY	ENDPOINT	LBTMSHI
1751-7	Y	1	SCREENIN...	-7	2013-12-2...	-7		0.8
1751-7		4	WEEK 2	14	2014-01-1			
1751-7		5	WEEK 4	28	2014-01-3			
1751-7		7	WEEK 6	42	2014-02-1			
1751-7		8	WEEK 8	56	2014-03-0			
1751-7		9	WEEK 12	84	2014-03-2...	84		0.8
1751-7		10	WEEK 16	112	2014-05-0...	126		0.8
1751-7		11	WEEK 20	140	2014-05-2...	140		0.8
1751-7		12	WEEK 24	168	2014-06-1...	168	Y	0.8
1751-7		13	WEEK 26	182	2014-07-0...	182		0.8
6768-6	Y	1	SCREENIN...	-7	2013-12-2...	-7		0.3
6768-6		4	WEEK 2	14	2014-01-1...	15		0.4
6768-6		5	WEEK 4	28	2014-01-3...	29		0.4
6768-6		7	WEEK 6	42	2014-02-1...	42		0.4
6768-6		8	WEEK 8	56	2014-03-0...	63		0.4
6768-6		9	WEEK 12	84	2014-03-2...	84		0.5
6768-6		10	WEEK 16	112	2014-05-0...	126		0.4

Label: LAB RESULT/UPPER LIMIT OF NORMAL
Mandatory = No
Datatype: float
Length: 8

As the current implementation is dependent on whether a ValueList has been assigned to the QNAM variable in the supplemental qualifier data set, the latter is also displayed, so that one can still always use the classic mechanism using the menu "Tools - Show parent record of SUPPQUAL record" (or using CTRL-S):

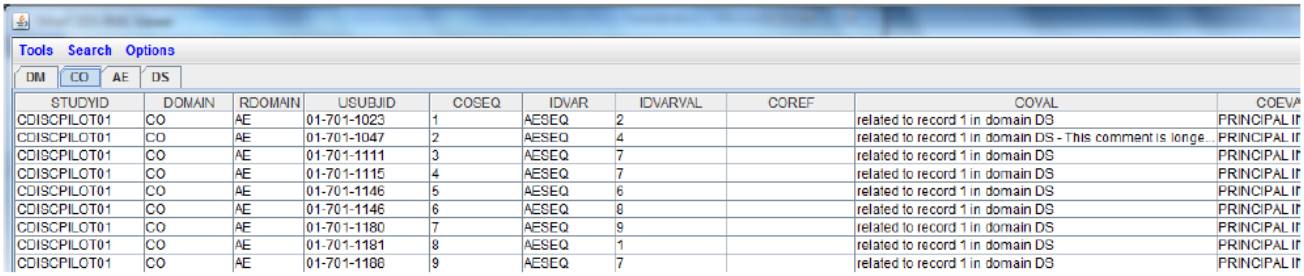


19. Working with Comments (CO) Dataset

A similar mechanism applies to records in the CO (Comments) dataset.

In the set of test files, you can also find a set of datasets where the supplemental qualifiers are not located anymore in SUPPxx datasets, but have been integrated in the parent datasets. This test set also contains a CO.xml dataset.

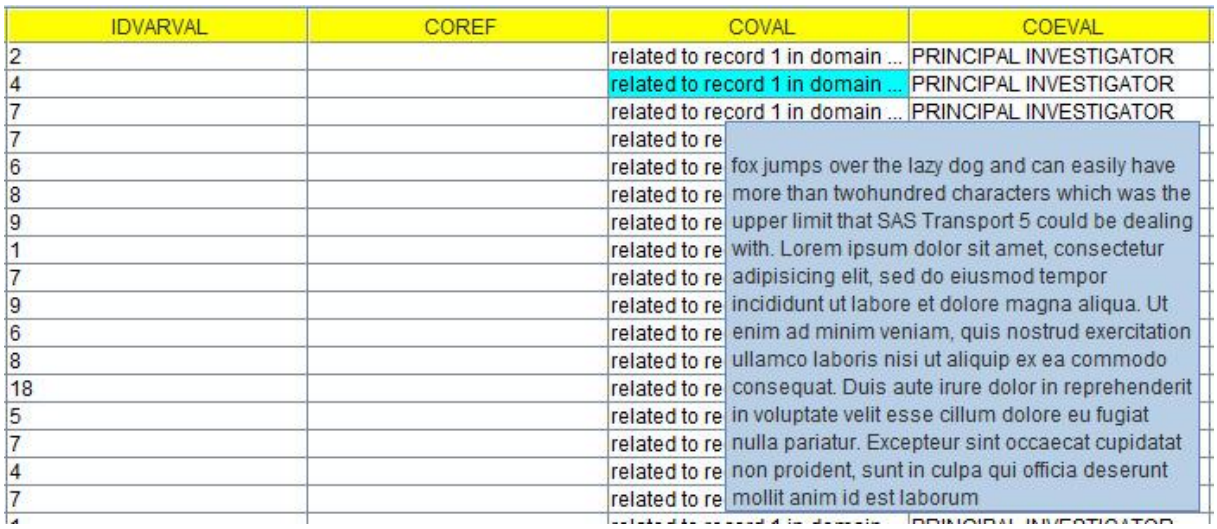
First load the define.xml files from this set of files. If you did not do already, also clear the list with files, and start adding the following files from the same directory: DM.xml, CO.xml, AE.xml and DS.xml. Start loading this set of files and then select the CO table. You will find:


Tools Search Options

DM CO AE DS

STUDYID	DOMAIN	RDOMAIN	USUBJID	COSEQ	IDVAR	IDVARVAL	COREF	COVAL	COEVA
CDISCPILOT01	CO	AE	01-701-1023	1	AESEQ	2		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1047	2	AESEQ	4		related to record 1 in domain DS - This comment is longe...	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1111	3	AESEQ	7		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1115	4	AESEQ	7		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1146	5	AESEQ	6		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1146	6	AESEQ	8		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1180	7	AESEQ	9		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1181	8	AESEQ	1		related to record 1 in domain DS	PRINCIPAL IN
CDISCPILOT01	CO	AE	01-701-1188	9	AESEQ	7		related to record 1 in domain DS	PRINCIPAL IN

Now hold the mouse over the COVAL cell in the second row. You will see:


IDVARVAL	COREF	COVAL	COEVAL
2		related to record 1 in domain ...	PRINCIPAL INVESTIGATOR
4		related to record 1 in domain ...	PRINCIPAL INVESTIGATOR
7		related to record 1 in domain ...	PRINCIPAL INVESTIGATOR
7		related to re	
6		related to re	
8		related to re	
9		related to re	
1		related to re	
7		related to re	
9		related to re	
6		related to re	
8		related to re	
18		related to re	
5		related to re	
7		related to re	
4		related to re	
7		related to re	

fox jumps over the lazy dog and can easily have more than twohundred characters which was the upper limit that SAS Transport 5 could be dealing with. Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

The value of COVAL contains more than 200 characters. When using SAS Transport 5, the value would then have to be split over several columns COVAL, COVAL1, COVAL2, ..., as in SAS Transport 5, the length of a variable value was limited to 200 characters. As Dataset -XML however uses XML technology, there is no such limitation anymore.

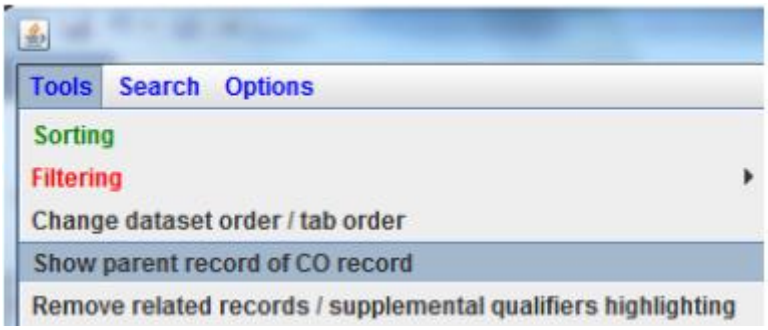
In order to find the parent record for a record in the CO domain, select any cell from the record you are interested in, e.g.:


Tools Search Options

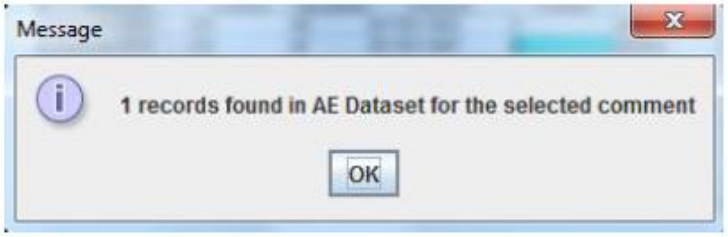
DM CO AE DS

STUDYID	DOMAIN	RDOMAIN	USUBJID	COSEQ	IDVAR	IDVARVAL	CO..	COVAL
CDISCPILOT01	CO	AE	01-701-1023	1	AESEQ	2		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1047	2	AESEQ	4		related to record 1 in domain DS - Th
CDISCPILOT01	CO	AE	01-701-1111	3	AESEQ	7		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1115	4	AESEQ	7		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1146	5	AESEQ	6		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1146	6	AESEQ	8		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1180	7	AESEQ	9		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1181	8	AESEQ	1		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1188	9	AESEQ	7		related to record 1 in domain DS
CDISCPILOT01	CO	AE	01-701-1211	10	AESEQ	9		related to record 1 in domain DS

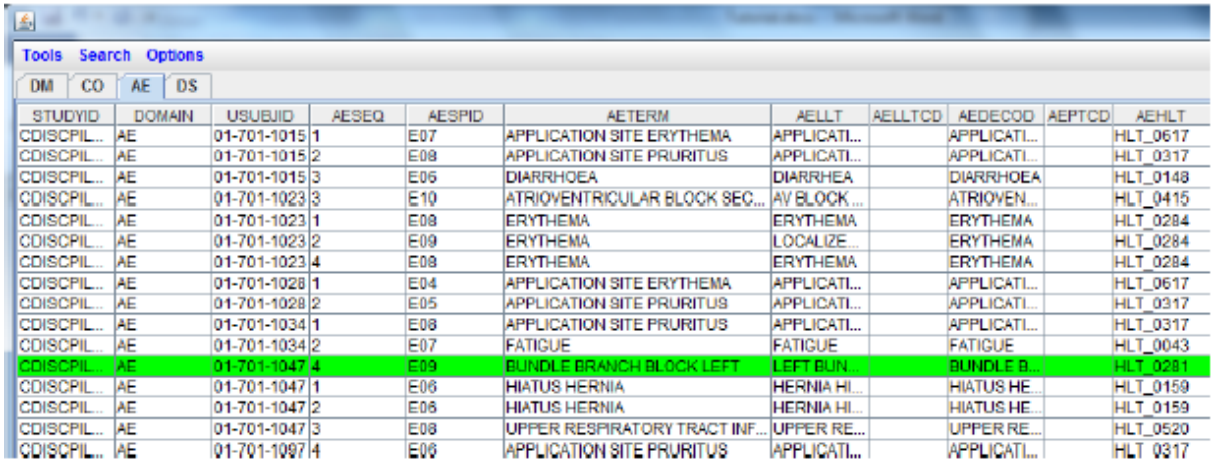
Then use the menu "Tools – Show parent record of CO record", or use the keyboard shortcut CTRL-C ("C" for "Comment")



The system will then start searching for the parent record of the selected CO record, using the information from RDOMAIN, IDVAR and IDVARVAL, and soon report:



When then selecting the "AE" table, the parent record(s) are highlighted:


Tools Search Options

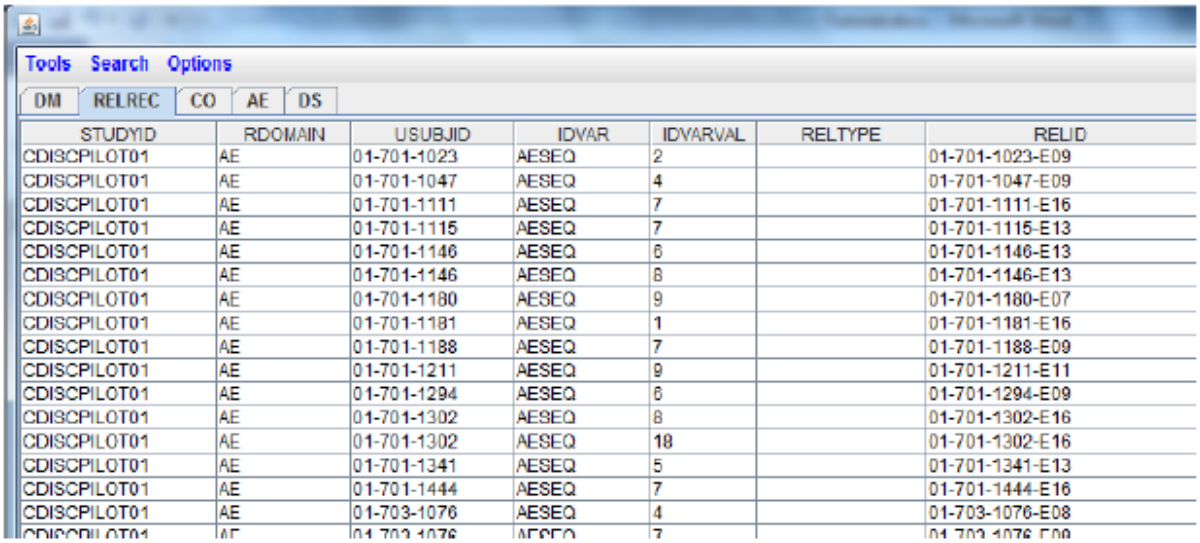
DM CO AE DS

STUDYID	DOMAIN	USUBJID	AESEQ	AESPID	AETERM	AELLT	AELLTCD	AEDECOD	AEPTCD	AEHLT
CDISCPIL...	AE	01-701-1015	1	E07	APPLICATION SITE ERYTHEMA	APPLICATI...		APPLICATI...		HLT_0617
CDISCPIL...	AE	01-701-1015	2	E08	APPLICATION SITE PRURITUS	APPLICATI...		APPLICATI...		HLT_0317
CDISCPIL...	AE	01-701-1015	3	E06	DIARRHOEA	DIARRHEA		DIARRHOEA		HLT_0148
CDISCPIL...	AE	01-701-1023	3	E10	ATRIOVENTRICULAR BLOCK SEC...	AV BLOCK ...		ATRIOVEN...		HLT_0415
CDISCPIL...	AE	01-701-1023	1	E08	ERYTHEMA	ERYTHEMA		ERYTHEMA		HLT_0284
CDISCPIL...	AE	01-701-1023	2	E09	ERYTHEMA	LOCALIZE...		ERYTHEMA		HLT_0284
CDISCPIL...	AE	01-701-1023	4	E08	ERYTHEMA	ERYTHEMA		ERYTHEMA		HLT_0284
CDISCPIL...	AE	01-701-1028	1	E04	APPLICATION SITE ERYTHEMA	APPLICATI...		APPLICATI...		HLT_0617
CDISCPIL...	AE	01-701-1028	2	E05	APPLICATION SITE PRURITUS	APPLICATI...		APPLICATI...		HLT_0317
CDISCPIL...	AE	01-701-1034	1	E08	APPLICATION SITE PRURITUS	APPLICATI...		APPLICATI...		HLT_0317
CDISCPIL...	AE	01-701-1034	2	E07	FATIGUE	FATIGUE		FATIGUE		HLT_0043
CDISCPIL...	AE	01-701-1047	4	E09	BUNDLE BRANCH BLOCK LEFT	LEFT BUN...		BUNDLE B...		HLT_0281
CDISCPIL...	AE	01-701-1047	1	E06	HIATUS HERNIA	HERNIA HI...		HIATUS HE...		HLT_0159
CDISCPIL...	AE	01-701-1047	2	E06	HIATUS HERNIA	HERNIA HI...		HIATUS HE...		HLT_0159
CDISCPIL...	AE	01-701-1047	3	E08	UPPER RESPIRATORY TRACT INF...	UPPER RE...		UPPER RE...		HLT_0520
CDISCPIL...	AE	01-701-1097	4	E06	APPLICATION SITE PRURITUS	APPLICATI...		APPLICATI...		HLT_0317

Just like in the case of SUPPQUAL records, one can now switch back and forth between the CO table and the AE table (e.g. using CTRL-B) to inspect further details.

20. Working with Related Records (RELREC Dataset)

Now also load the RELREC.xml dataset from the same directory. The final result is:

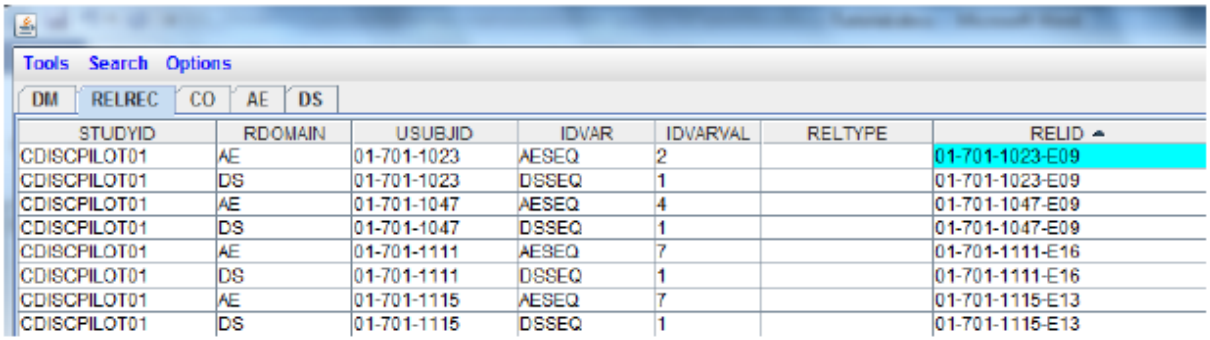


Tools Search Options

DM | RELREC | CO | AE | DS

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
CDISCPILOT01	AE	01-701-1023	AESEQ	2		01-701-1023-E09
CDISCPILOT01	AE	01-701-1047	AESEQ	4		01-701-1047-E09
CDISCPILOT01	AE	01-701-1111	AESEQ	7		01-701-1111-E16
CDISCPILOT01	AE	01-701-1115	AESEQ	7		01-701-1115-E13
CDISCPILOT01	AE	01-701-1146	AESEQ	6		01-701-1146-E13
CDISCPILOT01	AE	01-701-1146	AESEQ	8		01-701-1146-E13
CDISCPILOT01	AE	01-701-1180	AESEQ	9		01-701-1180-E07
CDISCPILOT01	AE	01-701-1181	AESEQ	1		01-701-1181-E16
CDISCPILOT01	AE	01-701-1188	AESEQ	7		01-701-1188-E09
CDISCPILOT01	AE	01-701-1211	AESEQ	9		01-701-1211-E11
CDISCPILOT01	AE	01-701-1294	AESEQ	6		01-701-1294-E09
CDISCPILOT01	AE	01-701-1302	AESEQ	8		01-701-1302-E16
CDISCPILOT01	AE	01-701-1302	AESEQ	18		01-701-1302-E16
CDISCPILOT01	AE	01-701-1341	AESEQ	5		01-701-1341-E13
CDISCPILOT01	AE	01-701-1444	AESEQ	7		01-701-1444-E16
CDISCPILOT01	AE	01-703-1076	AESEQ	4		01-703-1076-E08
CDISCPILOT01	AE	01-703-1076	AESEQ	7		01-703-1076-E09

In the RELREC dataset, records with the same value for "RELID" (last column in the view) contain references to records that are related. For example, for RELID with value 01-701-1023-E09 there are two entries in the dataset. This is however only visible after sorting the table on RELID:

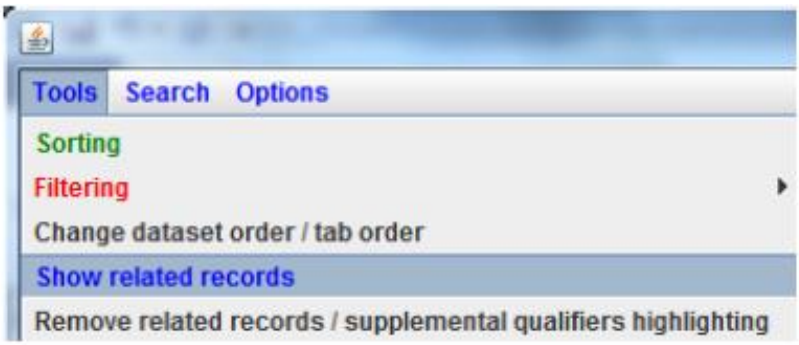


Tools Search Options

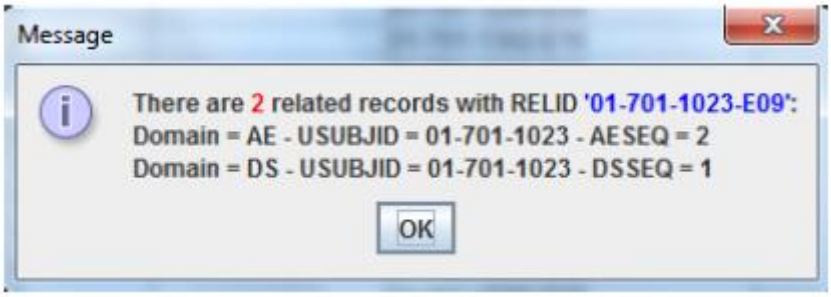
DM | RELREC | CO | AE | DS

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID ▲
CDISCPILOT01	AE	01-701-1023	AESEQ	2		01-701-1023-E09
CDISCPILOT01	DS	01-701-1023	DSSEQ	1		01-701-1023-E09
CDISCPILOT01	AE	01-701-1047	AESEQ	4		01-701-1047-E09
CDISCPILOT01	DS	01-701-1047	DSSEQ	1		01-701-1047-E09
CDISCPILOT01	AE	01-701-1111	AESEQ	7		01-701-1111-E16
CDISCPILOT01	DS	01-701-1111	DSSEQ	1		01-701-1111-E16
CDISCPILOT01	AE	01-701-1115	AESEQ	7		01-701-1115-E13
CDISCPILOT01	DS	01-701-1115	DSSEQ	1		01-701-1115-E13

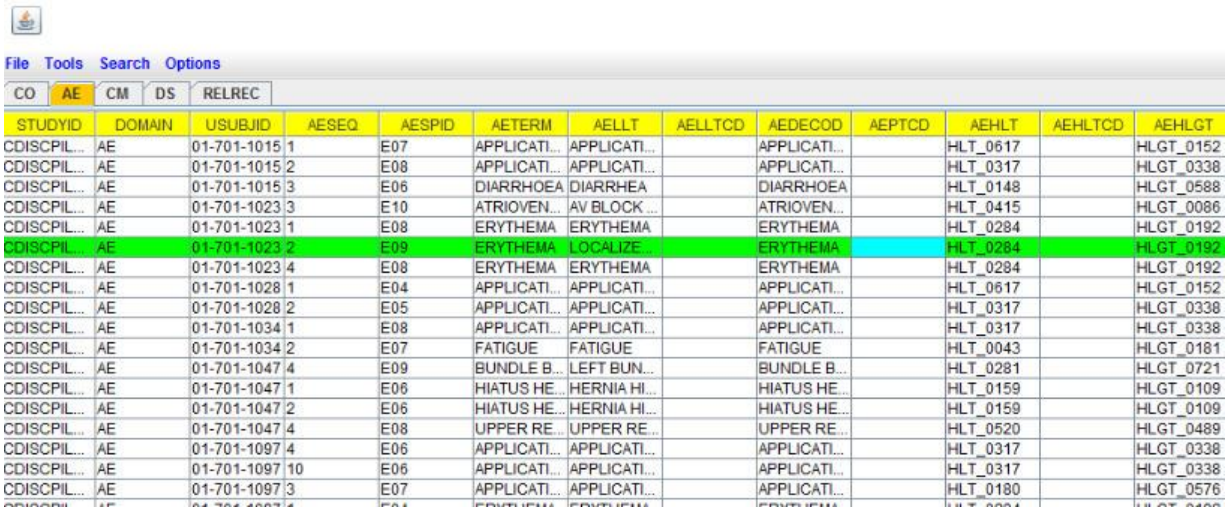
We can now easily find the related records in the AE and DS domain by selecting a cell with a RELID (in this case 01-701-1023-E09) and then using the menu "Tools – Show related records":



The system will then look up the related records based on the values of RDOMAIN, USUBJID and IDVAR and IDVARVAL. Very soon the result is obtained:

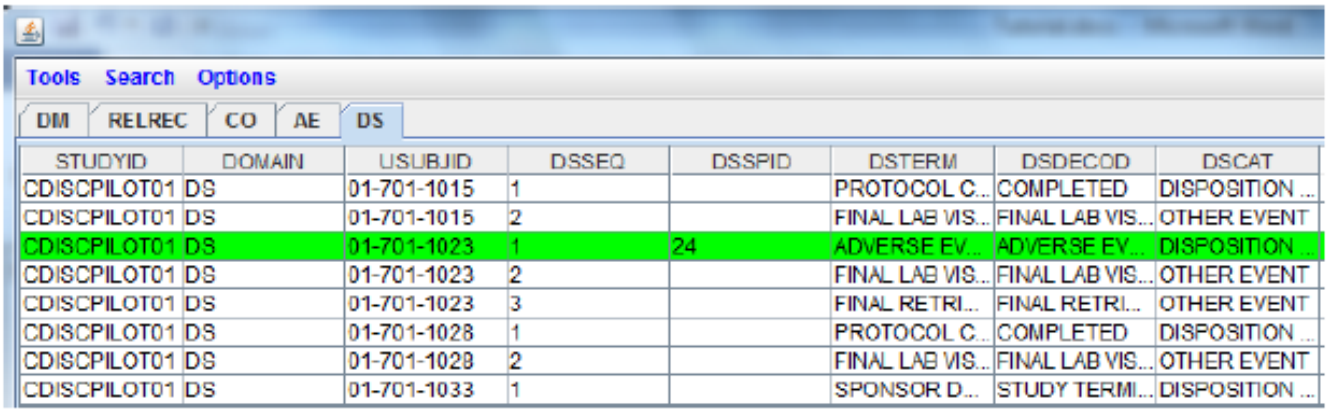


If we now click the tab for the AE table, the system automatically scrolls to the related record and highlights it:



STUDYID	DOMAIN	USUBJID	AESEQ	AESPID	AETERM	AELLT	AELLTCD	AEDECOD	AEPTCD	AEHLT	AEHLTCD	AEHLGT
CDISCPIL...	AE	01-701-1015	1	E07	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0617		HLGT_0152
CDISCPIL...	AE	01-701-1015	2	E08	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0317		HLGT_0338
CDISCPIL...	AE	01-701-1015	3	E06	DIARRHOEA	DIARRHEA		DIARRHOEA		HLT_0148		HLGT_0588
CDISCPIL...	AE	01-701-1023	3	E10	ATRIOVEN...	AV BLOCK ...		ATRIOVEN...		HLT_0415		HLGT_0086
CDISCPIL...	AE	01-701-1023	1	E08	ERYTHEMA	ERYTHEMA		ERYTHEMA		HLT_0284		HLGT_0192
CDISCPIL...	AE	01-701-1023	2	E09	ERYTHEMA	LOCALIZE...		ERYTHEMA		HLT_0284		HLGT_0192
CDISCPIL...	AE	01-701-1023	4	E08	ERYTHEMA	ERYTHEMA		ERYTHEMA		HLT_0284		HLGT_0192
CDISCPIL...	AE	01-701-1028	1	E04	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0617		HLGT_0152
CDISCPIL...	AE	01-701-1028	2	E05	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0317		HLGT_0338
CDISCPIL...	AE	01-701-1034	1	E08	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0317		HLGT_0338
CDISCPIL...	AE	01-701-1034	2	E07	FATIGUE	FATIGUE		FATIGUE		HLT_0043		HLGT_0181
CDISCPIL...	AE	01-701-1047	4	E09	BUNDLE B...	LEFT BUN...		BUNDLE B...		HLT_0281		HLGT_0721
CDISCPIL...	AE	01-701-1047	1	E06	HIATUS HE...	HERNIA HI...		HIATUS HE...		HLT_0159		HLGT_0109
CDISCPIL...	AE	01-701-1047	2	E06	HIATUS HE...	HERNIA HI...		HIATUS HE...		HLT_0159		HLGT_0109
CDISCPIL...	AE	01-701-1047	4	E08	UPPER RE...	UPPER RE...		UPPER RE...		HLT_0520		HLGT_0489
CDISCPIL...	AE	01-701-1097	4	E06	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0317		HLGT_0338
CDISCPIL...	AE	01-701-1097	10	E06	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0317		HLGT_0338
CDISCPIL...	AE	01-701-1097	3	E07	APPLICATI...	APPLICATI...		APPLICATI...		HLT_0180		HLGT_0576

Similarly, selecting the tab for the DS table, the system scrolls to the related record and highlights it:



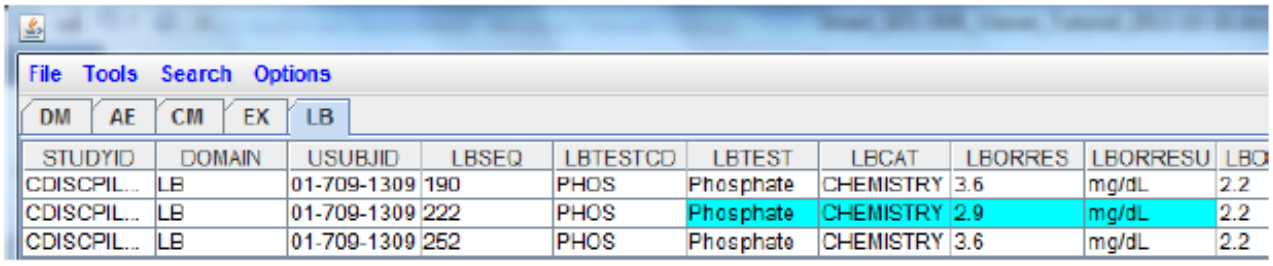
STUDYID	DOMAIN	USUBJID	DSSEQ	DSSPID	DSTERM	DSDECOD	DSCAT
CDISCPILOT01	DS	01-701-1015	1		PROTOCOL C...	COMPLETED	DISPOSITION ...
CDISCPILOT01	DS	01-701-1015	2		FINAL LAB VIS...	FINAL LAB VIS...	OTHER EVENT
CDISCPILOT01	DS	01-701-1023	1	24	ADVERSE EV...	ADVERSE EV...	DISPOSITION ...
CDISCPILOT01	DS	01-701-1023	2		FINAL LAB VIS...	FINAL LAB VIS...	OTHER EVENT
CDISCPILOT01	DS	01-701-1023	3		FINAL RETRI...	FINAL RETRI...	OTHER EVENT
CDISCPILOT01	DS	01-701-1028	1		PROTOCOL C...	COMPLETED	DISPOSITION ...
CDISCPILOT01	DS	01-701-1028	2		FINAL LAB VIS...	FINAL LAB VIS...	OTHER EVENT
CDISCPILOT01	DS	01-701-1033	1		SPONSOR D...	STUDY TERMI...	DISPOSITION ...

Clearly demonstrating that the disposition event for the record with DSSPID=24 is related to the AE with AESPID=E09.

In order to remove all highlighting of related records, use the menu "Tools – Remove related records / supplemental qualifiers highlighting", or just use the keyboard shortcut CTRL-H.

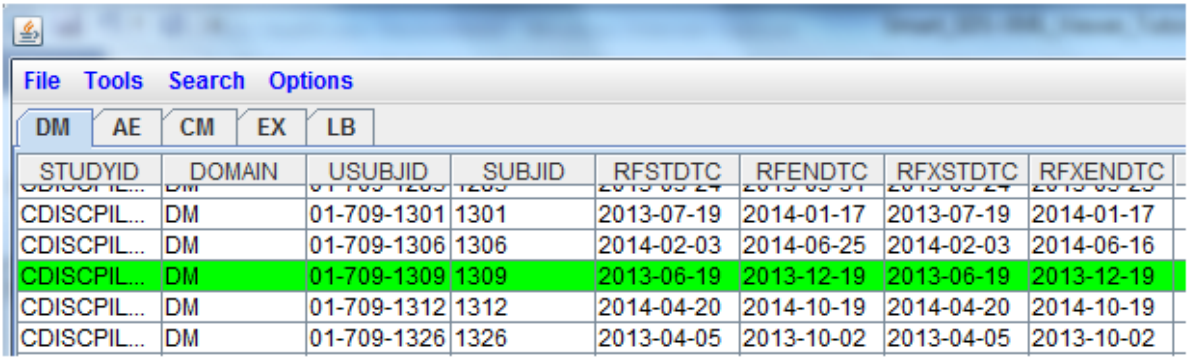
21. Jumping to Corresponding Data in the DM Dataset

When inspecting data in a subject-related dataset, one can always jump to the corresponding record in the DM dataset by selecting a line or cell in the other dataset for that subject, and then either use the menu "Tools – View – Show corresponding DM record", or simply by using the keyboard combination "CTRL-D" ("D" standing for "Demographics"). This automatically opens the "DM" tab (when loaded) and selects and highlights the corresponding record.
For example, when inspecting a lab data point for a specific subject:



STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBO
CDISCPIL...	LB	01-709-1309	190	PHOS	Phosphate	CHEMISTRY	3.6	mg/dL	2.2
CDISCPIL...	LB	01-709-1309	222	PHOS	Phosphate	CHEMISTRY	2.9	mg/dL	2.2
CDISCPIL...	LB	01-709-1309	252	PHOS	Phosphate	CHEMISTRY	3.6	mg/dL	2.2

Using "CTRL-D" automatically immediately shows the corresponding record in the DM dataset:



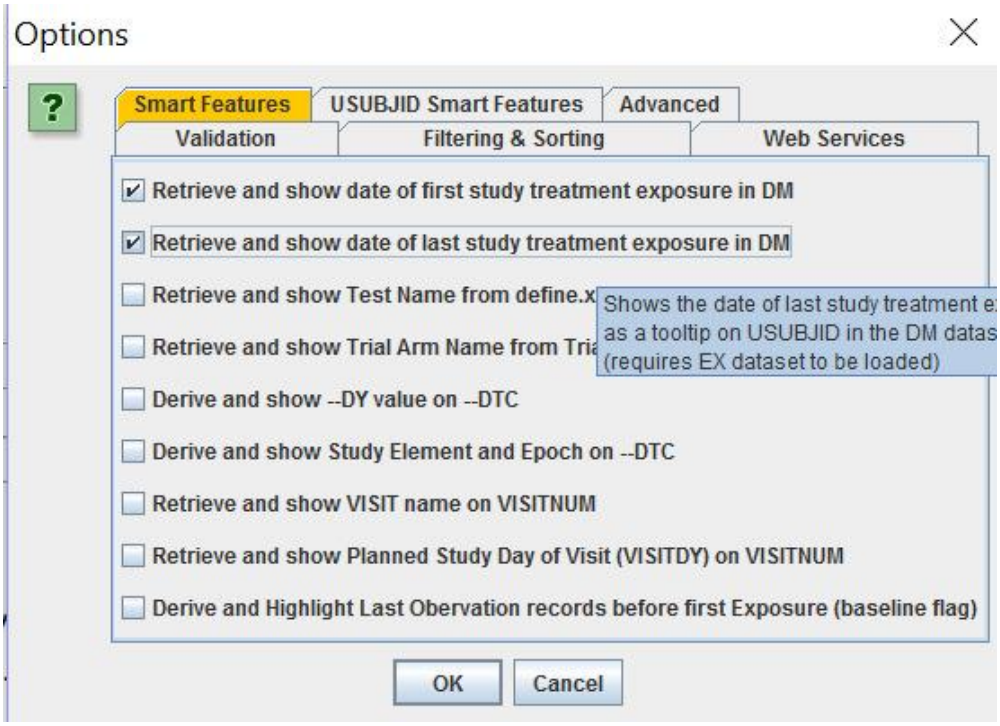
STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC
CDISCPIL...	DM	01-709-1301	1301	2013-07-19	2014-01-17	2013-07-19	2014-01-17
CDISCPIL...	DM	01-709-1306	1306	2014-02-03	2014-06-25	2014-02-03	2014-06-16
CDISCPIL...	DM	01-709-1309	1309	2013-06-19	2013-12-19	2013-06-19	2013-12-19
CDISCPIL...	DM	01-709-1312	1312	2014-04-20	2014-10-19	2014-04-20	2014-10-19
CDISCPIL...	DM	01-709-1326	1326	2013-04-05	2013-10-02	2013-04-05	2013-10-02

One can then toggle between the two datasets either using the mouse and clicking the tab at the top, or using the menu "Tools – View – Show last selected table" or even easier by using the keyboard combination "CTRL-B" ("B" for "back"). So in the above example, subsequently using "CTRL-B" combination toggles between the record in the LB dataset and the record in the DM dataset.

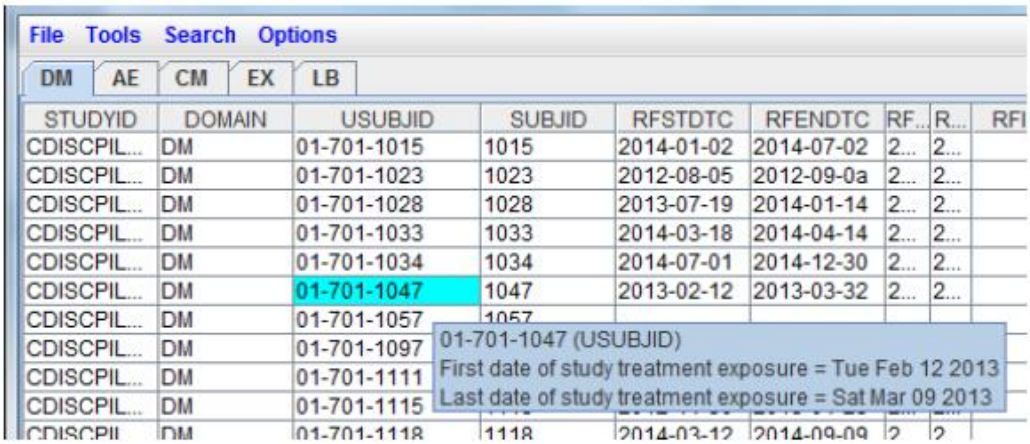
22. Showing Date of First and Last Study Treatment in the DM Dataset

In the last few years, the SDTM DM domain has been overloaded with new variables that contain information in domains such as DM that is already present in other datasets. The reason for this is that the SASViewer is not able to make joins between tables (it also has no instructions to do so). For example, the variables RFXSTDTC (Date/Time of first study treatment) and RFXENDTC (Date/Time of last study treatment) have been added on request of the FDA, although they are already present in the EX dataset (earliest value of EXSTDTC and latest value of EXENDTC). Essentially, duplicating such information is bad practice, as the dates do not correspond, it is unclear which of them is valid (data redundancy error). So RFXSTDC and RFXENDTC were only added as the visualization was not capable of combining both pieces of information.

For making the life of reviewers easier, we have incorporated such features in the "Smart Dataset -XML Viewer". When both the DM and the EX datasets will be loaded, one can use the option "Retrieve and show date of first study treatment exposure in DM" and "Retrieve and show date of last study treatment exposure in DM" in the "Options" tab "Smart features":



When then loading the datasets in the DM, holding the mouse over a cell with a USUBJID, then a tooltip will show up displaying the date/time of first study treatment and the date/time of last study treatment in a user-friendly format, e.g.:



This shows that the usage of "smart" tools like ours can in future allow CDISC to eliminate redundant variables from the model again. For example, the column --DY could be auto-generated by a software tool from RFSTDTC in the DM domain and the variable --DTC in the observation domain. Or, the software tool could automatically calculate LBDY when LBDTC is provided[11].

[11] The usage of a --DY variable when --DTC is present is essentially a violation of the third normal form rule for good database design.

23. On-the-Fly Calculation of Derived Variable Values

As stated before, the SDTM standard has been overloaded with derived variables as the SASViewer and other FDA tools are unable to make joins between tables. So in SDTM there is a large amount of variables that violates the well-known "normal forms" for good database design. Examples are:

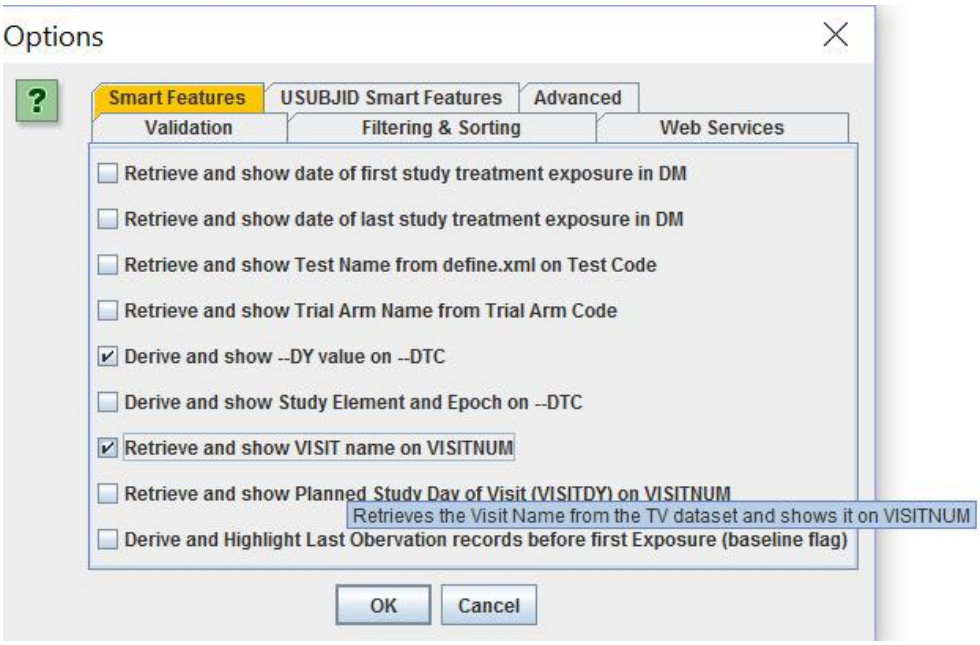
- All --DY variables - derived from --DTC and RFSTDTC
- The VISIT (visit name) - derived from VISITNUM - lookup in table TV
- --TEST - derived from --TESTCD (there is a 1:1 relationship)

In the latter case, --TEST can be obtained by a lookup in the define.xml, or from a database with controlled terminology, e.g. through a web service.

For demonstration purposes, we implemented automated calculation / lookup for:

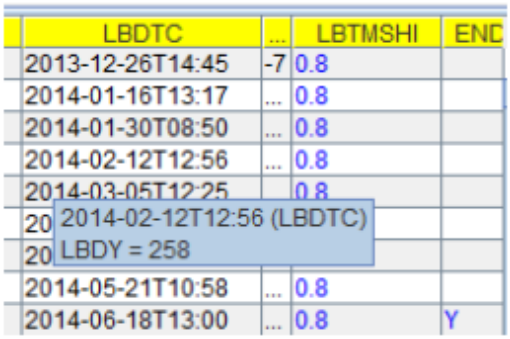
- Age from birthdate (when the latter is present)
- Automated, on-the-fly calculation of --DY variables (except for BRTHDTC) from RFSTDTC in the DM table (when it is loaded)
- Automated, on-the-fly calculation of VISIT (visit name) from VISITNUM, when the TV table is present

To do so, one should check the corresponding checkboxes "Check age from birthdate and reference start date", "Show --DY value on --DTC" and/or "Get and show VISIT name on VISITNUM", in the "Options" panel before starting generating the tables.



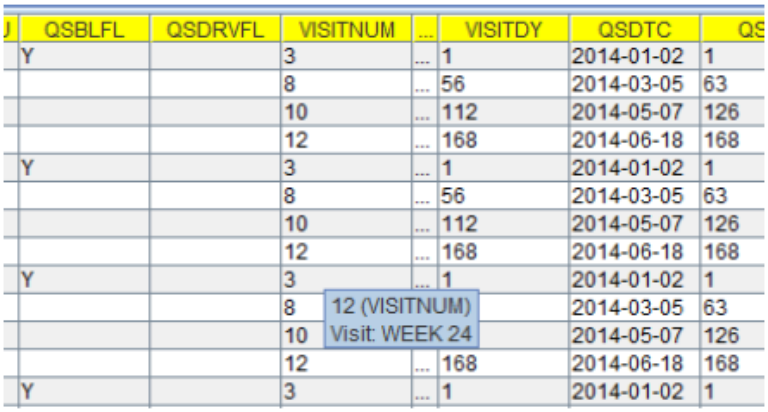
When then loading is then executed, the tooltip on the respective cells also shows the "on-the-fly calculated" or "looked up" value, independently from whether the --DY variable and VISIT is present.

For example:



This shows the calculated LBDY from LBDTC and RFSTDTC – so **NOT** taken from the column LBDY[12].

In case "Get and show VISIT name on VISITNUM", a join is made with the TV table and the "visit name" (VISIT) is looked up there, and shown as a tooltip:

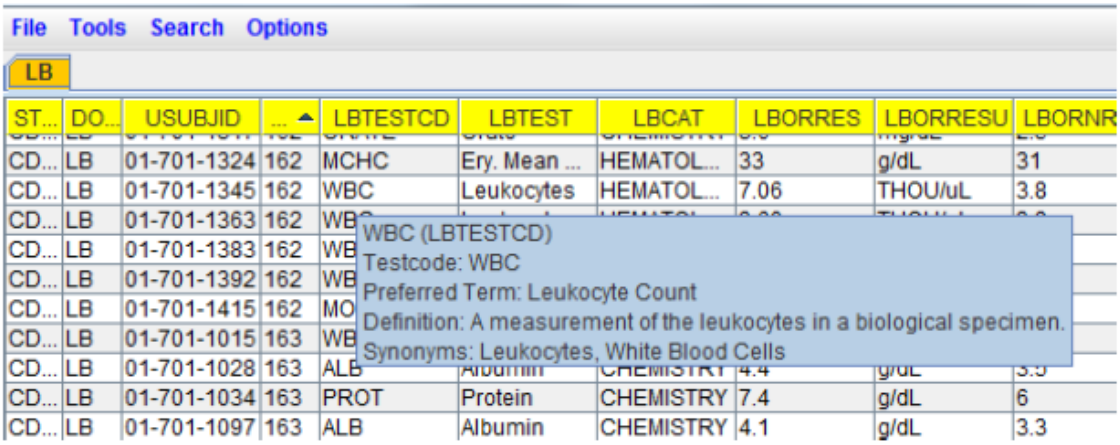


During testing these features, we found that in several SDTM submissions, the --DY values were not always correctly calculated, and/or the VISIT (name) was not always correctly listed. Without our features however, there was no way to validate whether the --DY and/or VISIT value is correct. Also OpenCDISC does not perform this validation.

Essentially, also "--TEST" is a derived or lookup variable. It can be regarded as a lookup from the SDTM-IG and/or CDISC controlled terminology, as there is a 1:1 relation between –TESTCD and --TEST. Recently, we also implemented a "fast lookup" for --TESTCD so that when the user hovers the mouse over a --TESTCD cell, a web service makes a lookup in a remote database (can be somewhere on the internet) and looks up the corresponding "test name" (--TEST), and

[12] Essentially this means that in future, all or most of the --DY variables can be removed from the SDTM model, as calculating them is done by the viewer or other tool anyway. This would seriously contribute to the data quality of SDTM submissions (removal of redundancy that is currently leading to possible data errors).

displays it as a tooltip. Alternatively, the --TEST column could be automatically generated and added by a web service. A screenshot how this looks like is given here.

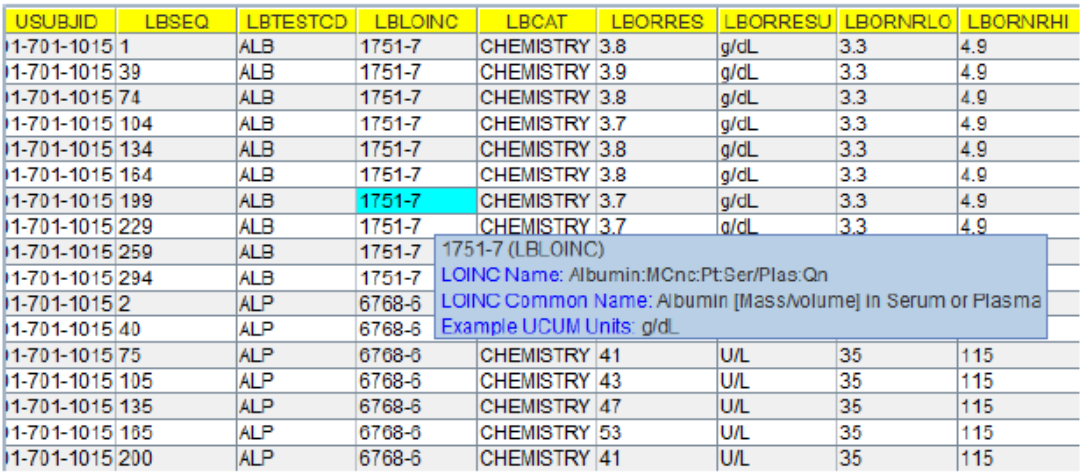


ST...	DO...	USUBJID	...	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNR
CD...	LB	01-701-1324	162	MCHC	Ery. Mean ...	HEMATOL...	33	g/dL	31
CD...	LB	01-701-1345	162	WBC	Leukocytes	HEMATOL...	7.06	THOU/uL	3.8
CD...	LB	01-701-1363	162	WB					
CD...	LB	01-701-1383	162	WB					
CD...	LB	01-701-1392	162	WB					
CD...	LB	01-701-1415	162	MO					
CD...	LB	01-701-1015	163	WB					
CD...	LB	01-701-1028	163	ALB					
CD...	LB	01-701-1034	163	PROT	Protein	CHEMISTRY	7.4	g/dL	6
CD...	LB	01-701-1097	163	ALB	Albumin	CHEMISTRY	4.1	g/dL	3.3

For further details, please see the separate document "Smart Dataset-XML Viewer Web Services ".

Remark that such features can be implemented for any coded values, also from external code lists and dictionaries (SNOMED-CT, RxNorm, …). For example, three of our students have recently developed a web service for looking up the details of LOINC codes. This enables to show the extended information of any LBLOINC value (if present - we think it should) as a tooltip. This feature further contributes to the integration with electronic health records, as in the latter (e.g. HL7-CDA) the use of LOINC codes is usually mandatory.

Here is a screenshot on how this web service has been implemented in the Smart Dataset-XML Viewer:



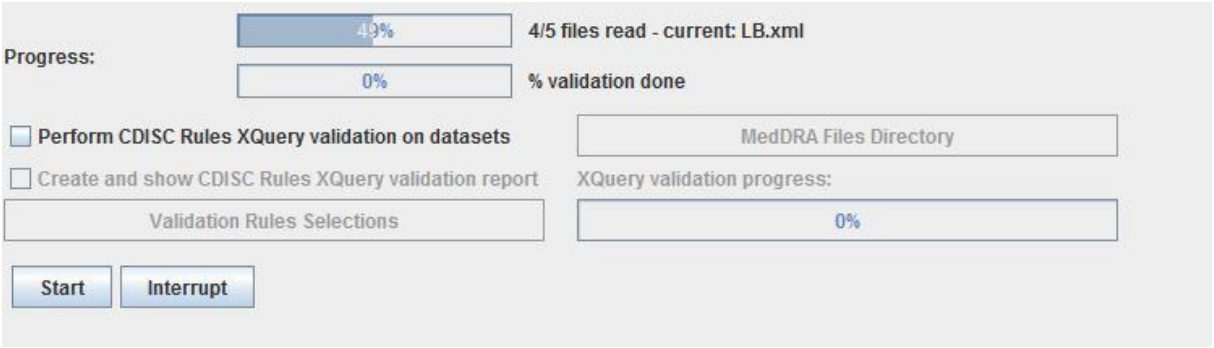
USUBJID	LBSEQ	LBTESTCD	LBLOINC	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI
01-701-1015	1	ALB	1751-7	CHEMISTRY	3.8	g/dL	3.3	4.9
01-701-1015	39	ALB	1751-7	CHEMISTRY	3.9	g/dL	3.3	4.9
01-701-1015	74	ALB	1751-7	CHEMISTRY	3.8	g/dL	3.3	4.9
01-701-1015	104	ALB	1751-7	CHEMISTRY	3.7	g/dL	3.3	4.9
01-701-1015	134	ALB	1751-7	CHEMISTRY	3.8	g/dL	3.3	4.9
01-701-1015	164	ALB	1751-7	CHEMISTRY	3.8	g/dL	3.3	4.9
01-701-1015	199	ALB	1751-7	CHEMISTRY	3.7	g/dL	3.3	4.9
01-701-1015	229	ALB	1751-7	CHEMISTRY	3.7	g/dL	3.3	4.9
01-701-1015	259	ALB	1751-7					
01-701-1015	294	ALB	1751-7					
01-701-1015	2	ALP	6768-6					
01-701-1015	40	ALP	6768-6					
01-701-1015	75	ALP	6768-6	CHEMISTRY	41	U/L	35	115
01-701-1015	105	ALP	6768-6	CHEMISTRY	43	U/L	35	115
01-701-1015	135	ALP	6768-6	CHEMISTRY	47	U/L	35	115
01-701-1015	165	ALP	6768-6	CHEMISTRY	53	U/L	35	115
01-701-1015	200	ALP	6768-6	CHEMISTRY	41	U/L	35	115

For further details and other currently implemented web services (such as from the National Library of Medicine), see the separate document "Smart Dataset-XML Viewer Web Services".

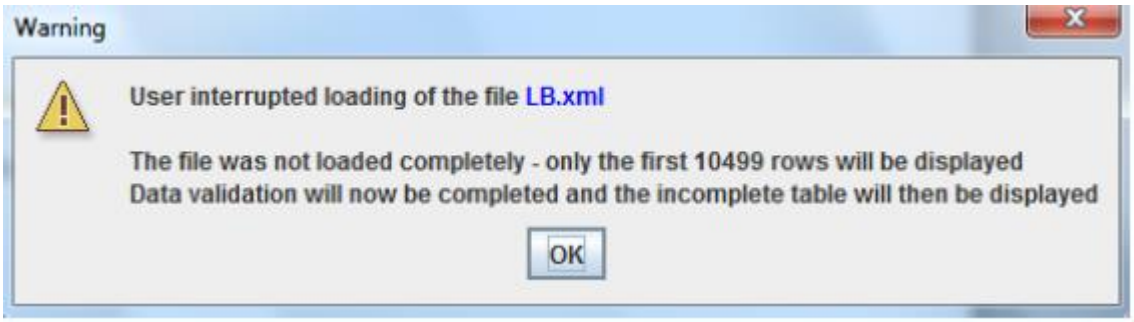
24. Additional Features: Interrupting the Loading Process

Loading very large datasets with millions of data points can take considerable time. As described before, it often makes sense to first make selections based on test codes and of subjects (subpopulations) and then load the data for this subpopulation only.

If one sees that loading a large dataset is taking too much time, and one decides to first define a subpopulation of subjects and/or tests, one can always interrupt the current loading process by clicking the "Interrupt" button. This button becomes enabled once the "Start" button has been clicked, and becomes disables once all datasets have been loaded:



When the "Interrupt" button is clicked, the system will complete loading the row that was busy loading, and then display the following message:

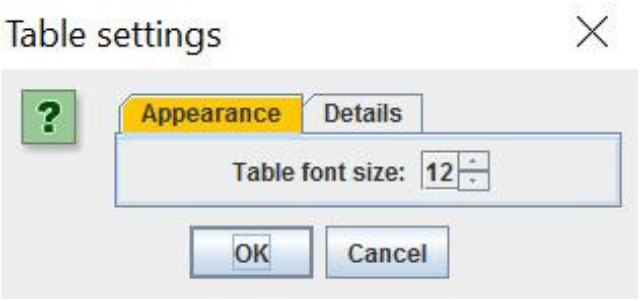


After clicking "OK", the tables are displayed as usual, but the "LB" table will be incomplete.

Datasets that were after "LB.xml" in the list of files will not be loaded, and no table will be generated for them.

25. Setting the Font Size for the Tables

The font size for the tables can be changed by using the menu "Options – Table font size". A dialog is then displayed allowing to set the requested font size using a spinner:

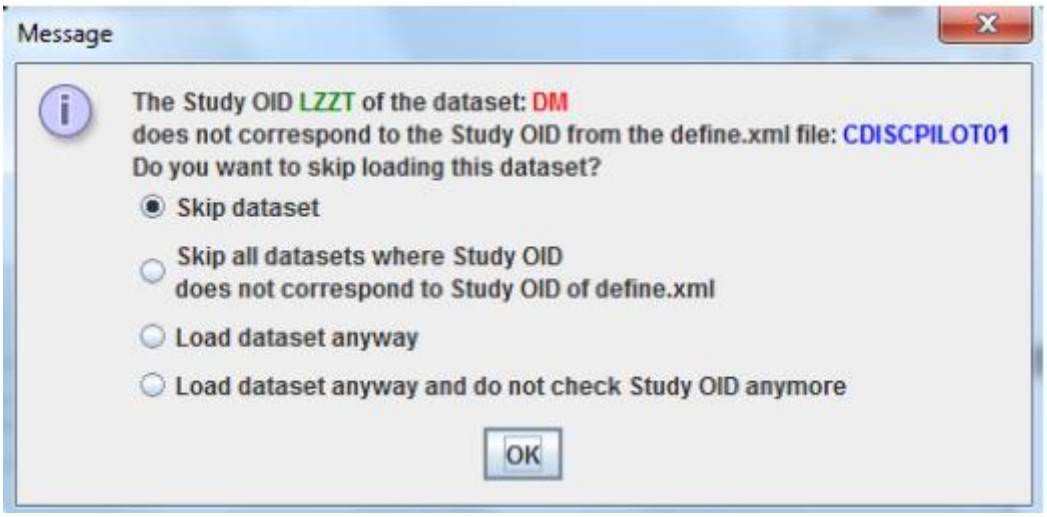


A font size between 8 and 16 can then be selected. The default font size is 12. Upon clicking, the tables are immediately updated using the new font size.

26. Validation: Checking the Study OID versus the Study OID Given in the define.xml File

Both the define.xml files (Dataset-XML files) contain information regarding which study is represented by the data. In the define.xml, the Study OID is given by the "OID" attribute on the "Study" element (Study/@OID). In the Dataset -XML files, the Study OID is given by the "StudyOID" attribute on either the "ReferenceData" element (in case of study design datasets) or the "StudyOID" attribute on the "ClinicalData" elements (in case of subject-related data).

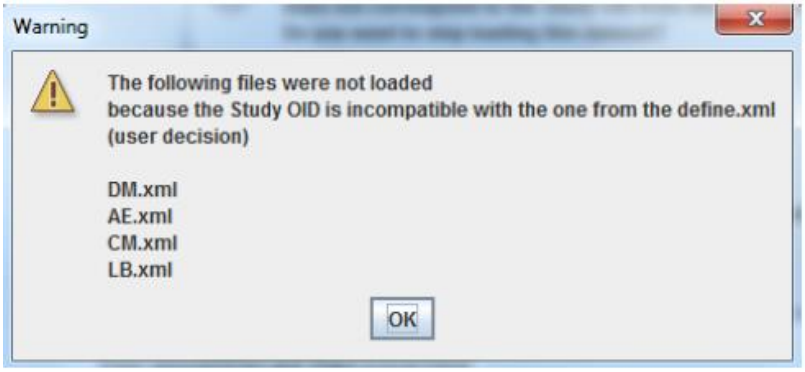
By default, the software checks whether the Study OID given in each dataset corresponds to the one in the define.xml. If there is a mismatch, the following dialogues is displayed:



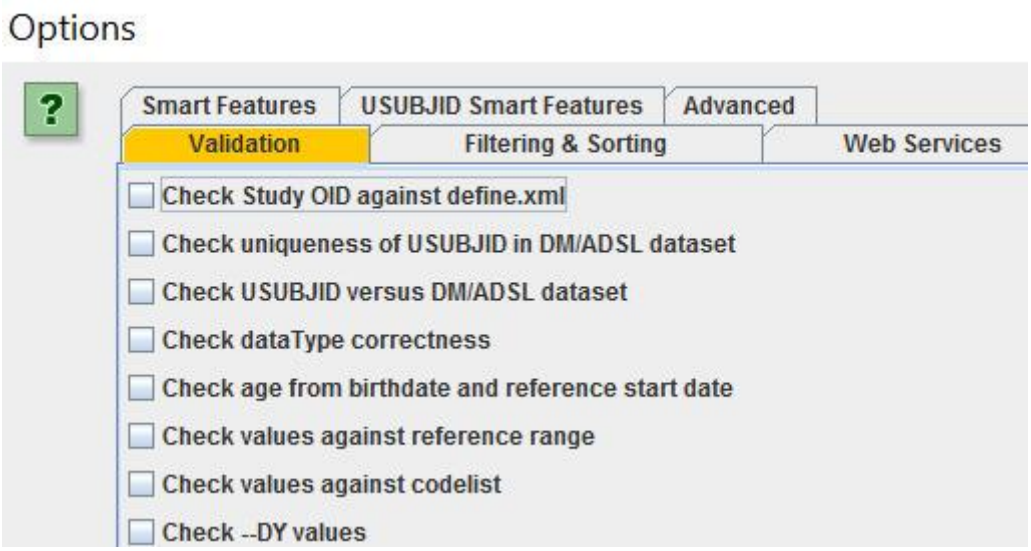
Four possibilities are then presented:

- Skip loading this dataset (user recognizes that this dataset does not correspond to the one for which the metadata is defined in the define.xml)
- Skip all datasets where Study OID does not correspond to Study OID of define.xml (skip all the datasets that seem not to belong to the submission represented by the define.xml)
- Load dataset anyway (user believes that the dataset corresponds to the submission the define.xml represents)
- Load dataset anyway and do not check Study OID anymore (same as "Load dataset anyway" but the following datasets are not checked for a corresponding Study OID) anymore.

In case the option "Skip dataset" is selected, the dataset will not be loaded. The system will continue with the next dataset and check the Study OID again. In case the option "Skip all datasets where Study OID does not correspond to Study OID of define.xml", a message will be generated at the end of the process, displaying a list of datasets that were skipped due to mismatch between the Study OIDs in the dataset and the define.xml:



In case the option "Load dataset anyway" the mismatch is ignored, the dataset is loaded, and the next dataset is again checked for correspondence between Study OID of the define.xml and that of the dataset. In case the option "Load dataset anyway and do not check Study OID anymore" is selected, the dataset is loaded, and further checking of the Study OID is disabled. Remark that this also unchecks the checkbox "Check Study OID against define.xml" in the "Options" dialogue:



If for the next loading, one would want to execute Study OID checking again, one needs to check the box "Check Study OID against define.xml" again.

27. Validation: Using OpenCDISC within the Smart Dataset-XML Viewer

to be implemented

28. Internationalization

One of the disadvantages of the SAS XPORT format was that only US-ASCII characters were allowed. As SDTM is also used in countries that use other character sets (e.g. Japan) this was a serious limitation for the use of SDTM in such countries and for submissions to local regulatory authorities.

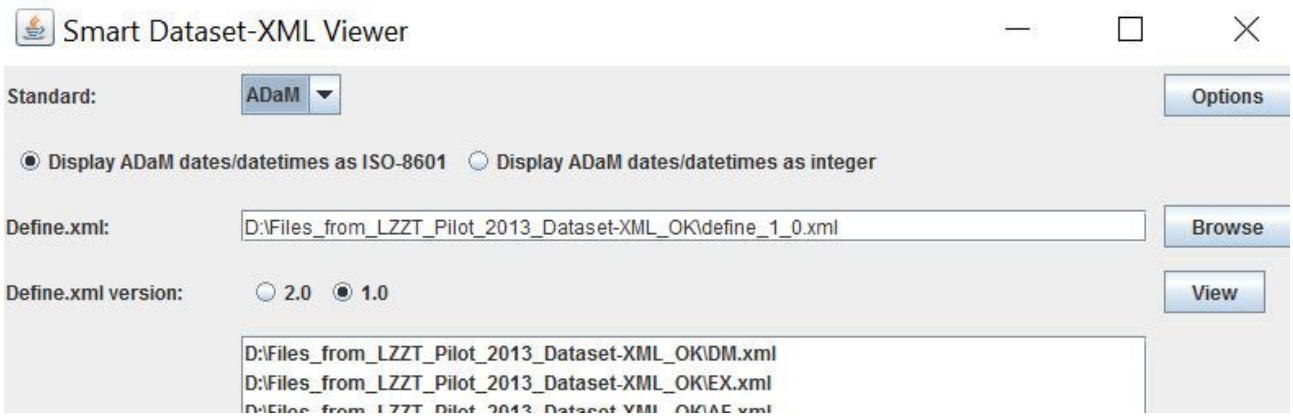
XML does not have such a limitation, and neither does Dataset-XML.

The "Smart Dataset-XML Viewer" supports Dataset-XML files that have values that have non-ASCII characters[13]. We currently have tested this on files using French (e.g. êèé), Norwegian (e.g. æøå), German (e.g. üäöß) and Japanese characters. For example:

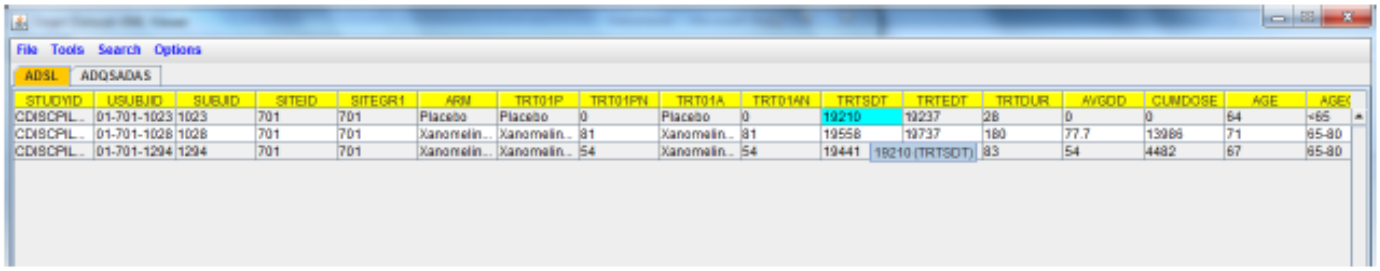
SEX	RACE
F	Weiß
M	WHITE

29. Display ADaM Dates as ISO-8601 in the Smart Dataset-XML Viewer

When the user selects "ADaM" as the standard, a choice between displaying ADaM dates (which are integers) as integer or as ISO-8601 is displayed:

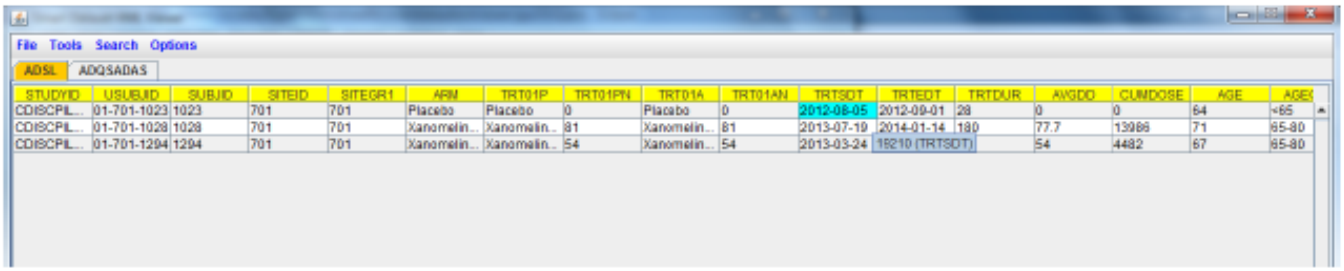


If the user selects to display ADaM dates as integers, the result is:



If the user selects to display ADaM dates as ISO-8601, the result is:

[13] We encourage the use of UTF-8 encoding for XML files in general



This is just the display, internally, ADaM dates are still stored as integers.

Appendix 1: Starting the Smart Dataset-XML Viewer from other programs and systems

The Smart Dataset -XML viewer has been designed in such a way that it can easily be integrated with or started from other programs. There is a clear API and the software can also be started with a parameter list in order to prepopulate some fields in the GUI and preset some settings.

The following shows how to start the Smart Dataset -XML viewer from another Java-based program. For this, the file "Smart_ Dataset -XML_Viewer.jar" needs to be in the classpath.

A simple example is:

```
import edu.fhjoanneum.ehealth.smartdatasetxmlviewer.*;
…
GUI gui = new GUI(); // sets up the Smart Dataset-XML Viewer GUI
// set the location of the define.xml file (case Windows system)
gui.setDefineFile (new
File("C:\\CDISC_SDTM_XML_Standard\\Files_from_MSG_XML\\define.xml"));
// set the define.xml version - default is "1.0" so only needed if you want to use a define.xml
1.0 file gui.setDefineVersion("1.0");
// sets the CDISC model for the viewer. The default is "SDTM, the two other possibilities are
"SEND" and "ADaM". Attention: case sensitive!
gui.setCDISCModel("SDTM");
// you can now define a set of Dataset-XML files to appear in the list in the GUI.
// This is done using a vector of "File" objects
Vector<File> sdsXMLFiles = new Vector<File>();
sdsXMLFiles.add(new
File("C:\\CDISC_SDTM_XML_Standard\\Files_from_MSG_XML\\dm.xml"));
sdsXMLFiles.add(new
File("C:\\CDISC_SDTM_XML_Standard\\Files_from_MSG_XML\\ae.xml"));
…
// and pass these to the viewer
gui.setSDSXMLFiles(sdsXMLFiles);
```

```
// get a JFrame and display it
JFrame f = gui.getFrame();
f.setVisible(true);
```

That's it!

The second possibility is to execute the Smart Dataset-XML Viewer from another program, script, … even from line command. In that case you can add a number of arguments with values. The following parameters and values are available:

Parameters	Values
-df filelocation	passes the location of the define.xml file
-f filelist	passes a list of Dataset-XML files
-d1	passes the information that the define.xml file is of version 1.0
-m	sets the CDISC model to be used. Default is "SDTM". Other allowed values are "ADaM" and "SEND" (case sensitive)
-dt	sets that the data is "typed", i.e. "typed" ItemData is used (default is untyped)

So for example, you would start the program (e.g. within a script by):

```
# set the classpath
set CLASSPATH=C:\Smart_Dataset-XML_Viewer\Smart_Dataset-XML_Viewer.jar;C\vtd-
xml.jar;.\log4j-1.2.13.jar
# run the software
java -Xms256M -Xmx1024M -cp %CLASSPATH%
com.xml4pharma.smartsdsxmlviewer.gui.GUI
-df C:\CDISC_SDTM_XML_Standard\Files_from_MSG_XML\define.xml
-f C:\CDISC_SDTM_XML_Standard\Files_from_MSG_XML\dm.xml
C:\CDISC_SDTM_XML_Standard\Files_from_MSG_XML\ae.xml
-m SDTM -d1 -dt
```

The "-Xms256M" and "-Xmx1024M" set the minimum and maximum amount of computer memory the software is allowed to claim. It is advised to claim no more than 60% of the available physical memory. So if you have 2GB memory machine, it is advised to not go beyond -Xmx1228M.

The "-cp %CLASSPATH%" applies the class path defined in the first line.

The second line (-df) sets the location of the define.xml file. The two following lines pass two files to be processed by the viewer, in the case the dm.xml and ae.xml files. In the last line, we pass the information that the SDTM model needs to be used (but that is the default anyway) and that the files use "typed" ItemData, through the "-dt" argument.

Although we speak of "lines" in the explanation, everything starting from "java" should go into a single line, as it is a single command.

Appendix 2: Perspectives for the Future

Due to its open-source nature, the Dataset-XML Viewer allows to develop very interesting extensions. For example, one could add a module that looks up additional information through a web service about what a specific medication (CMTRT) exactly is and what active ingredients it contains. E.g. in the above example, holding the mouse over a cell with CMTRT=LOPID could lead to triggering a web service (e.g. querying RXList), and then display additional information about LOPID as a tooltip or using a popup window.

At the moment of writing, eHealth students at the Applied University FH Joanneum in Graz have recently developed such a web service for LOINC, which could then be used for obtaining additional information of specific LBLOINC, EGLOINC and VSLOINC values.

Also, this could e.g. be used to look up the details of a SNOMED-CT code. In future, it would also e.g. enable to connect to systems that have "artificial intelligence".

References

[1] Blog – Working on and with CDISC Standards	http://cdiscguru.blogspot.com/2016/08/why-lobxfl-should-not-be-in-sdtm.html
[2] Download "Smart Dataset-XML Viewer"	https://sourceforge.net/projects/smart-sds-xml-viewer/files/
[3] JRE Oracle	http://www.oracle.com/