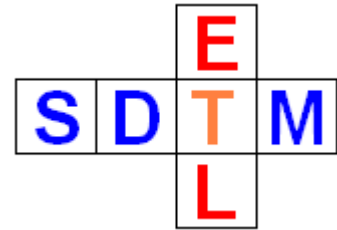


SDTM-ETL 5.0 User Manual and Tutorial: Working with the DC Domain (Demographics as Collected)



Author: Jozef Aerts, XML4Pharma

Last update: 2025-04-31

Table of contents:

Introduction.....	1
Availability of the DC template	1
Adding DC to the study	1
Future Development.....	4

Introduction

After years of discussions, it has recently become clear that FDA and CDISC have finally agreed how to have "multiple screenings" treated. This is now explained in a formal document from CDISC, which is the "draft specification" of the next SDTM-IG, which will succeed SDTMIG-3.4. This documents describes how the "DC" domain, usually named "Demographics as Collected", should be used for the case of multiple screenings¹.

Therefore, we developed a template for DC based on the recent CDISC publications.

IMPORTANT REMARK: DC should only be used for the case of "multiple screenings" and when the SDTM datasets are intended to be submitted to a regulatory authority that explicitly recommends the use of DC (FDA).

Do NOT use DC when you do not have multiple screenings or do not intend to submit the information for these to a regulatory authority.

Availability of the DC template

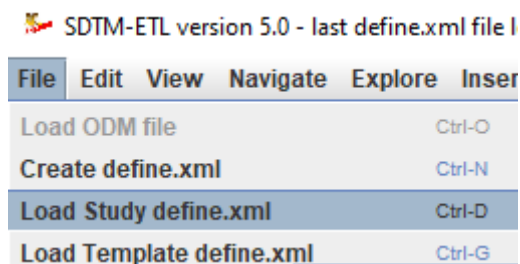
If your distribution does not already contain the template file for the DC domain (it is named "define_template_DC_domain_SDTMIG.xml" in the folders "define_2_1" and "define_2_0"), just send us a mail, and we will send it to you or make a download available.

Remark that the template is currently meant for SDTMIG-3.4, but can as well be used for earlier versions of the SDTMIG. Also remark that DC is a non-standard domain, so in the case of Define-XML 2.1, it has the attribute 'def:IsNonStandard="Yes"'.

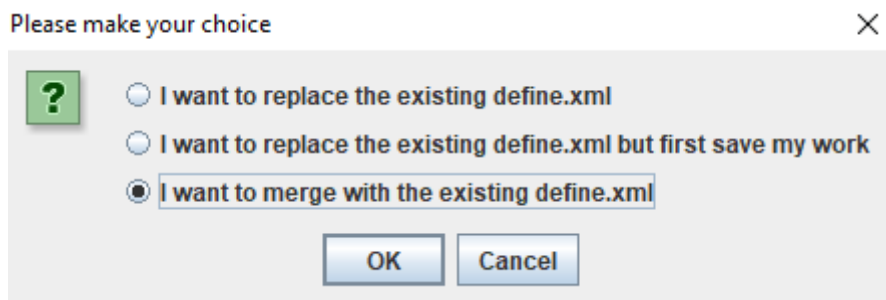
Adding DC to the study

After having loaded the template for a specific SDTMIG version, or having loaded earlier mappings for your study, use the menu "File - Load Study define.xml". Do NOT use the menu "File - Load Template define.xml", as this will clear the already loaded SDTM table.

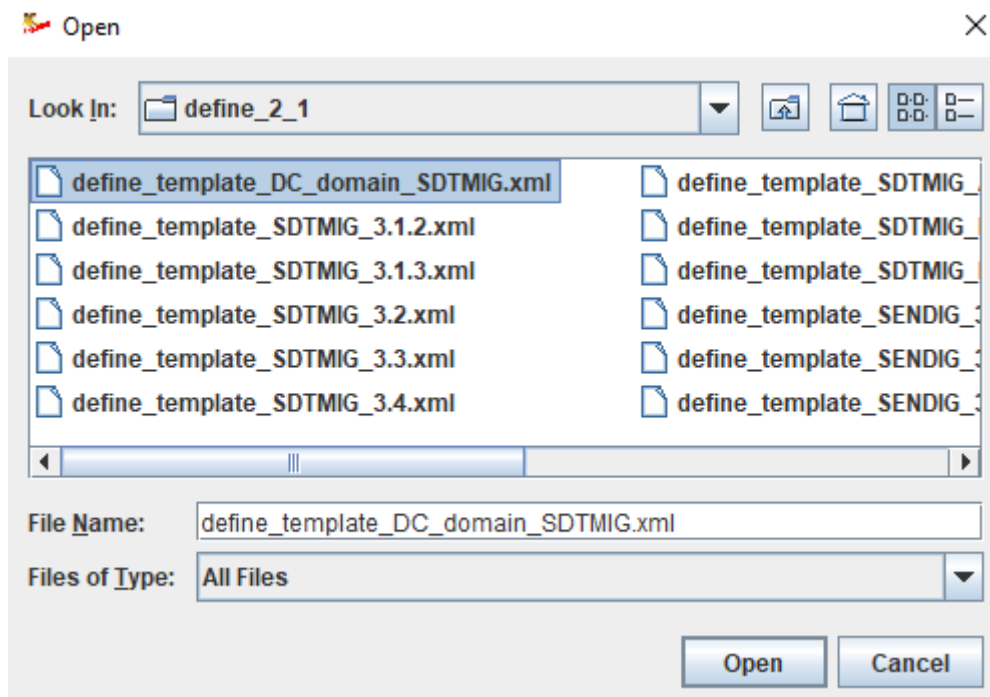
¹ In the draft version of the next SDTM-IG, the DC domain got the designation "Demographics for Multiple Participations (DC)"



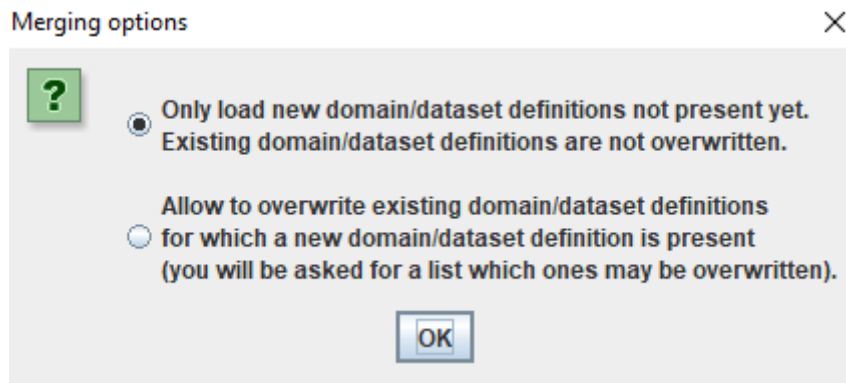
and select "I want to merge with the existing define.xml":



Then navigate to either "define_2_1" folder (or "define_2_0" folder if you still use Define-XML 2.0) of your distribution, and select "define_template_DC_domain_SDTMIG.xml"

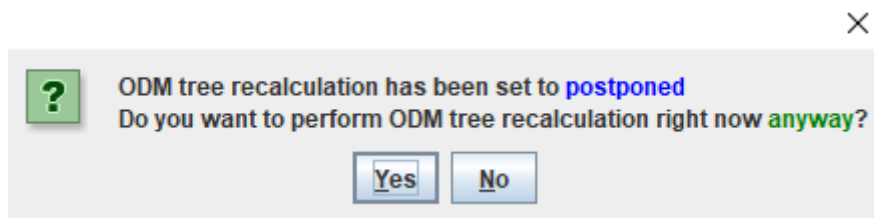


After clicking "OK", the following dialog appears after a few seconds:

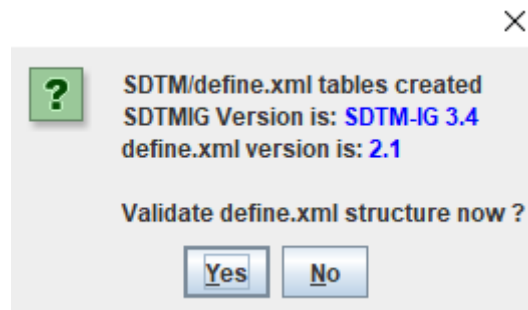


As one has not loaded DC before, one can just accept "Only load new domains/dataset definitions" and then proceed.

It will then take 1-2 minutes before everything is reassembled, after which the following dialog appears:



where one can click "No", as no new study-specific information (mappings) was added anyway. It is then requested whether one wants to validate the define.xml structure (again):



This will usually not be necessary. The result then finally is e.g.:

TU	STUDYID	DOMAIN	USUBJID	TU.TUSEQ	TU.TUGRPID	TU.TL
TR	STUDYID	DOMAIN	USUBJID	TR.TRSEQ	TR.TRGRPID	TR.TF
RS	STUDYID	DOMAIN	USUBJID	RS.RSSEQ	RS.RSGRPID	RS.RS
VS	STUDYID	DOMAIN	USUBJID	VS.VSSEQ	VS.VSGRPID	VS.VS
FA	STUDYID	DOMAIN	USUBJID	FA.FASEQ	FA.FAGRPID	FA.FA
SR	STUDYID	DOMAIN	USUBJID	SR.SRSEQ	SR.SRGRPID	SR.SR
RELREC	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELT
SUPQUAL	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM
DC	STUDYID	DOMAIN	USUBJID	SUBJID	DC.FOCID	DC.D
CES:GLOBAL	REFSTARTDATE					
CES:DM	STUDYID	DOMAIN	USUBJID	SUBJID	DM.RFSTDTC	DM.R
CES:SU	STUDYID	DOMAIN	USUBJID	SU.SUSEQ	SU.SUGRPID	SU.SI
CES:AE	STUDYID	DOMAIN	USUBJID	AE.AESEQ	AE.AEGRPID	AE.AE
CES:LBBL	STUDYID	DOMAIN	USUBJID	LB.LBSEQ	LB.LBGRPID	LB.LE
CES:LBCH	STUDYID	DOMAIN	USUBJID	LB.LBSEQ	LB.LBGRPID	LB.LE

where one sees that "DC" has been added to the existing templates.

Then in order to start developing mappings, just like for all other domains", one still need to make a "study-specific instance, which can be done by "drag-and-drop" to the bottom (or use the menu "Edit - Copy Domain/Dataset" followed by "Edit - Paste Domain/Dataset". Very often, one will want to have the "study-specific" DC instance immediately after the "study-specific" DM instance, e.g.:

SR	STUDYID	DOMAIN	USUBJID	SR.SRSEQ	SR.SRGRPID	SR.SRREFID
RELREC	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE
SUPQUAL	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM
DC	STUDYID	DOMAIN	USUBJID	SUBJID	DC.FOCID	DC.DCSEQ
CES:GLOBAL	REFSTARTDATE					
CES:DM	STUDYID	DOMAIN	USUBJID	SUBJID	DM.RFSTDTC	DM.RFENDTC
CES:DC	STUDYID	DOMAIN	USUBJID	SUBJID	DC.FOCID	DC.DCSEQ
CES:SU	STUDYID	DOMAIN	USUBJID	SU.SUSEQ	SU.SUGRPID	SU.SUSPID
CES:AE	STUDYID	DOMAIN	USUBJID	AE.AESEQ	AE.AEGRPID	AE.AEREFID
CES:LBBL	STUDYID	DOMAIN	USUBJID	LB.LBSEQ	LB.LBGRPID	LB.LBREFID
CES:LBCH	STUDYID	DOMAIN	USUBJID	LB.LBSEQ	LB.LBGRPID	LB.LBREFID
CES:VS	STUDYID	DOMAIN	USUBJID	VS.VSSEQ	VS.VSGRPID	VS.VSSPID

One also sees that for DC "SUBJID" is a "looping variable" (by the cyan border on the cell), i.e. one will usually have "One record per SUBJID per USUBJID".

One can now start developing the mappings as usual.

Future Development

"DC", with as label "Demographics for Multiple Participations (DC)" will be a formal, standard CDISC domain in the next SDTMIG.

For the moment, it must however still be treated as a "non-standard" domain.

The advantage of the draft of the next SDTMIG however is that it is now finally clear how DC has to be structured, and that is how we have implemented it.

In the next version of SDTM-ETL (v.5.1), we will have a more explicit way of adding DC to the study - the current way of doing is only a temporary solution.

For example, when additionally loading DC is SDTM-ETL 5.1, a warning dialog will appear stating that DC may only be used for the case of "multiple screenings" and only when the regulatory authority recommends the use of DC for the case of "multiple screenings".