

SDTM-ETL 5.0 User Manual and Tutorial: CodeLists and CodeList mapping

Author: Jozef Aerts, XML4Pharma

Last update: 2025-03-02

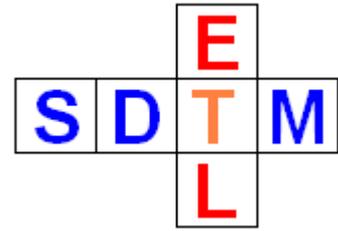


Table of contents:

Introduction.....	1
Problems with the mapping wizard.....	1
ODM item is free text but needs to to be mapped to an SDTM variable that is under controlled terminology	3

Introduction

CodeList mapping is one of the major problems encountered when generating SDTM/SEND datasets from sources like ODM. Although it is always advised to use CDISC Controlled Terminology (CDISC-CT) already during development of the study design, this is not always done (in some cases people using their own codelists that are not compatible with the SDTM/SEND codelists), or cannot always be done (e.g. as there is no CDISC-CT for that item).

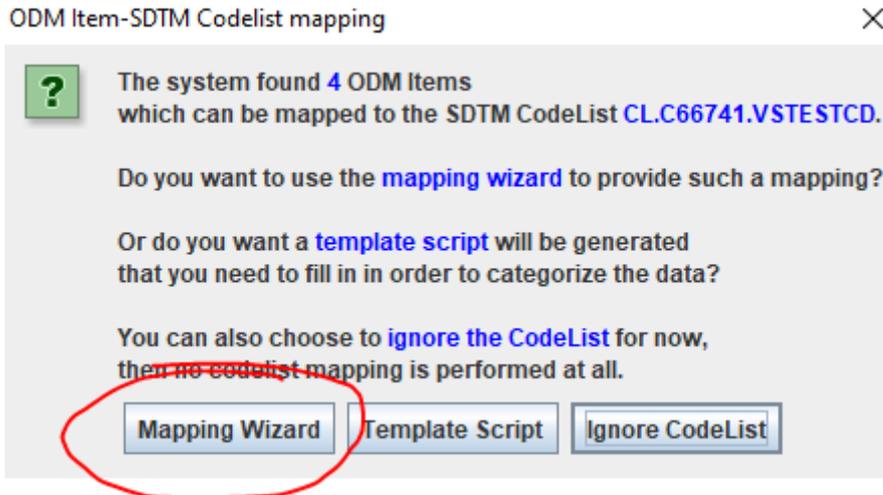
CodeList mapping is already treated in several of the tutorials. In this document we concentrate on common problems that have been reported by users when doing codelist mapping. As such, this document will regularly be updated.

Problems with the mapping wizard

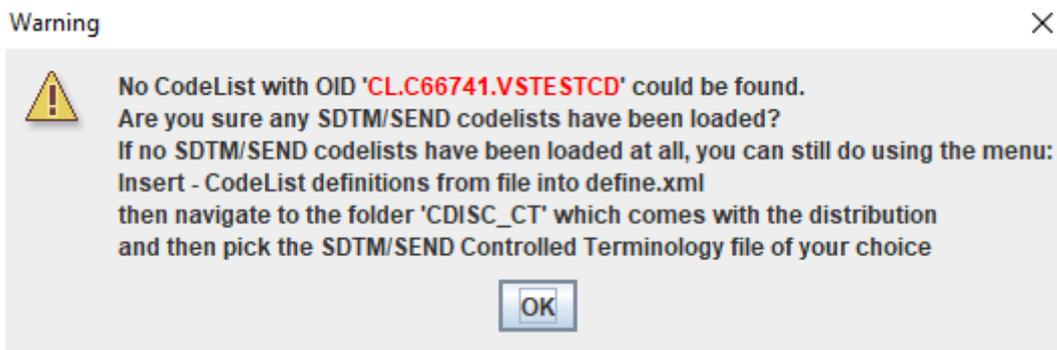
The ideal case is when for the mapped item, such as an --TESTCD, there is an associated codelist for the ODM item, and an associated codelist for the SDTM/SEND variable, or that the the user has selected one or more items (in the latter case using "Generalize for all items", and then selecting the ones needed) from the ODM tree that need to be mapped to the codelist of the SDTM variable.

For example, there are 4 items selected from the ODM tree, like "height", "weight", systolic blood pressure" and "diastolic blood pressure".

After drag-and-drop, the following dialog then shows up:

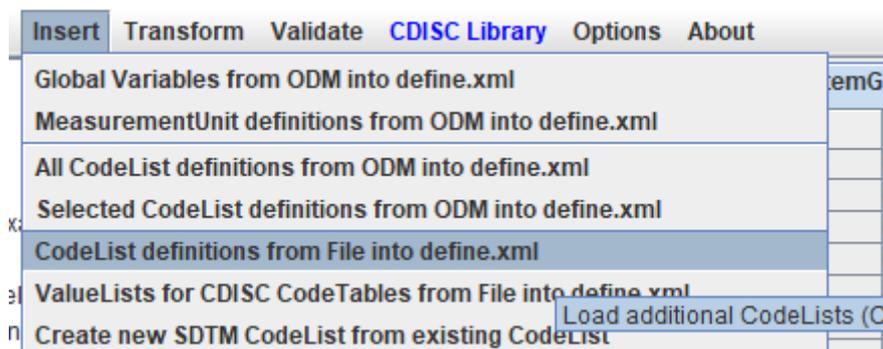


Something that has been reported by a user is, that when attempting to use the mapping wizard, nothing happens, or the following message appears:



The reason is that the user forgot to select a CDISC controlled terminology version when loading a define.xml template (for a specific SDTM-IG or SEND-IG version), or did decide to not select one yet. In such a case, the SDTM variable (like VSTESTCD) is associated with a codelist, but the codelist itself is not present (yet), and so the mapping wizard cannot be started, as the codes for the SDTM side are missing.

The user can then still load Controlled Terminology, using Insert - CodeList definitions from file into define.xml:



e.g.:

Insert CodeList definitions from ODM 1.3 file



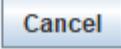
 Overwrite existing CodeLists with same OID

File with CodeLists in ODM Format

D:\SDTM-ETL_4_0\CDISC_CT\SDTM_Terminology_2021-12-17.xml 

Store as new CodeList version: 2021-12-17

For Standard: SDTM 

Message



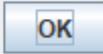
 All imported Controlled Terminology will be registered as a new CT version **2021-12-17** for the Standard **SDTM**



A message will then be shown that the codelists have been loaded.



984 CodeLists added



After that, one can again attempt to start a codelist mapping using the wizard.

ODM item is free text but needs to to be mapped to an SDTM variable that is under controlled terminology

This is close to a worst-case scenario, but ... happens....

Suppose for example that the CRF has the following questions that have "free text" answers:

- what was the lab test?
- result of the lab test:
- unit for the lab test:

or for an Adverse Event form:

- what is the severity of the adverse event?

SDTM/SEND is all about categorization of collected data + a lot of (mostly unnecessary¹) derivations. This means that for many variables, the mapping will contain if - else if - else constructs. This is also the case when free text results need to be categorized, as the corresponding SDTM/SEND variable is under controlled terminology (CDISC-CT).

Mapping from free text to CDISC-CT can be very tricky, and if done, will need serious validation back to the source record.

Let us take the example of the (bad practice) open question "what is the severity of the adverse event?". The result will need to go into SDTM AESEV (Adverse Event Severity) which is under controlled terminology (codelist CL.C66769.AESEV: "Severity/Intensity Scale for Adverse Events") with the allowed values "MILD", "MODERATE" and "SEVERE".

One can then write a script similar to the following:

```
$T1 = xpath(.....); # path to the ODM variable
$T2 = upper-case($T1);
$AE.AESEV = "";
if(contains($T2,'MILD') {
    $AE.AESEV = 'MILD';
} elseif (contains($T2,'MODERATE') {
    $AE.AESEV = 'MODERATE';
} elseif (contains($T2,'SEVERE') {
    $AE.AESEV = 'SEVERE';
} else {
    $AE.AESEV = $T1;
}
```

In the first line (usually generated by drag-and-drop) the path to the source item in the ODM is provided. The free text is then made upper case in the second line in the temporary variable \$T2. An initial value of blank is then assigned to AESEV.

In the next lines, using an if-elseif-else construct, "MILD" is assigned when the free text contains the word "mild" (whenever lower case or upper case), "MODERATE" is assigned when the free text contains the word "moderate", and "SEVERE" is assigned when the free text contains the word "severe". If none of these is found, the original text is put in AESEV.

If the latter happens, one will of course get a validation error, which one will need to explain the reviewers guide.

Again, such categorization from free text can be very tricky.

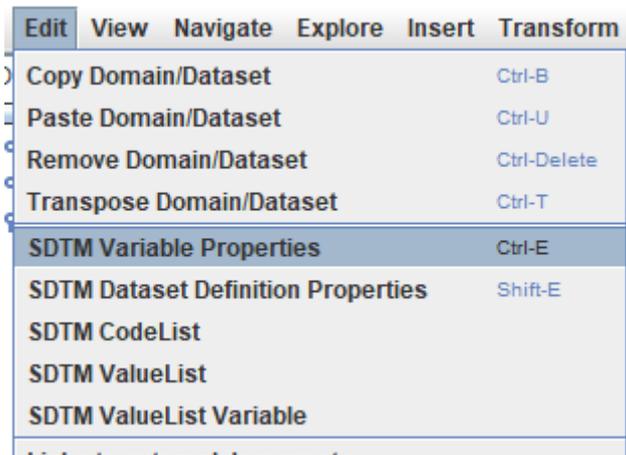
If you do so, it is also not a bad idea to submit the original free text as a "supplemental qualifier variable" in SUPPAE.

Another possibility, which one will also need to explain in the reviewers guide, is to remove the codelist association from the SDTM/SEND variable.

The way to do this in SDTM-ETL is the following:

- select AESEV in the study-specific instance in the table on the right.
- use the menu "Edit - SDTM Variable Properties"

¹ A typical example are the -DY variables. These have been introduced on request of the FDA, as the tools of the FDA are not capable to derive them automatically.



- the following window will then appear:

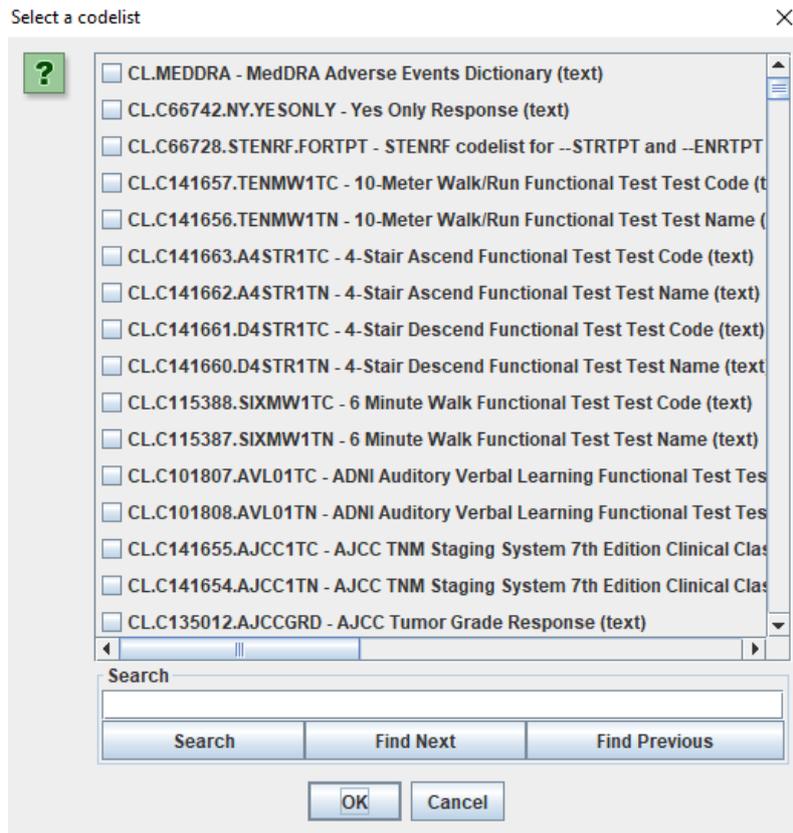
Edit Properties for SDTM Variable AE.AESEV

OID:	AE.AESEV
Name:	AESEV
SASFieldName:	AESEV
Data type:	text
Current Length:	80
<input type="checkbox"/> New Length:	80
Current Significant Digits:	
<input type="checkbox"/> New Significant Digits:	-1
Current Role:	Record Qualifier
<input type="checkbox"/> New Role	Record Qualifier
Current Role CodeList:	
<input type="checkbox"/> New Role CodeList	CL.MEDDRA - MedDRA Adverse Events Dictionary (text)
Current Origin:	NONE DEFINED YET
<input type="checkbox"/> Edit Origin:	Edit
Comment:	
External document for comment	
Current CodeList	CL.C66769.AESEV
<input type="checkbox"/> New CodeList:	Select CodeList
Description:	Severity/Intensity
current def:DisplayFormat:	
<input type="checkbox"/> New def:DisplayFormat:	
current ValueList OID:	NO VALUELIST ASSIGNED

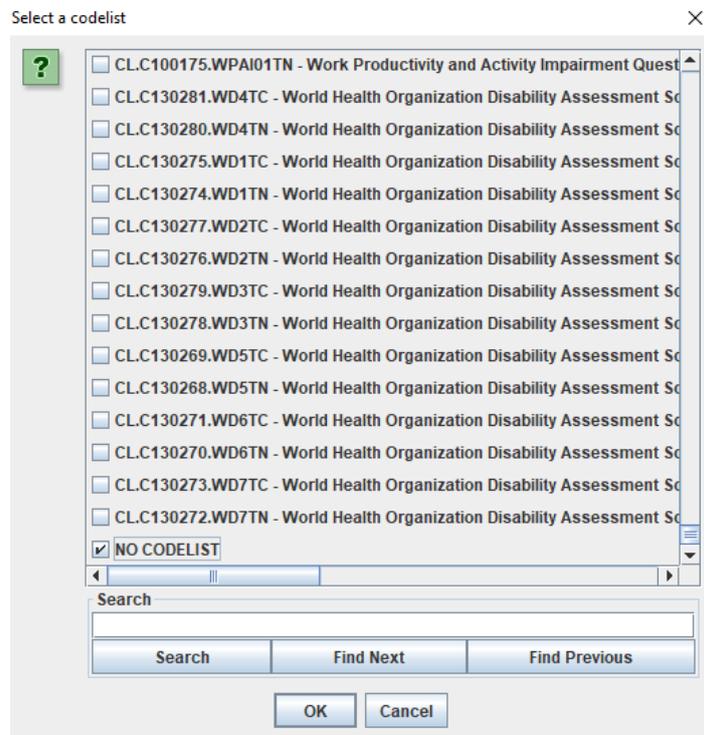
- check the checkbox "New CodeList". The "Select CodeList" button becomes enabled.

Current CodeList	CL.C66769.AESEV
<input checked="" type="checkbox"/> New CodeList:	Select CodeList
Description:	Severity/Intensity

- now click the button "Select CodeList". The following window is displayed:



- now scroll to the bottom and select "NO CODELIST"

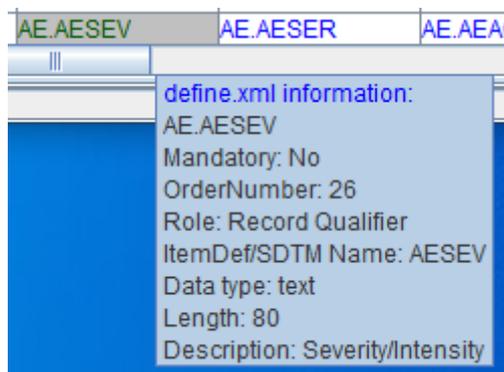


- click "OK". The window closes and in the prior window "NO CODELIST" is shown for "Current CodeList"



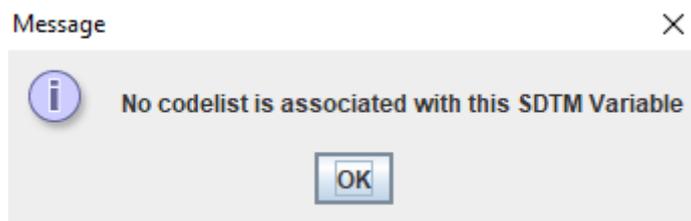
- Also click OK for this window.

The main window is again shown and when one hovers the mouse over AESEV, the following tooltip is displayed:



with no mention of a codelist anymore.

Also, when using the menu "View - SDTM associated CodeList", after having selected the AESEV cell in the table, the following message appears:



And one can now just drag-and-drop the item with the free text to AESEV, so that it will contain the original, free text.

Please remind that this is NOT good practice, and you will need to document this change (and the thus generated validation errors) in the reviewers guide, and explain why you removed the codelist from the variable.

This is surely also one of the things you will want to discuss in a "pre-submission meeting" with the regulatory authority, as it may be that this is not an acceptable solution for the reviewer assigned to your submission.