Adding mappings for the Vital Signs domain

After loading and validating the ODM file with metadata, your load existing define.xml with mappings, or create a new define.xml (using menu “File – Create define.xml”).

In our case, where we already created the mappings for DM (Demographics) and EC (Exposure as collected), the right part of our screen will look like:

- We now create a “study-specific instance” of the VS domain just by drag-and-drop of the “VS” row (from the template) to the bottom. The following dialog shows up:

  Copy Domain VS

  ![Copy Domain VS dialog](image)

  Which we all accept by clicking “OK”. The result is:

  ![Study-specific instance of the VS domain](image)

  We see that mappings have already automatically be created for STUDYID, DOMAIN, USUBJID, VSSEQ – these fields in the table have a grey color. We also see that the field for VSTESTCD has a blue border, meaning that it is a “looping variable”, i.e. a record will be created for each vital signs test code found in the source. If we double-click the first cell “MyStudy:VS”, the following dialog is displayed:
Also displaying the looping structure, which essentially is “One record per subject per vital signs test code”. This can be visualized by clicking the “Validate” button:

We will keep this for now, and come back later and extend this when necessary.

It is always a good idea to first provide the mapping for the looping variables. So we will develop the mapping for VSTESTCD first.

**Generating the mapping for VSTESTCD**

If we look into the ODM tree, we find:
and see that vital signs are only collected in the “Treatment” visit, which is a repeating visit. Further expanding the tree view leads to:

We observe that some of the fields have a cyan background color, meaning that they do already have an “SDTM annotation”, i.e. the developers of the study design already annotated these with SDTM information, stating where the data point will later be used in SDTM. For example, if we hold the mouse of “Systolic Blood Pressure”, a tooltip is displayed:

We can now start developing the mapping for VSTESTCD.
First drag-and-drop from one of the appropriate entries from the ODM tree to the cell “VS.VSTESTCD”. If we e.g. take “Systolic Blood Pressure” and drag-and-drop it, the following dialog shows up:

Now, we do not want to import the value of the data point from the set of clinical data, we want to import its data point code, and will then map that to SDTM. So, we check the checkbox “Import XPath expression for another ItemData attribute” (second checkbox). This leads to:

What is happening here? When getting the clinical data, the software will pick up the identifier of the data point (in our case it is “IT.SYSBP”) and will use that for mapping to VSTESTCD.

Now, we do have of course more vital signs collected than just the systolic blood pressure, and we also want to get the others too. In order to get them, click the checkbox “Generalize for all Items”.

This will select all the data points in the group “Vital Signs”, including “Vital Signs Date”, “Vital Signs” time and so on. These two are not vital signs of course, so we will want to exclude them, or alternatively, we only want to include the data point codes that really represent.

This can easily be done using one of the buttons “Except for …” or “Only for …”. If we use “Only for …” the following dialog is displayed:

We than select only these items that really represent a vital signs test code, i.e.:
We do not select “Vital Signs Date”, “Vital Signs Time”, “Vital Signs DateTime”, and “Weight Units” as these do not represent vital signs test codes.

Clicking OK now leads to:

Stating that we will have 5 different vital signs test codes.

If we had the case that vital signs would also have been collected in other visits (study-events) than “treatment”, we should select the checkbox “Generalize for all StudyEvents” and use the button “Except for …” or “Only for …” and then select the visits in which vital signs have been collected. This will then take care that these visits are also taken into account.

Clicking “OK” then leads to a new dialog:
Stating that there is an SDTM codelist for VTESTCD in SDTM, and that we can map our 5 test codes in our ODM to the SDTM codelist. Using the wizard for this will simplify our work a lot, so we click “Yes”, leading to:

For each of the entries on the left (our ODM codes) we can now easily map to an SDTM “vitals signs test code”, by a simple selection on the right. For example, for “Systolic Blood Pressure”:

In many cases, just typing the first character of the desired test code, allows to find the desired test code. This will e.g. lead to:
For the last row “MISSING VALUE”, we should also make a choice. Usually, one will select the “NULL” value, i.e. scroll to the bottom of the list and select the last entry which is the empty entry:

With the result:

The case of a “missing value” should however never occur in practice, as VSTESTCD is a “looping variable” and we iterate only over data points that represent one of these 5 test codes.

Clicking “OK” then generates the mapping script:
If needed, one can then still adapt the mapping script manually. In many cases, this will not be necessary. One then accepts the mapping script (which is then stored in the underlying define.xml) by clicking the “OK” button.

One can now do a first test, by using the menu “Transform – Generate Transformation (XSLT) Code for SAS-XPT”. This leads to:

Most of the EDC vendors (>90%) use the “classic” ODM export, with “non-typed ItemData”. Only a few use the “Typed ItemData” flavor of ODM. If you have doubts, ask your EDC vendor which of both is exported, or just try both here – if you have the wrong one, the output will simply be empty.

The following dialog is displayed after clicking “OK”:

---

1 If one does not want to use SAS-XPT as the output format, but use the new Dataset-XML format, one of course uses the menu “Transform – Generate Transformation (XSLT) Code for Dataset-XML”
Showing you the generated XSLT. If you want to use it for offline SDTM generation, you can save it now. You can also skip this step by using the menu “Options – Settings” and checking “Skip display of generated XSLT”.

Then click the button “Execute Transformation (XSLT) Code”, leading to:
Now select an input file with ODM clinical data (as exported from your EDC system) by using the first “Browse” button. For example:
In most cases, you will not need to check the checkbox “Metadata in separate ODM file”. This will only be the case e.g. when you need “decoded” codelist values, or the “description” or “question text” of a data point.

If you already want to generate SAS-XPT datasets, also check the checkbox “Save Result SDTM tables as SAS XPORT file”, and then select a directory where these need to be written to. During testing however, this will usually not be necessary, as the generated SDTM tables will be displayed
by the software itself.

All the other options will be explained later.

Clicking “Execute Transformation on Clinical Data” starts the transformation. In case of very complicated mappings and large amounts of data this can take a few minutes.

In our case, the results display after a few seconds:

![Table showing the data points collected for the first subject (1001).](image)

Showing that 17 vital signs data points have been collected for the first subject (1001). Remark that “VSSEQ” has been generated automatically, this does not need to be programmed by the user.

**Continuing with VSORRES**

Once the mapping for VSTESTCD has been developed (which can be a bit more challenging when the tests are divided over different forms), generating the mapping for VSORRES is pretty easy.

Select the cell “VS.VSORRES” and then drag-and-drop one of the tree nodes representing a vital signs test to the cell “VS.VSORRES”. Again, a dialog is displayed.
This time we must import the value of the data point, not the code. So, we let the checkbox “Import … for ItemData Value attribute …” selected. We also see that the system remembers the “generalization” with the 5 inclusions, so there is nothing to be changed there.

Also notice the line “ODM ItemDef Length” (near the bottom) stating that in the ODM the value has been declared as of a maximum length of 3, whereas in SDTM the value from the template is 80. So, not a bad idea to check the checkbox “Set SDTM Variable Length to ODM ItemDef Length” which will set the value for the maximum length in SDTM to 3 too:

Clicking “OK” leads to:

After accepting the mapping script, testing on our clinical data leads to:
We also see that something is not entirely ok. In the ODM, it was defined that the maximum length (as text) is 3, but we see that there are values that are of length 4. Although this can be automatically corrected when generating the SAS-XPT files, it is not a bad idea to set the maximum length for VSORRES to 4 already now in the underlying define.xml.

In order to do so, select the cell VS.VSORRES and then use the menu “Edit – SDTM Variable Properties”. This leads to:

Checking the box “New Length” allows us to set the new maximum length to 4:

Clicking “OK” updates the variable properties.
When we then hold the mouse over the cell “VS.VSTESTCD” we see:
And see that the maximum length has indeed been updated.

Generating the mapping for VSTEST

Generating the mapping for VSTEST is almost identical to the one for VSTESTCD. Drag-and-drop one of the vital sign test entries from the ODM tree to the VS.VSTEST cell, then select to use the “ItemOID” from the clinical data:

![Import ItemDef: Systolic Blood Pressure - for SDTM Variable VS.VSTEST](image)

The system still knows about our 5 different vital sign test, so we can just continue, leading to:

![ODM Item-SDTM Codelist mapping](image)

Accepting with “Yes” then leads to:
Which can then easily be completed to:

Which then automatically generates the mapping script:

And after testing using Transform – Generate Transformation …”, leads to:
Nice! We are making good progress …

Generating other mappings for Vital Signs - VSORRESU

If we now look at our table again, we see:

That we still need to add the information for “VSORRESU” (original result units) and for VSSTRES, VSTRES (standardized results, character and numeric) and VSSTRESU (standardized result units). We will start with VSORRESU (original units).

We see at the ODM side that for systolic and diastolic blood pressure, the units are not provided. The reason is that the value is always mmHg (UCUM notation\(^2\): mm[Hg]). For “Weight” however, the field in the EDC system was essentially a free text field, as we can see from the ODM:

And when we ask for the associated codelist using the menu “View – Item CodeList Details” we get … nothing. We can also see that there was no codelist, by selecting “Weight Units” and inspecting the line at the bottom:

So, what were the reported weight units? In order to see them, use the menu “View – ODM ClinicalData”. This leads to:

\(^2\) Unfortunately, CDISC and FDA do not allow units in UCUM notation yet, which would simplify many things, such as automated unit conversions.
And clicking “View ODM Clinical Data” then shows us a table:

And we see that “kg” has always been used. If we had found a multitude of different units here, we would have had to read some extra code, e.g. taking into account the different ways of writing “kilogram” and “pounds”.

Inspecting the collected data is always a good idea when the item was a free text item and needs to be standardized.

In order to provide the mapping for VSORRESU, again drag-and-drop one of the vital signs entries from the ODM to the VS.VSORRESU cell, and select to import the “ItemOID”, leading to:
Continued by the dialog:

And when accepting with “Yes”, the mapping wizard is displayed:

Which we can easily complete, as typing the first character already leads us to the neighbourhood of what we need. The result will be:
Suppose now that we want to read the value for “weight unit” from the ODM, and add the value here. We can easily do so by adding some extra scripting. We first drag-and-drop “Weight Units” to “VSORRESU”. As there is already a mapping present, the system asks us:
In our case, it is a good idea to append the mapping to the top of our script, so we select “Append to existing mapping at top”. We can also give a new name to the new variable that is then created, e.g.:

After clicking “OK”, we get:
However, this time, we only want to import the value for “Weight Units”, so do not want to “generalize for all items” (within the group), so we must unselect that checkbox:

And after clicking “OK” we get:

Which will generate a script that we will need to complete. Clicking “OK” leads to:

Such a “template” script is extremely useful when one needs to “categorize” collected data that comes as free text into categories or coded values, as is the case here. In our case however, it is only about “weight”, so we can remove that template script again, as we will just “pick up” the value as it was collected. This leads to:
The first set of lines takes the collected value for VSORRESU, the other lines “hardcode” the value, as it was not collected, but e.g. prescribed by the protocol.

There is now still one line we need to correct, which is the line in the “if” statements that is about weight units. We change it into:

Saying that for “Weight”, we will take the collected value, and not the hardcoded value.

Running the scripts then leads to:
Later, we will also learn how to create “Value lists” for VSORRES for the individual tests. See the tutorial “Working with the WhereClause in define.xml 2.0“. With SDTM-ETL, generating value lists is extremely easy, whereas it is a nightmare with some of our competitor software offerings.

Creating the mapping for VSPOS

VSPOS (Vital Signs Position) is a “permissible” variable, meaning that one can omit it when no data was collected for it, which is the case here. We will however still add it to demonstrate the use of “Origin” in define.xml.

Suppose that it was stated in the protocol that the vital sign tests blood pressure and pulse must be taken in sitting position. For “weight”, “height” (which is not collected here) this of course doesn’t make sense. So we will create a small mapping script in which we state that in the former case, the value of VSPOS is “SITTING”, whereas it is empty (“null”) in the latter case.

Let us first have a look at the allowed values for VSPOS. In order to do this, select the cell VS.VSPOS and use the menu “View - SDTM Associated CodeList“.

The following dialog is displayed:
It also provides the maximum length for the values in the list, which can be used to adapt the maximum length for the variable VSPOS itself.

In SDTM-ETL, one can always reuse variables in the same row that come before the current variable and for which a mapping exists. As VSPOS comes after VSTESTCD, we can easily reuse it in a set of one or more “if” statements.

So we can write:

```plaintext
$VS.VSPOS = '';
if($VS.VSTESTCD='DIABP' or $VS.VSTESTCD='SYSBP' or $VS.VSTESTCD='PULSE') {
    $VS.VSPOS = 'SITTING';
} else {
    $VS.VSPOS = '';
}
```

leading to:
If the “position” was collected, one can of course generate the mapping using simple drag-and-drop.

As the subject’s position for vital signs measurement was defined in the protocol, we should add this information in the (underlying) define.xml. We can do this at the variable level (VSPOS) or at the value level, i.e. individually for SYSBP, DIABP, WEIGHT, ... The mechanism is however essentially the same. Just for the sake of the tutorial, we will define the “origin” at the variable level here.

Select the VS.VSPOS cell, and then use the menu “Edit – SDTM Variable Properties”. The following dialog is displayed:

One sees that no “origin” has been defined yet. So, check the checkbox “Edit Origin”, and then click the button “Edit” on the right of it. The following dialog is displayed:
If the subject’s position would have been collected, we would set the “origin” to “CRF”, and then add the page numbers in the annotated CRF. If the subject’s position would have been assigned by us (not recommended), we would set the origin to “Assigned”. In our case however, as it comes from the protocol, we select the radiobutton “Protocol”:

One also sees that the fields about page numbers are then disabled. Click “OK” to confirm. This leads to the prior dialog:
We also see that somehow, the maximum length was set to 380, which does not make sense. We set it to 7 which is the length of the word “SITTING”.

We can also add a comment to the “origin” statement, and even make a reference to an external document if our comment is too long for keeping it in the define.xml. Just let us add a short statement that one can find the instruction about taking the vital signs in a sitting position on page 37 of the protocol. For example:

Now click “OK” to confirm.

When we move the mouse over the cell “VS.VSPOS” a tooltip is displayed:
Mappings for VSSTRESN, VSSTRESC and VSSTRESU

In our case, all collected values are already in standard units, so there is not much to do. Also, as VSSTRESN, VSSTRESC and VSSTRESU come after VSORRES and VSORRESU, we can reuse the values. For example for VSSTRESN, we can simply state:

And similar for VSSTRESC and VSSTRESU (which can be copied from VSORRESU).

Just for the sake of the tutorial, suppose that we also had collected “frame size” (VTESTCD=FRMSIZE) which has values “SMALL”, “MEDIUM” and “LARGE” and thus no units, and as it is not a numeric value, VSSTRESN (“Numeric Result/Finding in Standard Units”), we could easily write the script as:

With a similar script for VSSTRESU

In SDTM-ETL, we usually do not have to care whether the value is numeric or text as long as we do not perform calculations. This is taken care of by the software itself.

Just as an example, let us suppose that some of the weight data was collected in (US) pounds. We know that 1 pound (UCUM notation: [lb_av]). We know that 1 pound = 0.4536 kg. If CDISC would allow UCUM notation, this conversion factor and the conversion itself could be done by a RESTful
web service. However, this is not the case yet, and we need to program this manually. For VSSTRESN the script could then look like:

```javascript
# For WEIGHT in pounds, a conversion is necessary
# For FRMSIZE, there is no numeric value
if($VS.VSTESTCD == 'FRMSIZE') {
    $VS.VSTRESN = ''; 
} elseif($VS.VSTESTCD == 'WEIGHT' and $VS.VSORRESU == 'pounds') {
    $VS.VSTRESN = 0.4536 * number($VS.VSORRES);
} else {
    $VS.VSTRESN = $VS.VSORRES;
}
```

Remark the use of the “number()” function to ensure that the value is treated as numeric. Also remark that in “if”- statements, the comparator “==” as well as “=” can be used. Both are treated equal.

The result when executing the transformations then is:

![Image of transformation result]

Generating the mappings for VSDTC and VSDY

For VSDTC (Vital Signs Date/Time of collection), we can simply drag-and-drop from the ODM item “Vital Signs DateTime” or by concatenating the values of “Vital Signs Date” and “Vital Signs Time”. Also here, first displaying the collected data is often of help.

For the case, we need to concatenate date and time, we first drag-and-drop from “Vital Signs Date” to VS.VSDTC, and manually change the variable name in e.g. “VISDATE”:
Then we drag-and-drop from “Vital Signs Time” to VS.VSDTC. As there is already a mapping present, the systems asks us what to do. We state that we want to append it to the existing mapping with another variable name, e.g. “VISTIME”:

Leading to:

```plaintext
# Mapping using ODM element ItemData with ItemOID IT.VSDATE
$VSDATE = xpath://StudyEventData[@StudyEventID='SE.VISITA']//FormData[@FormOID='FORM.VS']
# Mapping using ODM element ItemData with ItemOID IT.VSTIM
$VISTIME = xpath://StudyEventData[@StudyEventID='SE.VISITA']//FormData[@FormOID='FORM.VS']
```

In case we are not sure that the time part has always been collected, we can take care of that in the script. For example:

```plaintext
# Mapping using ODM element ItemData with ItemOID IT.VSDATE
$VSDATE = xpath://StudyEventData[@StudyEventID='SE.VISITA']//FormData[@FormOID='FORM.VS']
# Mapping using ODM element ItemData with ItemOID IT.VSTIM
$VISTIME = xpath://StudyEventData[@StudyEventID='SE.VISITA']//FormData[@FormOID='FORM.VS']

if ($VISTIME !== '') {
    $VS.VSDTC = createdatetime($VSDATE, $VISTIME);
    # alternatively, we can use: $VS.VSDTC = concat($VSDATE, 'T', $VISTIME);
} else {
    $VS.VSDTC = $VSDATE;
}
```

We can either use the “createdatetime” function (which is a build-in function) or perform the concatenation ourselves using the “concat” function.
Our result is:

(remark that we moved the VSTESTCD column somewhat to the right for a better visualization of the results).

Once we have VSDTC, the derivation of VSDY is very simple. Here is the script:

```
# difference in days between date/time of collection and reference
$DAYS = datediff($Vs.VSDTC,$RFSTDTC);
# for non-negative differences, add 1
# as VSDY is not allowed to be 0
if($DAYS >= 0) {
  $Vs.VSDY = $DAYS + 1;
} else {
  $Vs.VSDY = $DAYS;
}
```

We use the function “datediff” to obtain the number of days between the date/time of collection and the reference start date RFSTDTC, which was previously defined as a global variable (see the tutorial “Generating mappings for the SV domain & creating global variables for reuse”), and then add 1 when the number of days is non-negative. The reason is that the FDA considers the reference start date as day 1, and there is no day 0 according to the FDA.

Generating mappings for VISITNUM and VISIT

VISITNUM (Visit Number) is an expected variable (integer datatype) in SDTM, VISIT a permissible variable. There is no requirement that visit numbers are subsequent. They just have to be unique for each planned visit, and need to be described in the “trial visits” (TV) domain. There is also no
requirement that they are positive numbers. So, we might e.g. assign negative numbers to all visits before the reference start date and positive numbers for the visits at or after the reference date. Unscheduled visits should be given a floating point number. For example, if there are two unscheduled visits between visit 2 and 3, these can be given visit number 2.1 and 2.2.

In our case, we have a single “pre-treatment” visit (ODM: Repeating=No) and several “treatment visits” (ODM: Repeating=Yes”). We will give the “pre-treatment” visit the visit number 1, and the repeating visits a number starting from 2. We could however also have started from e.g. 11 for the repeating visits, that is completely OK.

The nice thing about ODM is that in the clinical data the repeat number is already given by the attribute “RepeatKey”. We can easily see this when using the menu “View – ClinicalData”, and select the checkbox “Also display RepeatKeys”. For example, for our data point “systolic blood pressure”:

We see that for subject 001, the “treatment” visit (SE.VISITA) has been repeated 3 times (RepeatKey values from 1 to 3). So we can easily use the value of “RepeatKey” for calculating the visit number.

In order to do so, drag-and-drop the “StudyEventDef Treatment” to the SDTM cell “VS.VISITNUM”:

A dialog shows up:
We do however not retrieve the OID (the identifier) but we want the “RepeatKey”, so we use the selection box as follows:

After clicking “OK”, the mapping script is generated:

```xml
# Mapping using ODM element StudyEventData using value from attribute StudyEventRepeatKey
$VS.VISITNUM = xpath(/StudyEventData[$StudyEventOID='SE.VISITA']/$StudyEventRepeatKey);
```

And adapt it to:

```xml
# Mapping using ODM element StudyEventData using value from attribute StudyEventRepeatKey
$REPEATKEY = xpath(/StudyEventData[$StudyEventOID='SE.VISITA']/$StudyEventRepeatKey);
$VS.VISITNUM = $REPEATKEY + 1;
```

For the SDTM variable “VISIT”, we can reuse the value from “VISITNUM”. For example, we can set it as:

```xml
$VS.VISIT = concat('TREATMENT VISIT ', $VS.VISITNUM);
```

Our result then becomes:
How the visits are exactly named is of course up to the mapper. This is just an example. Also, do not forget to add the visit numbers and names to the TV (trial visits) dataset.

**The baseline flag**

The last mapping we want to do is the one for the „baseline“ flag (VSBLFL). How to do that is to be decided on by the sponsor. Often, the baseline flag measurement is assigned to the last measurement for each test before first exposure to the study treatment.

As there are different ways to generate a mapping for the baseline flag, depending on how the reference start date is set and what needs to be taken as the baseline flag, we developed a separate tutorial “Generating baseline flags”. It explains these different ways in great detail. Please use this tutorial for understanding the different ways of assigning the baseline flag.

Also remark that in SDTM 1.5, for which there is no Implementation Guide yet as of the time of writing this tutorial (July 2017), a new variable –LOBXFL (Last Observation Before Exposure Flag) has been added. Our SDTM-ETL software already implemented an algorithm for calculating the values for this new variable. It requires a post-processing step, which can be initiated at execution time using the checkbox ”Perform post-processing for assigning “LOBXFL”

---

3 Also see our article „Why LOBXFL should not be in SDTM”, which can be found at: [http://cdiscguru.blogspot.com/2016/08/why-lobxfl-should-not-be-in-sdtm.html](http://cdiscguru.blogspot.com/2016/08/why-lobxfl-should-not-be-in-sdtm.html)
Again, please note that no SDTM-IG for SDTM v.1.5 has been published yet (and thus that v.1.5 is not accepted yet by the regulatory authorities), so that one should not use this feature until there is a new SDTM-IG. We did however already add this feature so that existing customers are not forced to acquire a software update once SDTM v.1.5 and its IG become accepted by the authorities.