

SDTM-ETL 4.4 User Manual and Tutorial

Author: Jozef Aerts, XML4Pharma

Last update: 2024-02-10

Handling Unscheduled Visits

Table of Contents

Table of Contents	1
Introduction.....	1
New feature in SDTM-ETL 4.4	2
Conclusions.....	7

Introduction

The way the SDTMIGs recommend to handle unscheduled visits is at least ... bizarre. First of all, for most domains, they require a variable VISITNUM (visit number) as a derivative of VISIT (visit name) with the notice "Numeric version of VISIT, used for sorting". Very probably meant is "for chronological sorting". As the chronology is however already provided by the -DTC (date/time of collection), it looks that the only reason for VISITNUM is that reviewers are incapable to sort on date/time-s, maybe that the only thing they can sort on is numeric values¹. Unfortunately, it looks as CDISC has once again given in to demands of the regulatory agencies to compensate for the primitiveness of their review tools.

Also very strange is that VISITNUM must come, though it is a derivative of VISIT, before VISIT, which is ... bad design ...

The way VISITNUM is defined, as a number, makes it complicated when there are unscheduled visits, i.e. visits that are executed between planned ("scheduled") visits. Essentially, such visits should not get a VISITNUM when the data is already coming (on basis of -DTC) in chronological order. Well, ideally, VISITNUM should not be present anyway when the data is already coming in chronological order. Essentially, the SDTM requirement for VISITNUM reminds me of the punch cards of the 80ies, where [we used column 73-80 for a sequence number](#) for the worst case that we dropped our stack of cards.

One popular of treating VISITNUM for an unscheduled visit is to assign it a high number like "99" or "999". Essentially, this should be more than sufficient when the data is coming in chronological order, as there is no reason at all for the reviewer to have to "sort" the data.

I consider this as the best practice.

Remark also that VISIT and VISITNUM for unscheduled visits should not appear in the trial design TV (Trial Visits) dataset, as this domain should only contain "planned" visits.

Another approach, also recommended by an, in the mean time, [retired Phuse paper](#), is to use fractional numbers (like "3.1", "3.2" ...) for VISITNUM for unscheduled visits.

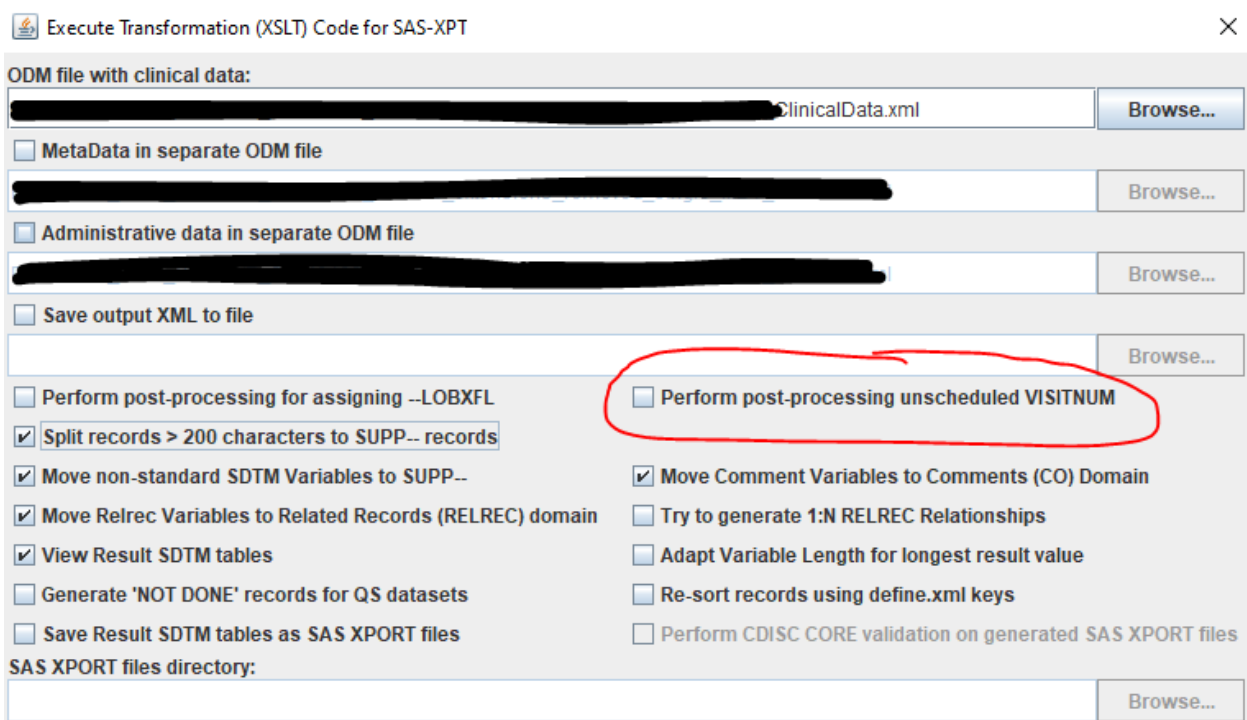
¹ As SDTM and SEND are using ISO-8601 notation, date/time-s can also simply be sorted alphabetically, giving the same result as when first converting to "calendar" date/time-s and then sorting.

This requirement asks the mappers to sort the data chronologically, then, for unscheduled visits, (re)assign the visit number. Essentially, this means (like for assigning the --LOBXFL flag) that a post-processing must be applied, which can easily become yet-another source for errors.

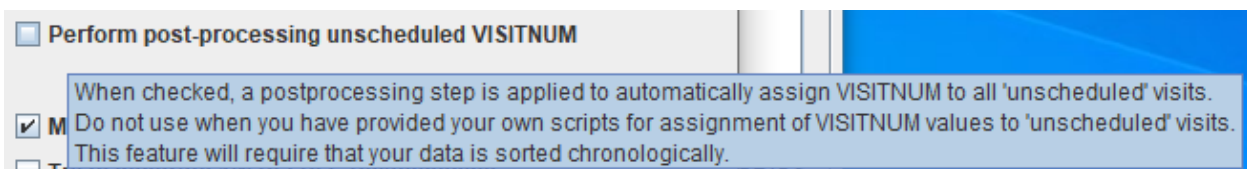
Our own recommendation is to not apply this additional post-processing step when the data already comes in chronological order, but some of our customers want to have it anyway. So we implemented it.

New feature in SDTM-ETL 4.4

As of version 4.4, when executing the mappings, one will notice an additional checkbox in the "Execute Transformation (XSLT) code ..." dialog:



When one hovers the mouse over it, one obtains more information:



And when checks the checkbox, a dialog is presented:

? In order to be able to correctly assign fractional VISITNUM values to 'unscheduled' visits some conditions need to be fulfilled:

- the data needs to be ordered/sorted chronologically.

If your data in the ODM 'ClinicalData' file is already ordered chronologically (oldest first), as the ODM standard prescribes, there is nothing additionally that must be done.

Alternatively, you can have the data sorted the data according to the define.xml keys. In order to do so, assign keys to the dataset definition, including a --STDTC variable or --DTC variable, and check the checkbox 'Re-sort using define.xml keys'.

- For unscheduled visits, the value for 'VISIT' needs to contain the word 'UNSCHEDULED' (uppercase), e.g. 'UNSCHEDULED', or 'UNSCHEDULED A' or 'UNSCHEDULED B' or ...

For multiple 'unscheduled' visits:

VISITNUM increment: 0.1

VISITNUM increment: 0.01

Concatenate generated VISITNUM value to value of VISIT

With blank (space) character as separator

With another character as separator

Separation character:

Displaying the requirements for correct use, and an option to choose between "0.1" and "0.01" as the increment value, i.e. the increase in number value VISITNUM should get for each subsequent unscheduled visit.

By default, the value for VISIT will be extended with the visit number, separated by a blank. For example, if, for the unscheduled visits, one has "\$VS.VISIT = 'UNSCHEDULED'", the final result will e.g. be "UNSCHEDULED 2.1", thus taking care of the 1:1 relation between VISIT and VISITNUM.

One can however switch this off by unchecking the checkbox "Concatenate generated VISITNUM value to value of VISIT". One can also provide another character (e.g. "-") As the separator instead of the blank character.

The requirements for correct use are important:

- The data needs to be sorted chronologically (see further)
- The value for "VISIT" (visit name) in the SDTM must be set to "UNSCHEDULED" (case-sensitive) by the mapping script, or at least be part of the visit name, otherwise the algorithm will not recognize that this is an unscheduled visit.

For example:

```

The Transformation Script
1 # Mapping using ODM element StudyEventData using value from attribute StudyEventOID
2 # Generalized for all StudyEvents
3 $VISITOID = xpath(/StudyEventData/@StudyEventOID/);
4 if($VISITOID = 'UNS') {
5     $VS.VISIT = 'UNSCHEDULED';
6 } else {
7     $VS.VISIT = $VISITOID;
8 }
    
```

- In the mapping for "VISITNUM", there must be a "null" placeholder value for the unscheduled visit. For example:

```

The Transformation Script
1 $VISITNAME = xpath(/StudyEventData/@StudyEventOID/);
2 if(starts-with($VISITNAME, 'V')) {
3     $VISITNAME = replace($VISITNAME, '_', '');
4     $VS.VISITNUM = substring-after($VISITNAME, 'V');
5 } elseif(starts-with($VISITNAME, '█')) {
6     $VISITNAME = substring-after($VISITNAME, '█');
7     $VS.VISITNUM = concat('10', $VISITNAME);
8 } elseif($VISITNAME = 'UNS') {
9     $VS.VISITNUM = '';
10 } else {
11     $VS.VISITNUM = "-999";
12 }

```

Where a "null" (empty) placeholder visit number is assigned when in the source data, a visit is declared as "unscheduled" by the identifier "UNS"².

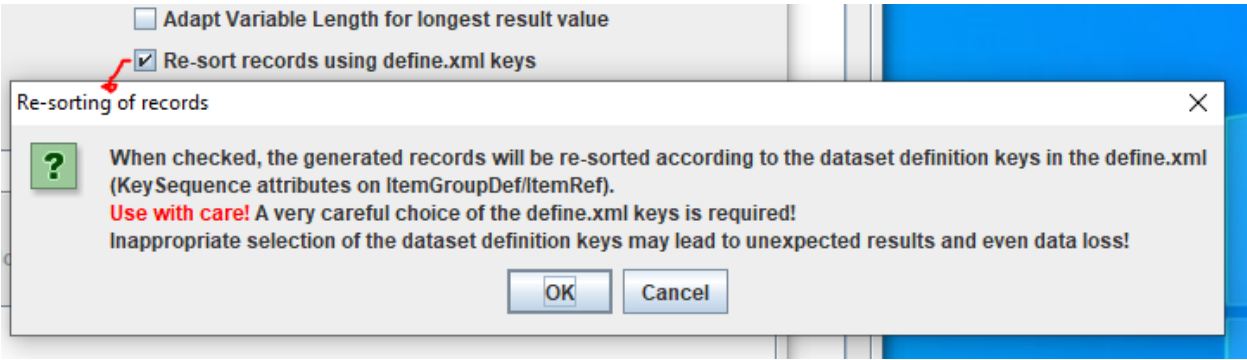
For the requirement that the data must be in chronological order, there are two possibilities:

- The source data is already ordered chronologically (as essentially required by the ODM standard, but not every EDC vendor adheres to this). In such a case, except for checking the checkbox "Perform post-processing unscheduled VISITNUM", nothing special needs to be done.
- The source data is **not** ordered chronologically (based on the date/time of collection). In that case, one will also need to check the checkbox "Re-sort records using define.xml keys"

Perform post-processing for assigning --LOBXFL
 Split records > 200 characters to SUPP-- records
 Move non-standard SDTM Variables to SUPP--
 Move Relrec Variables to Related Records (RELREC) domain
 View Result SDTM tables
 Generate 'NOT DONE' records for QS datasets
 Save Result SDTM tables as SAS XPORT files
 Perform post-processing unscheduled VISITNUM
 Move Comment Variables to Comments (CO) Domain
 Try to generate 1:N RELREC Relationships
 Adapt Variable Length for longest result value
 Re-sort records using define.xml keys
 Perform CDISC CORE validation on generated SAS XPORT files

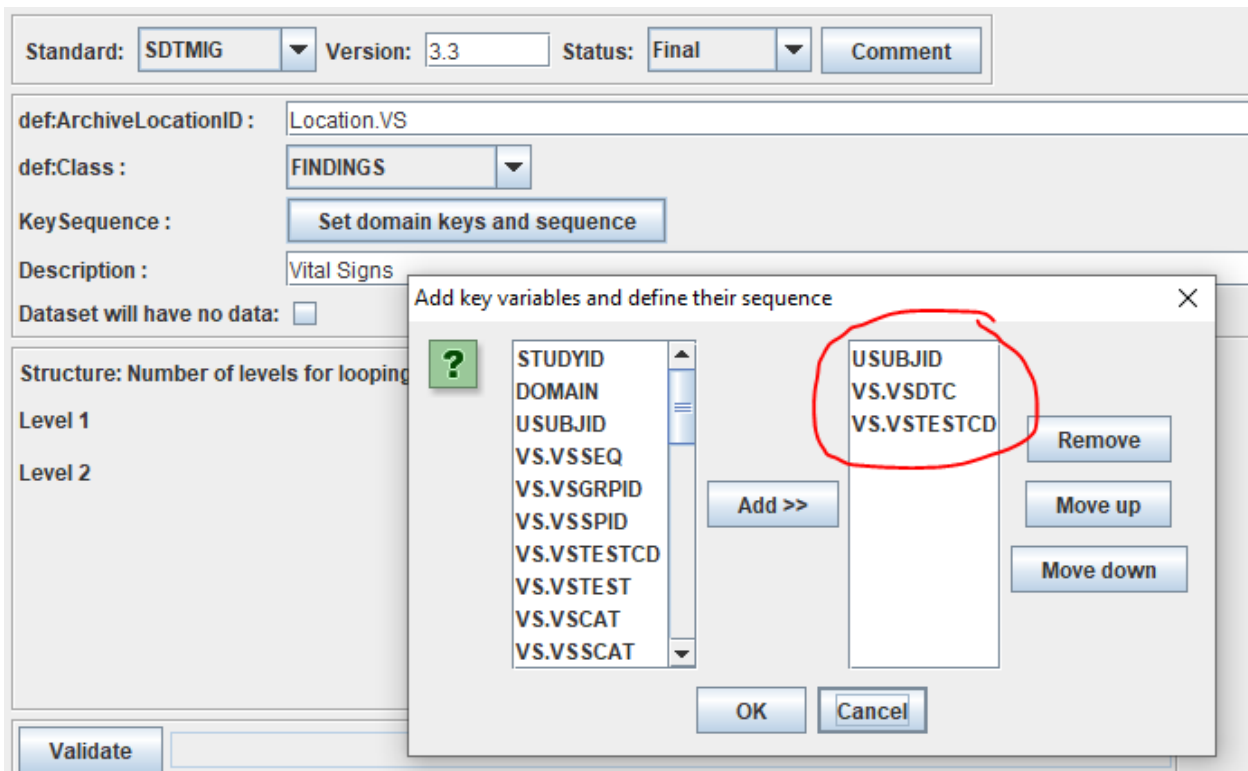
SAS XPORT files directory:

which pops up an information dialog:



and take care that the "sorting keys" are well defined, meaning that e.g. -DTC is among them. For example:

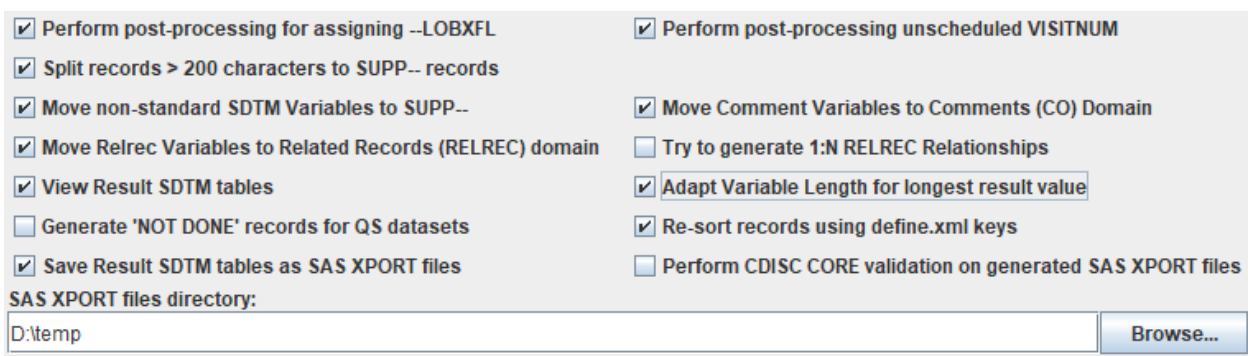
² There can be of course other ways the visit can be identified as "unscheduled" ...



Please take into account that using VISITNUM as a sort key is not a good idea here, as we want to generate VISITNUM ...³

Just as a reminder: setting the keys for a dataset (and for other dataset properties) can be achieved by a double-click on the first cell in the row for that dataset definition.

So, in our case, we would then e.g. have:



When then executing, this e.g. leads to:

³ In the final define.xml that is submitted, one can of course than use VISITNUM as a key again, though there is no extra advantage to do so.

VS.VSTESTCD	VS.VSTEST	VS.VSORRES	VS.VSORRESU	VS.VSSTAT	VS.VSREASND	VS.VSLOBXFL	VS.VISITNUM	VS.VISIT	VS.VSDTC
DIABP	Diastolic Blood Pres...	82	mmHg		Y		1 V1		2022-09-22T14:05
HR	Heart Rate	72	beats/min		Y		1 V1		2022-09-22T14:05
RESP	Respiratory Rate	16	breaths/min		Y		1 V1		2022-09-22T14:05
SYSBP	Systolic Blood Press...	129	mmHg		Y		1 V1		2022-09-22T14:05
TEMP	Temperature	36.8	C		Y		1 V1		2022-09-22T14:05
DIABP	Diastolic Blood Pres...	78	mmHg				1 V1		2022-09-22T15:30
HR	Heart Rate	70	beats/min				1 V1		2022-09-22T15:30
RESP	Respiratory Rate	14	breaths/min				1 V1		2022-09-22T15:30
SYSBP	Systolic Blood Press...	124	mmHg				1 V1		2022-09-22T15:30
TEMP	Temperature	36.7	C				1 V1		2022-09-22T15:30
DIABP	Diastolic Blood Pres...	86	mmHg				1.1 UNSCHEDULED 1.1		2022-09-26T11:00
HR	Heart Rate	77	beats/min				1.1 UNSCHEDULED 1.1		2022-09-26T11:00
RESP	Respiratory Rate	12	breaths/min				1.1 UNSCHEDULED 1.1		2022-09-26T11:00
SYSBP	Systolic Blood Press...	119	mmHg				1.1 UNSCHEDULED 1.1		2022-09-26T11:00
TEMP	Temperature	37.1	C				1.1 UNSCHEDULED 1.1		2022-09-26T11:00
DIABP	Diastolic Blood Pres...	87	mmHg				1.2 UNSCHEDULED 1.2		2022-09-26T12:00
HR	Heart Rate	78	beats/min				1.2 UNSCHEDULED 1.2		2022-09-26T12:00
RESP	Respiratory Rate	13	breaths/min				1.2 UNSCHEDULED 1.2		2022-09-26T12:00
SYSBP	Systolic Blood Press...	120	mmHg				1.2 UNSCHEDULED 1.2		2022-09-26T12:00
TEMP	Temperature	37.2	C				1.2 UNSCHEDULED 1.2		2022-09-26T12:00
DIABP	Diastolic Blood Pres...	80	mmHg				3 V3		2022-10-05
HR	Heart Rate	60	beats/min				3 V3		2022-10-05
RESP	Respiratory Rate	12	breaths/min				3 V3		2022-10-05

Where one sees that 2 unscheduled visit have been taken place between the planned visits V1 and V3, for which VISITNUM=1.1 and VISITNUM=1.2 has been assigned.

Important remark: if the data is not ordered chronologically, and one asks to assign VISITNUM for unscheduled visits automatically, incorrect values for VISITNUM will be assigned.

Some companies prefer to have VISITNUM for unscheduled visits still based on e.g. "99", but with a fraction added, e.g. "99.1", "99.2", meaning that it is not looked at what the value of VISITNUM was in the planned visit that took place immediately before the unscheduled visit. This is also possible in SDTM-ETL.

To do so, just assign "99" (or whatever other number one may want to use as the "base") to VISITNUM in the mapping script, e.g.:

```

0      $VISITNAME = SUBSTRING_INDEX($VISITNAME, '|', 1);
7      $VS.VISITNUM = concat('10', $VISITNAME);
8 } elseif($VISITNAME = 'UNS') {
9      $VS.VISITNUM = '99';
10 } else {
11     $VS.VISITNUM = "-999";
12 }

```

When then having the checkbox "Perform post-processing unscheduled VISITNUM" and having taken care that the data is ordered chronologically, the outcome will e.g. be:

VS.VISITNUM	VS.VISIT	
1	V1	2022-0
1	V1	2022-0
1	V1	2022-0
1	V1	2022-0
99.1	UNSCHEDULED 99.1	2022-0
99.1	UNSCHEDULED 99.1	2022-0
99.1	UNSCHEDULED 99.1	2022-0
99.1	UNSCHEDULED 99.1	2022-0
99.1	UNSCHEDULED 99.1	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0
99.2	UNSCHEDULED 99.2	2022-0

Conclusions

VISITNUM should not be an SDTM variable. It looks as it is only there because reviewers at regulatory authorities wanted it as they are incapable to sort data by collection data (–DTC or –STDTC values). The introduction of VISITNUM (for "sorting") causes additional problems, rather than solving some, especially for unscheduled visits. When the data is already sorted chronologically, the use of "99" or "999" for VISITNUM surely makes sense. However, many sponsors want to have fractional numbers for VISITNUM for unscheduled visits, based on the VISITNUM of the last scheduled visit.

The feature to do so, using a post-processing step has now been added to SDTM-ETL.