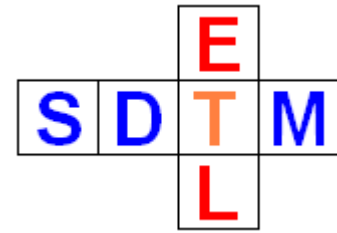# SDTM-ETL 4.0: Creating Records for QS "Logically Skipped Items"

Author: Jozef Aerts, XML4Pharma

Last update: 2020-10-22

Introduction

In some cases, the FDA publishes rules for SDTM that override existing SDTM rules (although also the FDA can comment during public review periods) and sometimes even conflicts with them.

The newest example of this is the rule that for QS (Questionnaires) datasets, also data of logically skipped questions must be submitted. See the latest "Technical Conformance Guide":

QS Domain (Questionnaires)

Some items in an instrument may be logically skipped per the instrument's instructions. Responses for logically skipped items should be (1) recorded and/or scored according to the instructions provided in the instrument's user manual, scoring manual, or other documentation provided by the instrument developer and (2) included in the submission dataset.

If instructions on how to record and/or score responses to logically skipped items are available from the instrument developer, then records for logically skipped items should be included in the submission dataset with the following:
- QSSTAT = 'NOT DONE';
- QSREASND = 'LOGICALLY SKIPPED ITEM'; and
- QSORRES, QSSTRESC, and QSSTRESN would be assigned according to the instrument's instructions.

If instructions on how to record and/or score responses to logically skipped items are not available from the instrument developer, then records for logically skipped items should be included in the submission dataset with the following:
- QSSTAT = 'NOT DONE';
- QSREASND = 'LOGICALLY SKIPPED ITEM'; and
- QSORRES, QSSTRESC, and QSSTRESN all set to null.

Remark that the above is not very consequent and violates SDTM rules, for example that population of "-ORRES" and -STAT" are mutually exclusive, i.e. only 1 of both is allowed to be populated.

This is unproblematic when such skipped items are collected and marked as such in the database, e.g. using a code "999". In most cases however (we checked with EDC and ePRO vendors), skipped questions are not explicitly collected and stored in the study database, as there is no data from them (logical, isn't it?), and because one can always do a lookup in the "skip rules" in order to see why a data point is missing. Such skip rules can easily be defined in the define.xml, but it looks as the FDA did not think about that.

So, this new rule forces SDTM generation tools to create records for non-collected data that usually are not in the database, and thus also not in the ODM export.

If the database however also contains records for skipped questions (such as "999" records), there is no problem at all.

Implementation in SDTM-ETL

In case the EDC system, thus also the ODM export, does not have information about questionnaire data that was not captured (as it was skipped due to the questionnaire workflow), a post-processing step is necessary.

In SDTM-ETL, we have solved this in an intelligent way. For the case that a questionnaire has "skip rules", such as the "Disability Rating Scale" questionnaire, for which also an annotated CRF exists, one can add a file "QS_skip_questions.txt" to the directory "QS_Skip_Questions".

This file in the "QS_Skip_Questions" folder then contains the information whether an item can be skipped, and will be used to generate "NOT DONE" records for the skipped questions when that information is not explicitly available in the source database and ODM.

The SDTM-ETL distribution already contains such an example file for the "Disability Rating Scale". The file is named "QS_skip_questions.txt". Here are some of the contents:

```
 6    DRS:QSDRS|ED102_1|ED1-Able to Communicate Clearly|COMMUNICATION ABILITY|false
 7    DRS:QSDRS|ED102_2|ED1-How They Communicate Primarily|COMMUNICATION ABILITY|true
 8    DRS:QSDRS|ED102_3|ED1-Correct Date and Time|COMMUNICATION ABILITY|true
 9    DRS:QSDRS|ED102_4|ED1-Few Words or Random Answers/Shouting|COMMUNICATION ABILITY|true
10    DRS:QSDRS|ED102_5|ED1-Moan/Groan/Sounds Not Understandable|COMMUNICATION ABILITY|true
11    DRS:QSDRS|ED104_1|ED1-Feed Independently Without Help|FEEDING|false
12    DRS:QSDRS|ED104_2|ED1-Understand Feeding Utensils|FEEDING|true
13    DRS:QSDRS|ED104_3|ED1-Know Meal Times|FEEDING|true
14    DRS:QSDRS|ED105_1|ED1-Use Toilet Independently|TOILETING|false
15    DRS:QSDRS|ED105_2|ED1-Manage Clothing When Toileting|TOILETING|true
16    DRS:QSDRS|ED105_3|ED1-Know When to Use Toilet|TOILETING|true
17    DRS:QSDRS|ED106_1|ED1-Can Dress/Groom Independently|GROOMING|false
18    DRS:QSDRS|ED106_2|ED1-Know How to Bathe/Wash|GROOMING|true
19    DRS:QSDRS|ED106_3|ED1-Understand How to Get Dressed|GROOMING|true
20    DRS:QSDRS|ED106_4|ED1-Start/Finish Grooming Activities|GROOMING|true
21    DRS:QSDRS|ED107_1|ED1-Function Completely Independently|LEVEL OF FUNCTIONING|false
22    DRS:QSDRS|ED107_2|ED1-Require Specific Aids/Equipment|LEVEL OF FUNCTIONING|false
23    DRS:QSDRS|ED107_3|ED1-Require Physical Assistance|LEVEL OF FUNCTIONING|false
24    DRS:QSDRS|ED107_4|ED1-Require Assistance Thinking Tasks|LEVEL OF FUNCTIONING|false
25    DRS:QSDRS|ED107_5|ED1-Require Assistance Managing Emotions|LEVEL OF FUNCTIONING|false
26    DRS:QSDRS|ED107_6A|ED1-Need a Helper Always Close By|LEVEL OF FUNCTIONING|false
27    DRS:QSDRS|ED107_6B|ED1-Need Help With All Major Activities|LEVEL OF FUNCTIONING|false
28    DRS:QSDRS|ED107_6C|ED1-Need 24-Hour Care|LEVEL OF FUNCTIONING|false
29    DRS:QSDRS|ED108_1|ED1-Independent Work/Social Situations|EMPLOYABILITY|false
30    DRS:QSDRS|ED108_2|ED1-Understand/Follow Directions|EMPLOYABILITY|false
31    DRS:QSDRS|ED108_3|ED1-Keep Track of Time/Schedules|EMPLOYABILITY|false
32    DRS:QSDRS|ED108_4|ED1-Perform Jobs, Manage Home/School|EMPLOYABILITY|false
33    DRS:QSDRS|ED108_5|ED1-Successful With Accommodations|EMPLOYABILITY|true
```

Fields in the file are separated by a vertical bar "|". The fields are:
- Field 1: The QS dataset identifier
- Field 2: The question identifier (OID of the ItemDef in the ODM file – for populating QSTESTCD)
- Field 3: The question label (for populating QSTEST)
- Field 4: The category (for filling QSCAT)
- Field 5: Boolean value whether the question can be skipped or not:
  when the value is "false", no "NOT DONE" record will ever be created for that item, as it could never be "logically skipped".

For example, the line:

```
13    DRS:QSDRS|ED104_3|ED1-Know Meal Times|FEEDING|true
```

indicates that for the dataset with identifier "DRS:QSDRS" (i.e. the QSDRS dataset in study "DRS") has an item "ED104_3" (the test code), with test name "ED1-Know Meal Times", with the category (QSCAT) "FEEDING" is an item that can be skipped under certain circumstances.
If the software then detects that for a certain subject/visit combination, there is no data for item "ED104_3", it will generated a record with QSORRES empty, QSSTAT="NOT DONE" and QSREASND="LOGICALLY SKIPPED ITEM".

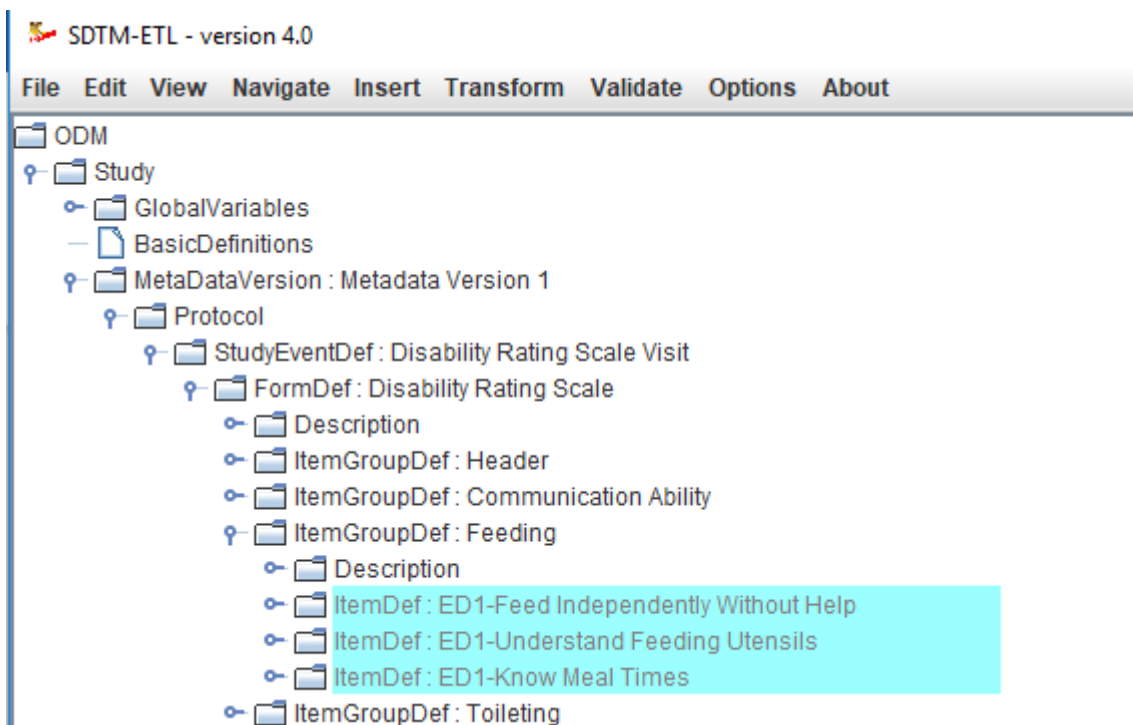This file can contain information for more than one QS datasets. For example:

As one may already guess, lines starting with a "#" are comment files and will be ignored by the software.

If the software finds more than one file, or none at all, in the directory "QS_Skip_Questions", it will ask which one to use, and/or allow the user to select such a file.

In the SDTM-ETL, the ODM file with metadata for the "Disability Rating Scale" questionnaire, when loaded looks like[1]:
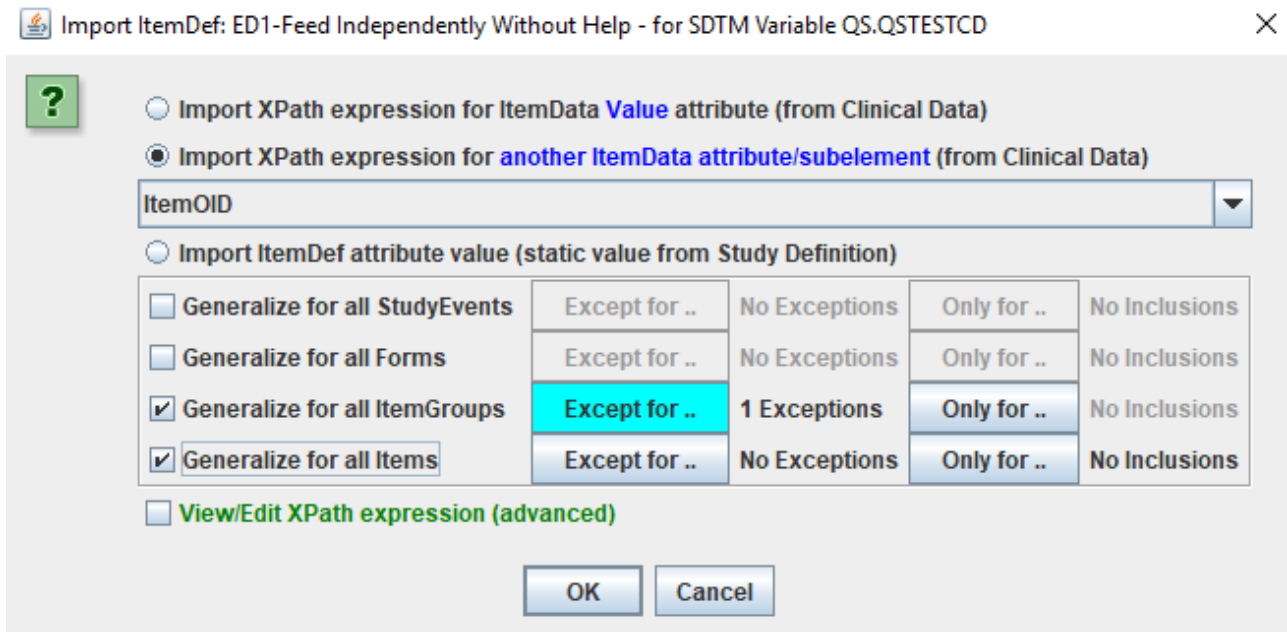

The SDTM-ETL script for QSTESTCD looks like:

Mapping using ODM element ItemData with ItemOID ED102_1 - value from attribute ItemOID
Generalized for all ItemGroups within the Form
Except for: IG.HEADER
Generalized for all Items within the ItemGroup
$QS.QSTESTCD =
xpath(/StudyEventData[@StudyEventOID='SE.DRS']/FormData[@FormOID='FO.DRS']/Item-
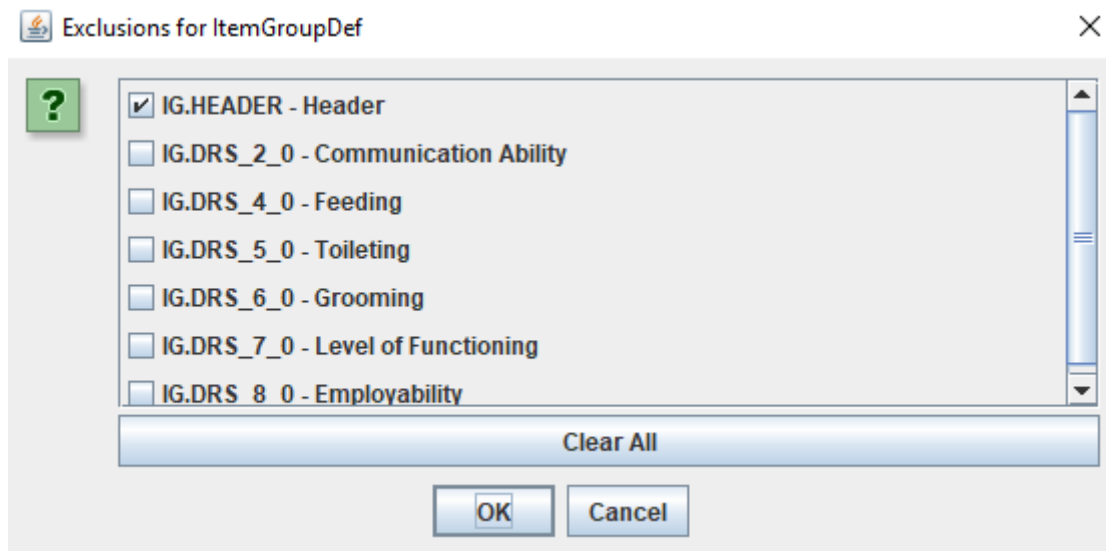GroupData[not(@ItemGroupOID='IG.HEADER')]/ItemData/@ItemOID);

Stating an iteration over all items in the form, except for the items in the "header". It was automatically created after a simple drag and drop and using the wizard:

---

[1] A copy of the "Disability Rating Scale" ODM file and the corresponding SDTM-ETL define.xml file that were used for testing this new feature, can be found in the distribution. Look for file "DRS_QRS_Metadata_ClinicalData.xml" in the folder "TestFiles\ODM1-3" and the folder "TestFiles\Define_2_0_Mappings\QRS_skip_questions".

With the one exception generated by clicking "Except for …" in "Generalize for all ItemGroups":

The script for "QSORRES" was generated in a very similar way and looks like:

?

Mapping Description and Link to external Document

SDTM-ETL mapping for QS.QSORRES | **External Document Link**

The Transformation Script

```
# Mapping using ODM element ItemData with ItemOID ED102_1
# Generalized for all ItemGroups within the Form
# Except for: IG.HEADER
# Generalized for all Items within the ItemGroup
# Generalized for all ItemGroups within the Form
# Except for: IG.HEADER
# Generalized for all Items within the ItemGroup
# Using decoded values from ODM CodeList CL.ED102_1
$CODEDVALUE = xpath(/StudyEventData[@StudyEventOID='SE.DRS']/FormData[@FormOID='FO.DRS']/ItemGroupData[not(@Item(
if ($QS.QSTESTCD == 'ED102_1') {
    $TEMP = decode($CODEDVALUE,'CL.ED102_1','en');
} elsif($QS.QSTESTCD == 'ED102_2') {
    $TEMP = decode($CODEDVALUE,'CL.ED102_2','en');
} elsif($QS.QSTESTCD == 'ED102_3') {
    $TEMP = decode($CODEDVALUE,'CL.ED102_3','en');
} elsif($QS.QSTESTCD == 'ED102_4' or $QS.QSTESTCD == 'ED102_5') {
    $TEMP = decode($CODEDVALUE,'CL.N_0_Y_1','en');
} elsif($QS.QSTESTCD == 'ED104_2' or $QS.QSTESTCD == 'ED104_3') {
    $TEMP = decode($CODEDVALUE,'CL.ALWAYS_TO_NEVER','en');
```

It uses the "decode()" function as the answers in the database are stored as numeric values, like "0" for "Always", "1" for "Most of the time", "2" for "Some of the time" and "3" for "Never". Storing the answers in the database as numeric values makes sense, as these numbers represent a score, from which the total score is calculated.

This can be easily visualized using the menu "View – Item Associated CodeList":

Details for CodeList: Always to Nover (OID: CL.ALWAYS_TO_NEVER) ✕

(i)

Coded Value	Language	Decoded Text
0	en	Always
1	en	Most of the time
2	en	Some of the time
3	en	Never

We also need take care that there is a "dummy" mapping specified in QSSTAT and QSREASND, as these will be automatically populated in a post-mapping run, i.e. this run will override the values in QSSTAT and QSREASND:

Mapping Description and Link to external I

SDTM-ETL mapping for QS.QSSTAT

The Transformation Script

```
$QS.QSSTAT = '';
```

Mapping Description and Link to external

SDTM-ETL mapping for QS.QSREASP

The Transformation Script

```
$QS.QSREASND = '';
```

When we now run "Transform – Generate Transformation Code" followed by "Execute Transformation Code", the system detects that there is a study-specific QS dataset instance defined, and displays an additional checkbox "Generate 'NOT DONE' records for QS datasets":

If we leave this checkbox unchecked, nothing special will happen, and our QSDRS dataset will look like:

	STUDYID	DOMAIN	USUBJID	QSSEQ	QSTESTCD	QSTEST	QSCAT	QSORRES	QSSTRESC	QSSTRESN	QSSTAT	QSREASND	QSEVAL	VISITNUM	QSDTC
1	DRS	QS	P001	1	ED102_1	ED1-Able to Communicate Clearly	COMMUNICA	Consisten	0	0			CAREGIVER	1	2015-02-16
2	DRS	QS	P001	2	ED102_2	ED1-How They Communicate Primari	COMMUNICA	Speech	0	0			CAREGIVER	1	2015-02-16
3	DRS	QS	P001	3	ED102_3	ED1-Correct Date and Time	COMMUNICA	Sometimes	2	2			CAREGIVER	1	2015-02-16
4	DRS	QS	P001	4	ED104_1	ED1-Feed Independently Without H	FEEDING	Yes	Yes	.			CAREGIVER	1	2015-02-16
5	DRS	QS	P001	5	ED104_2	ED1-Understand Feeding Utensils	FEEDING	Always	0	0			CAREGIVER	1	2015-02-16
6	DRS	QS	P001	6	ED104_3	ED1-Know Meal Times	FEEDING	Always	0	0			CAREGIVER	1	2015-02-16
7	DRS	QS	P001	7	ED105_1	ED1-Use Toilet Independently	TOILETING	Yes	Yes	.			CAREGIVER	1	2015-02-16
8	DRS	QS	P001	8	ED105_2	ED1-Manage Clothing When Toileti	TOILETING	Always	0	0			CAREGIVER	1	2015-02-16
9	DRS	QS	P001	9	ED105_3	ED1-Know When to Use Toilet	TOILETING	Always	0	0			CAREGIVER	1	2015-02-16
10	DRS	QS	P001	10	ED106_1	ED1-Can Dress/Groom Independentl	GROOMING	No	No	.			CAREGIVER	1	2015-02-16
11	DRS	QS	P001	11	ED106_2	ED1-Know How to Bathe/Wash	GROOMING	Always	0	0			CAREGIVER	1	2015-02-16
12	DRS	QS	P001	12	ED106_3	ED1-Understand How to Get Dresse	GROOMING	Most of t	1	1			CAREGIVER	1	2015-02-16
13	DRS	QS	P001	13	ED106_4	ED1-Start/Finish Grooming Activi	GROOMING	Some of t	2	2			CAREGIVER	1	2015-02-16
14	DRS	QS	P001	14	ED107_1	ED1-Function Completely Independ	LEVEL OF	No	1	1			CAREGIVER	1	2015-02-16
15	DRS	QS	P001	15	ED107_2	ED1-Require Specific Aids/Equipm	LEVEL OF	No	0	0			CAREGIVER	1	2015-02-16
16	DRS	QS	P001	16	ED107_3	ED1-Require Physical Assistance	LEVEL OF	0	0	0			CAREGIVER	1	2015-02-16
17	DRS	QS	P001	17	ED107_4	ED1-Require Assistance Thinking	LEVEL OF	1	1	1			CAREGIVER	1	2015-02-16
18	DRS	QS	P001	18	ED107_5	ED1-Require Assistance Managing	LEVEL OF	0	0	0			CAREGIVER	1	2015-02-16
19	DRS	QS	P001	19	ED107_6A	ED1-Need a Helper Always Close B	LEVEL OF	No	0	0			CAREGIVER	1	2015-02-16
20	DRS	QS	P001	20	ED107_6B	ED1-Need Help With All Major Act	LEVEL OF	No	0	0			CAREGIVER	1	2015-02-16
21	DRS	QS	P001	21	ED107_6C	ED1-Need 24-Hour Care	LEVEL OF	No	0	0			CAREGIVER	1	2015-02-16
22	DRS	QS	P001	22	ED108_1	ED1-Independent Work/Social Situ	EMPLOYABI	Always	0	0			CAREGIVER	1	2015-02-16
23	DRS	QS	P001	23	ED108_2	ED1-Understand/Follow Directions	EMPLOYABI	Most of t	1	1			CAREGIVER	1	2015-02-16
24	DRS	QS	P001	24	ED108_3	ED1-Keep Track of Time/Schedules	EMPLOYABI	Most of t	1	1			CAREGIVER	1	2015-02-16
25	DRS	QS	P001	25	ED108_4	ED1-Perform Jobs, Manage Home/Sc	EMPLOYABI	Uncertain	1	1			CAREGIVER	1	2015-02-16
26	DRS	QS	P001	26	ED108_5	ED1-Successful With Accommodatio	EMPLOYABI	Certain o	0	0			CAREGIVER	1	2015-02-16
27	DRS	QS	J001	1	ED102_1	ED1-Able to Communicate Clearly	COMMUNICA	No	2	2			CAREGIVER	1	2019-03-07
28	DRS	QS	J001	2	ED102_4	ED1-Few Words or Random Answers/	COMMUNICA	Yes	1	1			CAREGIVER	1	2019-03-07
29	DRS	QS	J001	3	ED104_1	ED1-Feed Independently Without H	FEEDING	Yes	Yes	.			CAREGIVER	1	2019-03-07
30	DRS	QS	J001	4	ED104_2	ED1-Understand Feeding Utensils	FEEDING	Always	0	0			CAREGIVER	1	2019-03-07

QSSTRESN (and thus also QSSTRESC) contain the numeric (coded) values from the source database, whereas QSORRES contains the text value of the answer. This is also as described in the CDISC document. Remark that "skipped questions" according to DRS rules are … skipped, and do not appear in the dataset.

When however the checkbox "Generate 'NOT DONE' records for QS datasets" is checked:

The result is:

We now see that there are "NOT DONE" records for ED102_4 to ED102_6 although the original ODM dataset does not contain any information about these "skipped questions". These records have been generated "post-mapping" from the information in the file "QS_skip_questions.txt".

Again, if your dataset also contains information about questions that were not asked or were skipped (e.g. as "999" records), all this is not necessary.

Further materials

Some interesting articles have been published about this new FDA rule and its implications and challenges for the generation of SDTM-QS datasets. Here is a short list (not claiming to be complete):

- [Challenges of ProcessingQuestionnaire Data from Collection to SDTM to ADaM and Solutions using SAS®](#).
- [CDISC Doesn't Tell You What to Collect, But Regulators Do!](#)
Phuse-US-2019 presentation Diane Wold, slides 22-25.
- [CDISC & PhUSE Roundtable: QRS (Questionnaires, Ratings and Scales) - Logically Skipped Questions in SDTM](#)
- [CDISC end-to-end: Logically Skipped Items in SDTM-QS](#)