

PhUSE 2011

Paper CD05

QS Domain: Challenges and Pitfalls - an Interpretation Guide to the Implementation Guide

Knut Mueller, UCB Biosciences GmbH, Monheim, Germany

ABSTRACT

In Clinical trials, patient reported outcomes (PROs) are widely used to support label claims or even to be part of the primary endpoint. According to the SDTM implementation guide, PROs are destined to be stored in the Questionnaires (QS) domain. The QS domain is one of the most complex domains as it has to capture a huge variety of different questionnaires. It is defined in the respective section of the SDTM implementation guide. However there is room for interpretation especially with regard to terminology and the way coded or derived variables are to be stored. A sponsor defined SDTM interpretation guide helps to ensure consistency across projects and studies. This paper describes the challenges and pitfalls of setting up the QS domain for SDTM and the implications for setting up respective ADaM datasets for PRO data.

PATIENT REPORTED OUTCOMES

A large number of efficacy parameters consist entirely or partially of PROs. One of the most typical PRO parameter is quality of life. But also other subjective parameter that describe the patient's state of health from the patient's perspective are by definition PROs. Some comprehensive disease activity indices combine clinician reported outcomes, objective measures like lab parameters, and also PROs.

These PROs are in clinical trials in general assessed using standardized patient questionnaires. Consisting of several items, they provide depending on the instrument one or more total or domain scores that are calculated using validated algorithms specified by the author of the instrument. These derived scores are the parameters of interest and are either part of the above mentioned composite scores or serve directly as the primary or secondary outcome and are analysed using appropriate statistical methods.

The subject's responses to the items on the standardized questionnaire are collected and entered into a database. Depending on the instrument, item responses have to be rescaled, imputed, recoded, before one or more scores can be derived. The complexity of the scoring algorithms for a PRO covers a wide range. In essence the total and domain scores are what is of interest for the analysis. In addition the original responses have to be stored as the source data. In addition the source data might also be presented in data listings.

How does data collected on a questionnaire move from the questionnaire to the final statistical analysis?

CDISC STANDARDS

CDISC SDTM and ADaM provide a framework how data coming from clinical trials is tracked from the CRF to the tables figures and listings that support the clinical trial report.

The Study Data Tabulation Model (SDTM) version 1.2 defines a standard structure for study data tabulations to support product applications to regulatory authorities. The SDTM implementation guide in its current version 3.1.2. describes in detail how the different SDTM domains are to be set up (1). The SDTM domain that is designed for representing data collected through patient questionnaires is called QS.

The Analysis Dataset Model (ADaM) version 2.1 and the corresponding ADaM implementation guide version 1.0 defines and describes the structure for analysis datasets to support statistical analysis.

However, there is still room for sponsor defined specifications, and to some extent room for interpretation, due to ambiguity or a lack of exactness in some of the definitions. A sponsor defined interpretation guide for each domain therefore is of help.

This paper focuses on the interpretation of the QS domain and especially on the interpretation of the description of

PhUSE 2011

the result variables. There are two perspectives, the collected data's perspective and the SDTM variable's perspective. The 2 main questions reflecting both perspectives are, how are the collected data to be stored in SDTM, and, how are the result variables in SDTM filled.

CHALLENGES IN SDTM

THE QS DOMAIN

The SDTM implementation guide is not explicit whether the QS domain is designed to contain only PROs. However, there is one big difference between clinician reported outcomes and PROs. Once the patient has left the site or sometimes even before that when he received the study medication, it is not possible to query the patient's responses as this inquiry would introduce a potential bias. Thus PROs might still contain some unexpected "dirty" data (5).

Questionnaire has to undergo a number of transformations until it is finally reported in TFLs. As it is also used as an example in the SDTM ImpG, the several steps are illustrated for the SF36. The SF36 is a widely used 36-item scale on health related quality of life. The steps for the SF36 are as follows:

- original item responses are entered into the database (or is directly entered into the database using an e-PRO) On the questionnaire there is a number printed beside the possible responses of each item representing the decoded value. For example "1 – All of the time"
- In case of a paper and pencil study, the entries have to be scanned for unexpected values (5)
- Some of the items require a recoding of the numeric value that is printed on the questionnaire. In one case the value of another item triggers the possible values.
- Subdomain scores are calculated by summing up the values and expand them to a scale ranging from 0 to 100
- Optionally the domain scores can be adapted to the norms of different populations
- Two weighted component scores are derived out of the 8 subdomain scores.

The subdomain scores and the 2 component scores, that are derived from the 36 original items, are the parameters that will be analysed. Towards the statistical analysis further derivations have to be done:

- study specific imputation rules such as LOCF
- determining the change from baseline for the domain and component scores.

Which results have to be stored in SDTM, where and how? Which results have to be stored in ADaM?

The SDTM provides a set of variables that are designed contain the result values.

QSORRES

As the source, the original result has to be stored. This is the answer given by the patient on the questionnaire to a specific question. Section 4.1.5.1.1. states that

"The --ORRES variable contains the result of the measurement or finding as originally received or collected." (2)

The description in column 'CDISC notes' for variable QSORRES in section 6.3.5. repeats that statement and elaborates it:

"Finding as originally received or collected (e.g. RARELY, SOMETIMES). When sponsors apply codelist to indicate the code values are statistically meaningful standardized scores, which are defined by sponsors or by valid methodologies such as SF36 questionnaires, QSORRES will contain the decode format, and QSSTRESC and QSSTRESN may contain the standardized code values or scores." (2)

To some extent the second statement contains a contradiction, as the result of the measurement as originally received or collected often is not the decode, but a coded number. In essence the intention to collect the decode of an original response is understood. However, sometimes the finding "as originally received or collected" is a number (although there is a decode defined, for example at the top of the page, as it is the case for a set of items of the SF36). So the instruction to record the "decode of the response as originally received or collected" in some cases may be ambiguous.

Leaving the SF36, there are other scales where a decode only exists for the extreme ends of a scale, eg. a 11 point numeric rating scale with 0 meaning "no pain at all" and 10 meaning "worst imaginable pain". So it is not clear whether to store a mixed format in QSORRES or stick to the numeric results. As the numbering is part of the scale one could say that the number is the finding as originally collected. That might be nitpicking but it has the potential to cause confusion.

QSSTRESC AND QSTRESN

PhUSE 2011

The 'CDISC notes' column for QSSTRESC in section 6.3.5 repeats the statement for QSORRES and elaborates it further:

"Contains the finding for all questions or sub-scores, copied or derived from QSORRES in a standard format or standard units. QSSTRESC should store all findings in character format; if findings are numeric, they should also be stored in numeric format in QSSTRESN. If question scores are derived from the original finding, then the standard format is the score. Examples: 0, 1. When sponsors apply codelist to indicate the code values are statistically meaningful standardized scores, which are defined by sponsors or by valid methodologies such as SF36 questionnaires, QSORRES will contain the decode format, and QSSTRESC and QSSTRESN may contain the standardized code values or scores ".(2)

This definition opens up a wide range of possibilities how these two variables can be filled. The naive understanding would be to store the number that is printed on the questionnaire in both QSSTRESN and QSSTRESC. But what is meant by statistical meaningful codelist? Should QSSTRESC be a copy of QSORRES or of QSSTRESN? With regard to the SF36 the statistical meaningful codelist may be the recoded values. However as mentioned before, the recoding of item 8 has to take into account the value of item 7. So strictly speaking, the recoding algorithm is also an imputation rule. The 1 to 1 mapping between QSORRES QSTRESC and QSTRESN gets lost in this case.

DERIVED SCORES

SDTM IG 3.1. section 4.1.5.1.3 suggests to store derived scores, such as total scores in SDTM datasets. That implies that all the rules specified in the manual of a standard instrument have to be applied to calculate the scores. These rules often include several steps, including the implicit or explicit imputation of missing items. Referring to the CDER Guidance to Review Divisions regarding CDISC Data on the FDA homepage (4), it is strongly discouraged to include imputed (or derived) data in SDTM datasets. The recommendation would be that all imputations (or derived variables) should be completed and documented in the ADaM datasets created to support specific PRO analyses. That lead us to another difficult issue:

MULTIPLE RESPONSES

With the approach, that multiple answers cannot generally be set to missing values without having the possibility to trace them back they have to be stored somewhere in the QS dataset. In some cases multiple answers still include valuable information that can lead to a meaningful result after imputation. Therefore, it is necessary to document the occurrence of multiple responses, in order to be able to run sensitivity analyses and to list the original responses in a data listing. Again, as SDTM datasets should not include imputed data, the linkage between QSORRES, QSSTRESN and QSSTRESC makes it difficult to store the responses in all three of them without violating a SDTM IG rule.

SDTM SOLUTION

SOLUTION FOR QSORRES QSTRESC AND QSSTRESN

To account for the CDER guideline and to circumvent the problem of defining when a code is meant to be statistical meaningful on the level of SDTM data, and still to be compliant with the SDTM implementation guide it was decided to omit QSSTRESN and just populate QSORRES and QSTRESC with the decoded values where applicable. QSSTRESN is a permissible variable.

If the questionnaire also represents sponsor or author defined numeric codes of the responses, these responses will populate the SUPPQS dataset.

SOLUTION FOR MULTIPLE RESPONSES:

After having discussed several other possibilities, like concatenate multiple responses into a single response, that all failed to be compliant with the SDTM implementation guide, it was decided to store multiple responses in different observations. This is described in the SDTM IG in section 4.1.2.8.2

"If multiple result values (--ORRES) are reported for a test in a Findings class dataset, multiple records should be submitted for that --TESTCD."(2)

Each response gets its own row. The sponsor-defined identifier QSSPID can be used to mark multiple responses.

SOLUTION FOR DERIVED DATA

That also leads to the decision that no derivations are done inside SDTM. QS only contains data that was originally collected. That also means that all derivations have to be done completely in ADaM.

CHALLENGES IN ADAM:

Questionnaire data typically is efficacy data and its representation in ADaM follows the ADaM Basic Data Structure (BDS). The ADaM implementation guide states that analysis datasets are described to be "analysis-ready". It should contain the data needed for analysis. Although it is possible to create additional variables and rows it is not necessary to put everything in one single analysis dataset. The optimum number of analysis datasets might change from study to study. Ideally, Analysis datasets are tailored to the analysis they support, they contain in essence only

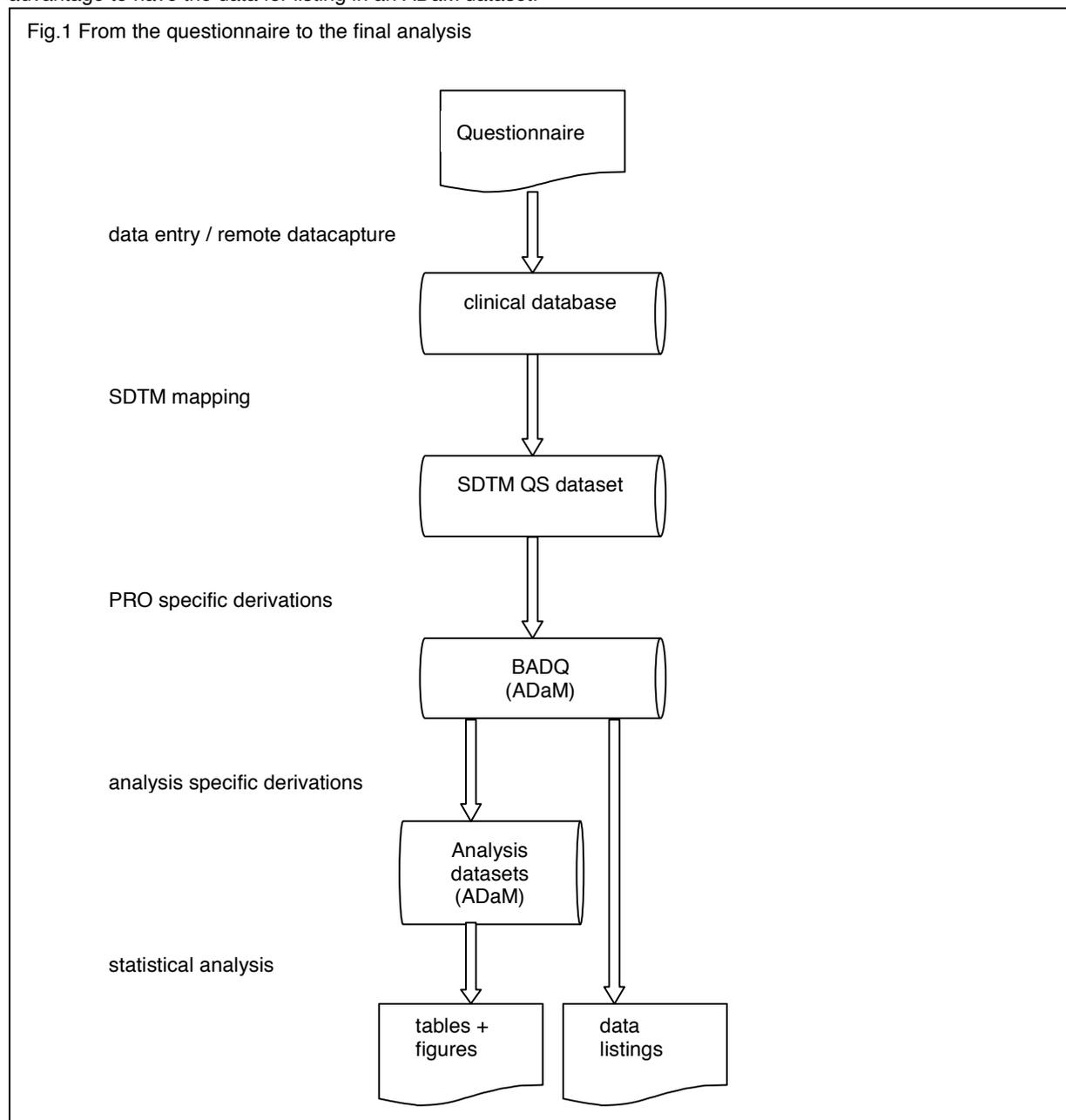
PhUSE 2011

the parameters and variables needed for a certain set of analyses.

For standardized questionnaire data this would mean that in each ADaM dataset the scores have to be derived again, as the SDTM dataset does not contain the derived scores.

As we decided not to provide derived results in SDTM it is necessary to provide them in analysis datasets. That is in essence data like domain scores, total scores etc. The ADaM ImpG explicitly mentions that it is not necessary to collect data in ADaM datasets only to for the purpose of data listings. However, data listings have to be created, and as the QS dataset does not contain all data in a format that it can be listed in a convenient way, it may be of advantage to have the data for listing in an ADaM dataset.

Fig.1 From the questionnaire to the final analysis



BASIC ADAM ANALYSIS DATASETS FOR QUESTIONNAIRE DATA

The solution to this would be a stepwise approach (see fig 1). First an analysis data set is created using the SDTM data as source data. This analysis dataset contains all questionnaire specific derivations that are analysis independent. In essence that is all the recoding and calculation of domain scores and total scores including imputation of missing data. This would provide a Basic Analysis Dataset for Questionnaire data (BADQ) which follows the ADaM BDS and provides all standardized scores, as well as all information that is needed for data

PhUSE 2011

listings. The BADQ are as standardized as SDTM datasets. These BADQs will thus serve as a kind of source data for further analysis datasets that are tailored for a specific statistical analysis. BADQs hold an intermediate position between SDTM datasets and "classical" analysis datasets. Nevertheless they should be fully ADaM BDS compliant. They do not support statistical analysis but contain all the basic derivations that are defined by the instrument itself. A good part of these parameters was optionally foreseen to be present in SDTM, but the aforementioned issues with the result variables, and the potential complexity of the derivation rules makes it more appropriate to provide these parameters in an analysis dataset.

SUMMARY

Both SDTM and ADaM implementation guide hold several possibilities how to store patient reported outcome data from the questionnaire to the final result.

To account for the complexity of PROs, the QS dataset in SDTM will only keep the original parameters. To be compliant with the SDTM implementation guide, only QSORRES and QSSTRESC are populated and both contain the decoded result value. A numeric representation of the result is stored in SUPPQS if a numeric response was printed on the questionnaire.

The QS dataset contains no derived data.

All derived parameters, that are defined by the authors of the standardised questionnaire will be stored in a BADQ.

The BADQ can be highly standardised as it reflects mostly instrument specific parameters. The BADQ also contains parameters that are needed for data listings. No analysis specific derivations are done within this dataset.

From the BADQs all analysis specific datasets are created. These analysis specific datasets are tailored to a specific analysis and contain exactly the parameters that are needed for the analyses they are created for.

CONCLUSION

- Questionnaire data provide some challenges with regard to be compliant with both SDTM and ADaM
- SDTM implementation guide still contain areas that are not clearly defined
- A sponsor specific Interpretation guide to the CDISC implementation guide is necessary to get consistency across studies and projects
- The introduction of Basic Analysis Datasets for Questionnaire Data provides a framework to facilitate standardization and to support data listings and serves as a basis for further analysis datasets.

REFERENCES

- (1) U.S. Department of Health and Human Services - Food and Drug Administration (2009): Guidance for Industry - Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims
- (2) Clinical Data Interchange Standards Consortium (2008): Study Data Tabulation Model Implementation Guide (Version 3.1.2)
- (3) Clinical Data Interchange Standards Consortium (2009): Analysis Model Implementation Guide (Version 1.0)
- (4) U.S. Department of Health and Human Services - Food and Drug Administration (2011): CDER Guidance to Review Divisions regarding CDISC Data. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM254113.pdf>
- (5) Mueller, K (2010): How to clean up dirty data in patient reported outcomes, Paper DH02, PhUSE conference 2010, www.phuse.com.

ACKNOWLEDGEMENTS

I would like to thank my colleague at UCB Biosciences GmbH, Irmgard Hollweck for valuable input during preparation of this paper

ABBREVIATIONS

AD	Analysis Dataset
ADaM	Analysis Dataset Model
BADQ	Basic Analysis Dataset for Questionnaire data
CDISC	Clinical Data Interchange Standards Consortium
FA	Findings About (SDTM Domain)
ImpG	Implementation Guide
IntG	Interpretation Guide
LOCF	Last observation carried forward
PRO	Patient Reported Outcome
QS	Questionnaires (SDTM Domain)
SF36	Short Form (36) Health Survey
SDTM	Standard Data Tabulation Model

PhUSE 2011

CONTACT INFORMATION

Knut Mueller
UCB Biosciences GmbH
Alfred-Nobel Str 10
D-40789 Monheim / Germany
Work Phone: +49 (0) 2173-482054
Fax: +49(0)2173-481947
Email:knut.mueller@ucb.com
Web:www.ucb.com