ADQRS: Basic Principles for Building Questionnaire, Rating and Scale Analysis Datasets

Nancy Brucken, INC Research/inVentiv Health, Ann Arbor, MI

Karin LaPann, Shire, Lexington, MA

ABSTRACT

Questionnaires, ratings and scales (QRS) are frequently used as primary and secondary analysis endpoints in clinical trials. The Submission Data Standards (SDS) QRS sub-team has compiled a considerable library of SDTM supplements defining standards for the collection and storage of QRS data. The ADaM ADQRS sub-team has been formed to develop addenda to these supplements, which will define standards for corresponding analysis datasets. This paper represents the current thinking of the ADQRS sub-team regarding basic principles for building QRS analysis datasets.

INTRODUCTION

The CDISC website provides the following definition of a questionnaire: "Questionnaires are named, stand-alone measures designed to provide an assessment of a concept. Questionnaires have a defined standard structure, format, and content; consist of conceptually related items that are typically scored; and have documented methods for administration and analysis. Questionnaires consist of defined questions with a defined set of potential answers. Most often, questionnaires have as their primary purpose the generation of a quantitative statistic to assess a qualitative concept." So, from an analysis viewpoint, questionnaires are a set of related questions that are generally combined in some manner to produce one or more scores, which are then summarized.

Ratings are a classification or ranking of someone or something based on a comparative assessment of their quality, standard, or performance. Functional tests are an example of ratings, and have been separated from questionnaires by the QRS team into the FT domain. The same is true for scales, which are defined by a set of criteria that makes up a single measurement, such as clinical classifications. Clinical Classifications instruments are also stored separately from questionnaires, in the Disease Response and Clinical Classification (RS) domain. In this sub-team, we are concentrating primarily on questionnaires in the QS domain, since their analysis is generally more complex.

Over the past several years, more than 100 different SDTM QRS supplements have been created to provide standards for the collection and storage of responses from questionnaires, ratings and scales. These supplements contain sample Case Report Forms (CRFs) for use in collecting the data, and controlled terminology for the QSTEST and QSTESTCD values representing the data in the SDTM QS domain. However, they stop short of describing corresponding datasets to be used for summarization and analysis of this data. The ADaM ADQRS sub-team was formed to carry this work forward, by developing basic standards for QRS analysis datasets that could be used by a variety of analysis methods.

ADQRS DATASET BASIC PRINCIPLES

During the process of developing the first ADaM QRS addendum, the ADQRS team has settled on some basic principles for analysis dataset structure and naming conventions. Even if you are creating analysis datasets for questionnaires, ratings or scales that are not covered by current QRS standards, following these conventions will help ensure that your datasets comply with CDISC standards. In all cases, the intent is to provide traceability from the analysis datasets back to their corresponding records in the SDTM QS domain.

DATASET NAMING CONVENTIONS

ADaM dataset names are currently limited to a maximum length of 8 characters, and must begin with 'AD', thus leaving only 6 characters available for the actual dataset name. Thus, we recommend excluding 'QS', 'RS' and 'FT' from the dataset name, in order to gain 2 more characters, and using the

remaining characters to hold an abbreviated version of the QRS name. For example, the QRS addendum for the Geriatric Depression Scale (Short Form) recommends using ADGDSSF for the dataset name. If there is concern that this practice might make it more difficult to distinguish ADQRS datasets from other analysis datasets, the dataset label may be used for this purpose; the QS origin may also be documented in define.xml.

This convention also implies that the QS domain should be split into separate analysis datasets for each individual questionnaire, rating or scale. These instruments are generally analyzed separately, and the associated datasets may be rather large. In addition, scoring calculations may vary widely for different questionnaires, so separating them simplifies the programming logic, and makes them easier to use.

DATASET VARIABLE CONVENTIONS

In the QS domain, QSCAT usually contains the acronym name of the actual questionnaire, rating or scale that is most commonly used. We recommend storing that value in PARCAT1 in the corresponding ADaM dataset, since per ADaM standards, PARCAT1 is a categorization of PARAM. While PARCAT1 will thus be identical on all records in the dataset, its presence will be helpful if multiple QRS analysis datasets are subsequently combined.

Similarly, we recommend that PARCAT2 be used to hold the associated questionnaire sub-scale name, if it exists. This practice allows for easier calculation of sub-scale scores, as well as the ability to quickly subset to only questions associated with a particular sub-scale. PARCAT2 can be null on some records if a question is not part of a sub-scale.

For maximum traceability back to QS, QSTEST and QSTESTCD values should be copied directly into PARAM and PARAMCD, respectively. QSSTRESN should also be copied over directly from QS into the analysis dataset, so that any changes to AVAL as a result of imputation or recoding can be easily compared back to their original value. Note that in the SDTM standard for QS, QSORRES contains the original text string of the answer, while QSSTRESC contains the character representation of QSSTRESN (e.g., when QSSTRESN = 2 then QSSTRESC = "2"). It is not necessary to copy QSSTRESC into AVALC, as it is unlikely that the character representation of the numeric code will be used for analysis or reporting. Instead, we suggest storing the value of QSORRES in AVALC, but only if it is needed for analysis. In any case, it may be helpful for traceability to bring QSORRES itself into the analysis dataset.

DATASET STRUCTURE CONVENTIONS

If sub-scale and/or total scores are included in QS because they were provided with the raw data as captured data, and are subsequently recalculated in the analysis dataset because of missing value imputation, or for any other reason, we recommend keeping the original QS records in the analysis dataset, with QSTEST and QSTESTCD copied into PARAM and PARAMCD. Separate records should then be created for the calculated scores, using the same QS QSTESTCD root for PARAMCD, plus additional characters representing the specific sub-scale or total score. For example, the individual GDSSF questions contain QSTESTCD/PARAMCD values beginning with 'GDS02', and the total score, if provided in QS, would have QSTESTCD='GDS0216'. The calculated GDSSF total score has a PARAMCD value of 'GDS02TOT'. This makes it easier to distinguish between collected and recalculated sub-scale and total scores, and allows for comparison between the two versions.

ABLFL (analysis baseline flag) and ANLzzFL (analysis flags) should be set only on sub-scale and total score records, unless individual question records are analyzed separately. This practice helps avoid cluttering the dataset with unnecessary flag values which won't be used for analysis. Additionally, CHG and BASE should only be computed for records that will be analyzed, and not for every individual question.

ADDITIONAL CONSIDERATIONS

Whether to include responses to all questions in the ADQRS data sets, or only summary records being analyzed, should be decided on a per-questionnaire basis, so this information will be included in the ADQRS addenda for the individual questionnaires. However, regulatory agencies have expressed a preference for including both the individual questions and summary records in the same dataset. They

recalculate all of the scores to compare against values provided by the sponsor as part of their review process, and it is easier for them if both individual questions and summary records are all stored in the same dataset. However, large datasets may take longer to process. A compromise is to create the full dataset, containing all individual and summary records, and then pull off only the summary records into a separate dataset, which is then used as input to the analysis and reporting programs.

All questionnaire scoring instructions should be documented in define.xml, along with all imputation rules used for handling of missing values.

ADQRS DATASET EXAMPLES

Here is a partial example of the variable-level metadata being proposed for the GDSSF analysis dataset:

Variable	Label	Type	Controlled Terms or Format	Source/Derivation
STUDYID	Study Identifier	Char		QS.STUDYID
USUBJID	Unique Subject Identifier	Char		QS.USUBJID
SITEID	Site Identifier	Char		ADSL.SITEID
ITTFL	Intent–To-Treat Population Flag	Char	"Y", "N"	ADSL.ITTFL *Conditional, must include at least one population flag.*
TRTP	Planned Treatment	Char		Derivation should be defined by sponsor according to analysis needs. *Conditional, should include at least one treatment variable, i.e. TRTP, TRTA, TRTSEQP or TRT01P.*
PARAM	Parameter	Char	SDTM codelist C106664 plus 'GDS02-Total Score'	QS.QSTEST for Responses to Individual Questions copied in from SDTM. Set PARAM to 'GDS02-Total Score' for the derived total score.
PARAMCD	Parameter Code	Char	SDTM codelist C106665 plus 'GDS02TOT'	QS.QSTESTCD for Responses to Individual Questions copied in from SDTM. Set PARAMCD to 'GDS02TOT' for the derived total score.
PARAMN	Parameter Number	Num		Number the questions in the order they appear on the form, and assign a higher value to the derived total score record.
PARCAT1	Parameter Category 1	Char	'GDS SHORT FORM'	Set to QS.QSCAT for records that come directly from SDTM and 'GDS SHORT FORM' for any derived records (i.e. Total Scores, Last Observed Value (LOV)).
VISIT	Visit Name	Char		QS.VISIT *Recommended for traceability.*
VISITNUM	Visit Number	Num		QS.VISITNUM *Recommended for traceability.*
AVISIT	Analysis Visit	Char		AVISIT represents the analysis visit for the record. Refer to the analysis plan for additional details.
AVISITN	Analysis Visit (N)	Num		A numeric representation of AVISIT.

Variable	Label	Type	Controlled Terms or Format	Source/Derivation
QSDTC	Date/Time of Finding	Char	ISO 8601	QS.QSDTC *Recommended for traceability.*
ADT	Analysis Date	Num	Date9.	Date portion of QS.QSDTC converted to numeric and displayed in a format such as DATE9. For derived total score records where PARAMCD="GDS02TOT" this information will just be copied over from the individual records that make up that questionnaire
ADY	Analysis Relative Day	Num		Example derivation: ADT – ADSL.TRTSDT + 1 if ADT ≥ TRTSDT, else ADT – ADSL.TRTSDT if ADT < TRTSDT
QSORRES	Finding in Original Units	Char		QS.QSORRES
AVAL	Analysis Value	Num		See Value Level Metadata in section 8.2.3.
AVALCAT1	Analysis Value Category 1	Char	"Normal", "Possible Depression", "Probable Depression"	Populate on derived total score records (where PARAMCD="GDS02TOT") - If 0<=AVAL<=5 then AVALCAT1="Normal"; else if 6<=AVAL<=9 then AVALCAT1="Possible Depression"; else if 10<=AVAL<=15 then AVALCAT1="Probable Depression". Source: categories described on GDSSF author's website, http://web.stanford.edu/~yesavage/GDS.html.
BASE	Baseline Value	Num		BASE=AVAL where ABLFL='Y'. Only the derived total score records where PARAMCD="GDS02TOT"should have BASE populated.
CHG	Change from Baseline	Num		Only populate for post-baseline total score records (where PARAMCD="GDS02TOT" and ADY>1) CHG=AVAL-BASE
DTYPE	Derivation Type	Char	"AVERAGE", "LOV", "MOV"	For PARAMCD="GDS02TOT": If the published scoring instructions are followed and 1-5 Responses are missing on the questionnaire then populate DTYPE with 'AVERAGE'. Other missing value imputations for individual questions are shown in the example with DTYPE equals 'MOV', and imputations for a completely missing total score are indicated with DTYPE equals 'LOV'. Otherwise leave NULL
ABLFL	Baseline Record Flag	Char	"Y"	Set ABLFL='Y' on the record defined as the baseline value. This flag will only be populated on the derived total score records (PARAMCD="GDS02TOT").
ANL01FL	Analysis Record Flag 01	Char	"Y"	Set ANL01FL='Y' for scheduled visits (VISIT is not of the form UNSCHEDULED XXX.XX) associated with the derived total score records (PARAMCD="GDS02TOT") *Permissible and useful for selecting records that appear in tables.*
QSSEQ	Sequence Number	Num		QS.QSSEQ. QSSEQ is included to provide traceability back to SDTM. *Recommended for traceability.*
INVNAM	Investigator Name	Char		ADSL.INVNAM *Permissible if needed for analysis.*

Table 1. Sample Variable Metadata

Note that in this dataset, we are keeping both the individual question and total score records. AVALC is not included, as the individual question responses do not have a corresponding text value. DTYPE indicates the type of imputation used for missing values, thus allowing for easy identification of imputed

responses. VISITNUM, QSDTC, QSORRES and QSSEQ have all been copied directly from the original SDTM records to provide traceability.

Another important aspect of questionnaire data is the need for detailed parameter value-level metadata (VLM). Most questionnaires contain one or more derived parameters. These need to be clearly defined when mapping from SDTM to ADaM, and the derivations for derived scores can be listed in greater detail in a VLM tab on the ADaM specification.

Here is an example being proposed for the GDSSF analysis dataset:

Variable	Where	Type	Controlled Terms or Format	Source/Derivation
AVAL	PARAMCD ≠ "GDS02TOT" (GDS02-Total Score)	Num		QS.QSSTRESN (Analysis Value is the result copied in from SDTM)
AVAL	PARAMCD = "GDS02TOT" (GDS02-Total Score)	Num		1) The sum of QS.QSSTRESN for QS.QSTESTCD in ('GDS0201'-'GDS0215') if no questions are missing 2) If <=5 responses are missing on a questionnaire then AVAL is 15*(Sum of QS.QSSTRESN for QS.QSTESTCD in ('GDS0201'-'GDS0215')/number of non-missing questions for QS.QSTESTCD in ('GDS0201'-'GDS0215')). If the result is a fraction, round up to the nearest integer. 3) If more than 5 responses are missing, do not calculate a total score, and do not create a total score record.

Table 2. Sample Parameter Value-Level Metadata

Complex questionnaires such as the Columbia Suicide Severity Rating Scale (CSSRS) provide additional challenges for ADaM. This scale is collected with different data entry screens for baseline, 1 month, 6 months, and lifetime. The QSTEST/QSTESTCD for these panels have different prefixes for each time point, and will need to be standardized to enable analysis across time points. Thus in this case, individual questions should be mapped to new PARAMCD/PARAM. For example, 'Non-Specific Suicidal Thought' will have a different QSTESTCD value at each time point as follows: QS.CSS0102, QS.CSS0202, QS.CSS0302A, QS.CSS1002A QS.CSS1002B, QS.CSS0302B. These can be aggregated under PARAMCD = 'NONSPEC' with different time points identified in ATPT (Baseline, Month 1, Month 6, Lifetime). QSTESTCD can be kept in the dataset as well to maintain traceability back to SDTM. The same logic applies to PARAM as well.

FUTURE PLANS

In order to help with the creation of future ADaM addenda, we are planning to create a document template for use in their development. We will also be developing an ADaM version of the current SDTM CDISC Operational Procedure 017 (COP 017), to enable teams to create their own ADaM addenda for ADQRS datasets, and submit them for CDISC approval. In addition, controlled terminology for the ADaM PARAM and PARAMCD values will be registered with the NCI team.

CONCLUSION

This paper has described some basic principles and approaches for developing analysis datasets for use with data from questionnaires, ratings and scales. Following these recommendations will help ensure that your analysis datasets conform to CDISC standards, in addition to providing traceability between the analysis datasets and the QS domain.

REFERENCES

CDISC Questionnaires, Ratings and Scales Web Page. 2017. Available at https://www.cdisc.org/foundational/qrs.

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Nancy Brucken
INC Research/inVentiv Health
Nancy.Brucken@inventivhealth.com

Karin LaPann
Shire
KLaPann0@Shire.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.