

Standardized, Customized or Both? Defining and Implementing (MedDRA) Queries in ADaM Data Sets

Richann Watson

Karl Miller

ABSTRACT

Investigation of drug safety issues for clinical development will consistently revolve around the experience and impact of important medical occurrences throughout the conduct of a clinical trial. As a first step in the data analysis process, Standardized MedDRA Queries (SMQs), a unique feature of MedDRA, provide a consistent and efficient structure to support safety analysis, reporting, and also address important topics for regulatory and industry users. A variance in working with SMQs is the ability to limit the scope for the analysis need (e.g., “Broad” or “Narrow”) but there is also the ability outside of the specific SMQs in allowing the ability to develop Customized Queries (CQs). With the introduction of the ADaM Occurrence Data Structure (OCCDS) standard structure, the incorporation of these SMQs, along with potential CQs, solidified the need for consistent implementation, not only across studies, but across drug compounds and even within a company itself. Working with SMQs one may have numerous questions: What differentiates the SMQ from a CQ and which one should be used? Are there any other considerations in implementation of the OCCDS standards? Where does one begin? Right here...

INTRODUCTION

In working with clinical trials and the data collected during study conduct, one of the most common aspects or areas that you will encounter in handling the safety data used in analysis will consistently be focused around the Adverse Events. One aspect, and the focus of this paper, is working with SMQs, which are pre-determined, validated sets of MedDRA terms, and the CQs which consist of specific terms of interest to the analysis.

Used to group MedDRA terms, to help define a specific medical condition, or define a particular area of interest SMQs and CQs are lists typically consisting of like Preferred Terms (PTs) developed to aid in the efficient identification of safety data key to analyze. Currently there exist over 100 SMQs with more in future development. Below is a list of some SMQs that are available:

- Anaphylactic reaction
- Cardiac failure
- Depression and suicide/self-injury
- Hepatic disorders
- Hypersensitivity
- Ischaemic heart disease
- Pregnancy and neonatal topics
- Rhabdomyolysis/myopathy
- Systemic lupus erythematosus

Like SMQs, CQs are used to group like adverse events, however, CQs are customized based on study need and therefore do not have a pre-defined list that can be referenced.

SPECIAL FEATURES OF SMQ

Since SMQs consist of sets of PTs, to avoid conflict with existing MedDRA terms, each SMQ name has '(SMQ)' appended at the end of the query. For example, a specific query looking for adverse events that would be classified under the SMQ of 'Anaphylactic reaction' would have 'Anaphylactic reaction (SMQ)'. Additionally, each SMQ is associated with a specific SMQ code that is 8 digits that starts with '2'.

Other special features, which we will discuss later, consist of useful options such as a scope search (“Narrow” vs “Broad”) and algorithmic searches using class or hierarchy structure. Through the use and implementation of these special features, this will allow users to further refine the SMQ list to potentially get the desired results without having to create a CQ.

SMQ SCOPE

Within an SMQ a PT can be classified as having a “Narrow” or “Broad” scope. The “Narrow” scope identify the PTs that are more likely to represent the condition or area of interest, where the “Broad” scope PTs may end up having little to minimal interest for use in the analysis upon further investigation. Thus, if a specificity analysis is needed then the PTs within the “Narrow” scope would be the terms of interest. However, if a sensitivity analysis is needed, then all PTs associated with the SMQ should be selected (i.e., PTs having either “Narrow” or “Broad” scope). Figure 1 illustrates the concept of specificity versus sensitivity analysis through the use of “Narrow” and/or “Broad” scope found with the SMQs.

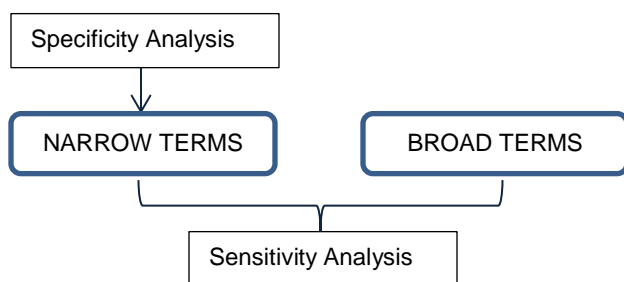


Figure 1 Specificity vs. Sensitivity Analysis in Relation to SMQ Scope

SMQ CLASS

To allow for an algorithmic search, each SMQ is further classified into classes (A, B, C or D). For example, there could be some instances where a SMQ may retrieve more terms than what is actually needed for a specificity analysis. In this case, the SMQ can then be further refined by selecting only the PTs that have CLASS = B. Table 1 provides a portion of the SMQ table for Acute Pancreatitis (SMQ). In this table you can see that for the terms with “Broad” scope that they are further classified into CLASS “B” and “C”. The terms associated with CLASS “B” are terms that are associated with laboratory values and the terms associated with CLASS “C” are the ones associated with signs and symptoms. Thus, if an analysis for Acute Pancreatitis is needed, dealing with laboratory results only, then the SMQ code can be refined even further just to look at the SCOPE = “BROAD” and the CLASS = “B”.

SMQ	SMQCODE	SCOPE	CLASS	PT	PTCODE
Acute pancreatitis (SMQ)	20000022	Broad	B	Amylase abnormal	10072327
Acute pancreatitis (SMQ)	20000022	Broad	B	Amylase creatinine clearance ratio abnormal	10073699
Acute pancreatitis (SMQ)	20000022	Broad	B	Amylase increased	10002016
Acute pancreatitis (SMQ)	20000022	Broad	B	Bilirubin conjugated abnormal	10067718
Acute pancreatitis (SMQ)	20000022	Broad	B	Blood bilirubin increased	10005364
Acute pancreatitis (SMQ)	20000022	Broad	B	Blood trypsin increased	10064751
Acute pancreatitis (SMQ)	20000022	Broad	C	Abdominal pain	10000081
Acute pancreatitis (SMQ)	20000022	Broad	C	Abdominal tenderness	10000097
Acute pancreatitis (SMQ)	20000022	Broad	C	Acute abdomen	10000647
Acute pancreatitis (SMQ)	20000022	Broad	C	Gastrointestinal pain	10017999
Acute pancreatitis (SMQ)	20000022	Broad	C	Nausea	10028813
Acute pancreatitis (SMQ)	20000022	Broad	C	Vomiting	10047700
Acute pancreatitis (SMQ)	20000022	Broad	C	Vomiting projectile	10047708

Table 1 Broad Terms Split into Class B and Class C

SMQ HIERARCHY

In addition to the scope and class feature, some SMQs may be broken down further into sub-queries. These sub-SMQs are related to another SMQ in a hierarchical nature, similar to the hierarchy of MedDRA, and can help refine the search to increase specificity of the selected terms for use in the analysis. The sub-SMQs can also be used to retrieve only those PTs of interest but be on alert, there can be up to four sub-queries within an SMQ.

For example, Depression and suicide/self-injury (SMQ) only has one sub-level (Figure 2) while Hepatic disorders (SMQ) has four sub-levels (Figure 3).

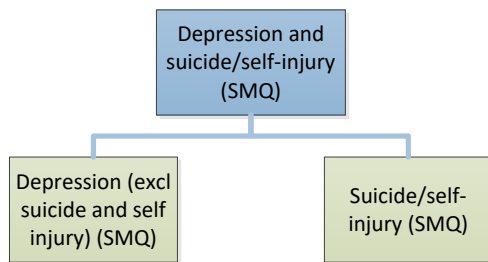


Figure 2 Depression and suicide/self-injury (SMQ) Hierarchy

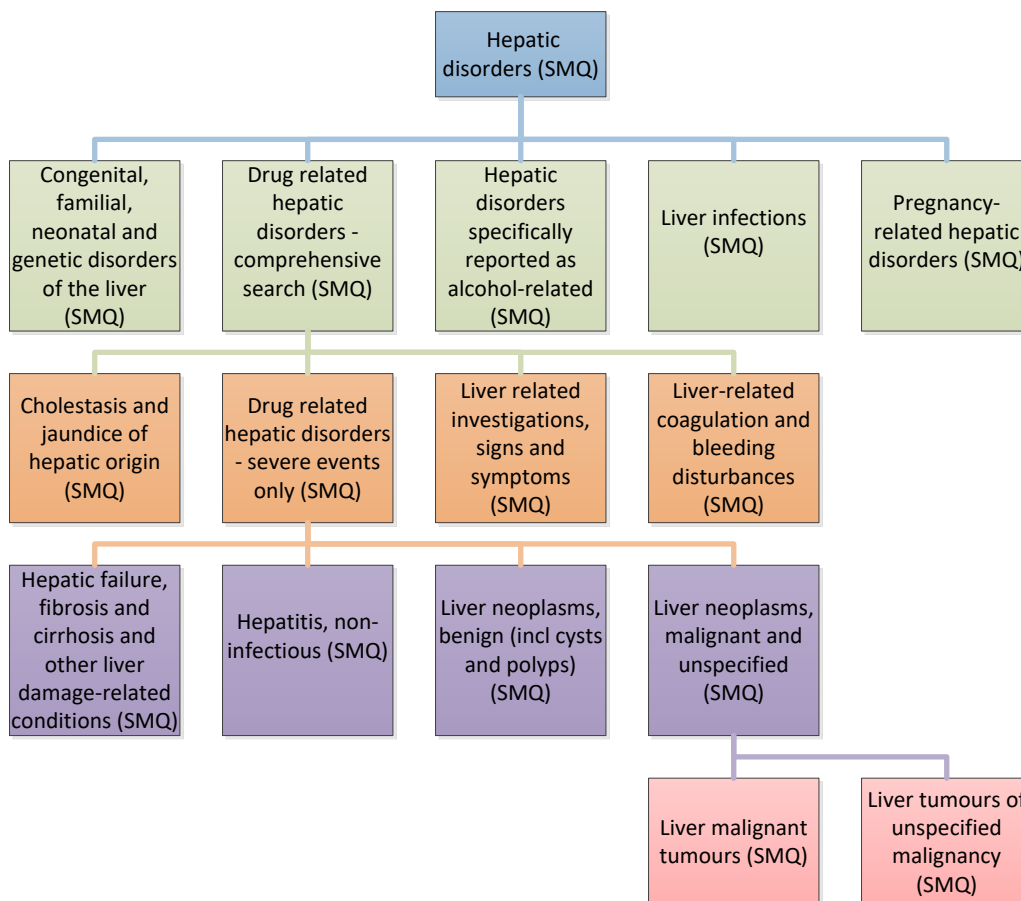


Figure 3 Hepatic disorders (SMQ) Hierarchy

DIFFERENCE BETWEEN SMQ AND CQ

In working with queries, through defining a set of conditions or simply grouping terms when applicable, there are differences in how to work with SMQs and CQs. One main difference between deciding to work with SMQs and CQs is that the SMQs are pre-defined lists that can be extracted using MedDRA PTs, where the CQs are lists that are specific to a study and are not pre-defined within MedDRA. In working with a set, pre-defined list versus a study-by-study base, this can have an impact on the efficiency of implementation through programming and the potential use of the query for future studies.

Another difference (and yet also a benefit to both) is that there are specific features built into the SMQ MedDRA dictionary that will allow for sub-setting and refining the list of PTs of interest. In the instance where you do implement the list of these pre-defined terms along with the pre-defined features, but the resulting list of PTs is not exactly what is needed for analysis or reporting purposes, then this is the case where the CQs come in handy by being able to expand the list of terms to encompass all the PTs of interest or by further sub-setting based on additional information.

Lastly, and key to implementation of CDISC standards, is the difference between the use of SMQ and CQ when it comes to implementing them into an ADaM data set for analysis and presentation purposes. As shown in Table 2, for SMQs there are four variables that can be used and therefore implemented into the specifications and data set, while at the same time for CQs there is only one variable.

	Variable Name	Variable Label	Type	Core
SMQs	SMQzzNAM	SMQ zz Name	Char	Cond
	SMQzzCD	SMQ zz Code	Num	Perm
	SMQzzSC	SMQ zz Scope	Char	Cond
	SMQzzSCN	SMQ zz Scope (N)	Num	Perm
CQs	CQzzNAM	Customized Query zz Name	Char	Cond

Table 2 SMQ and CQ Variables

Don't let the number differences between the variables used fool you on which will be more efficient to implement in your ADaM data set. In looking at the list of variables for queries, this can be the case where less is not always more; just because the use of a CQ requires one variable and SMQs can have up to at least four, does not always translate into the best option. For example, with SMQs the majority of the work is done up-front, with the pre-defined lists and implementation methods, so there is no need for the creation of special look-up tables and custom programming in implementation.

IMPLEMENTATION OF SMQ AND CQ VARIABLES

SMQ VARIABLES

When generating the specifications for the required analysis and implementation of SMQs, it is always good to keep in mind that each SMQ you are using will have its own set of variables. More than one SMQ cannot be stored within or under the same SMQzzNAM, SMQzzCD, SMQzzSC variables in your data set. See Table 3 for an example of how SMQ variables should be implemented into the data when there is more than one SMQ. One item of note is that it is not a requirement for zz to be sequential when implementing multiple SMQs.

Row	AEDECOD	SMQ11NAM	SMQ11CD	SMQ11SC
1	Acute fatty liver of pregnancy	Hepatic disorders (SMQ)	20000005	Narrow
2	Acute hepatitis B	Hepatic disorders (SMQ)	20000005	Narrow
3	Acute on chronic liver failure	Hepatic disorders (SMQ)	20000005	Narrow
4	Agitation neonatal			
5	Vomiting			
6	Ascites	Hepatic disorders (SMQ)	20000005	Narrow
7	Cholestasis of pregnancy	Hepatic disorders (SMQ)	20000005	Narrow
8	Neonatal cholestasis	Hepatic disorders (SMQ)	20000005	Narrow
9	Jaundice	Hepatic disorders (SMQ)	20000005	Narrow

Row	SMQ20NAM	SMQ20CD	SMQ20SC	SMQ23NAM	SMQ23CD	SMQ23SC
1	Pregnancy and neonatal	20000185	Narrow			

Row	SMQ20NAM	SMQ20CD	SMQ20SC	SMQ23NAM	SMQ23CD	SMQ23SC
	topics (SMQ)					
2						
3						
4	Pregnancy and neonatal topics (SMQ)	20000185	Narrow			
5				Acute pancreatitis (SMQ)	20000022	Broad
6				Acute pancreatitis (SMQ)	20000022	Broad
7	Pregnancy and neonatal topics (SMQ)	20000185	Narrow			
8	Pregnancy and neonatal topics (SMQ)	20000185	Narrow			
9				Acute pancreatitis (SMQ)	20000022	Broad

Table 3 Illustration of Implementation of SMQ Variables

In the example, the zz values were purposely set to various numbers to illustrate that the numbering does not have to start at '01' and does not have to be sequential. This can be seen through the use of pre-defined standard specifications across an entire compound, or even an entire therapeutic area, where specific SMQs can be assigned to their own variable naming convention across all clinical trials within a company.

CQ VARIABLE

Similar to SMQ, each CQ will have its own variable when implementing them into the ADaM data set. Just as with SMQs, more than one CQ cannot be stored within the same CQzzNAM variable within a data set of terms. In the case where a specific term may be included for multiple custom lists of PTs that are of interest for analysis, this helps keep the lists clear in their use and purpose for the analysis outputs.

HOW TO DETERMINE A CQ

With queries focused around the PTs of clinical events, where is the line between whether a query should be considered a SMQ or not? SMQs group terms and SOCs related to a defined medical condition or area of interest that can assist in case identification, through many related signs, symptoms, and other findings specifically related to the medical condition or area of interest. However, being stringent and rigid, SMQs do not cover all medical topics or safety issues that may be of interest, and will likely evolve and need to be refined as future progress is made.

With CQs there is more leniency and room to work outside of the defined lists of SMQs, how can you determine when you're using a CQ? CQs can be determined in several different ways:

- Provided list
 - A list that is built from one or more SMQs that are modified or joined together to provide a lookup table
 - A list of PTs, typically provided by a clinician, that are not already defined in an SMQ
- Programmatically determined
 - Criteria can be a SMQ with specific PTs removed that could not be removed using one of the built in features
 - Criteria provided by statistician, clinician or in Statistical Analysis Plan (SAP)
- Field captured on the case report form (CRF) to indicate the term/event is of special interest

The following examples illustrate the various way CQs can be determined. In Table 4 we illustrate a sample lookup table that is provided by either a clinician or statistician in the form of a spreadsheet. This lookup table can then be converted to a SAS® data set and then used to assign the correct CQzzNAM value as illustrated in Table 5.

Row	Criteria	Terms Needed
-----	----------	--------------

Row	Criteria	Terms Needed
1	Anaphylaxis per Criteria 1	Anaphylactic reaction
2	Anaphylaxis per Criteria 1	Dyspnoea
3	Anaphylaxis per Criteria 1	Angioedema
4	Anaphylaxis per Criteria 1	Migraine
5	Anaphylaxis per Criteria 1	Presyncope
6	Anaphylaxis per Criteria 1	Urticaria
7	Anaphylaxis per Criteria 2	Angioedema
8	Anaphylaxis per Criteria 2	Hyperhidrosis

Table 4 Sample Customized Query Lookup Table Provided as a List

Row	AEDECOD	CQ01NAM	CQ02NAM
1	Anaphylactic reaction	Anaphylaxis per Criteria 1	
2	Angioedema	Anaphylaxis per Criteria 1	Anaphylaxis per Criteria 2
3	Liver function test abnormal		
4	Migraine	Anaphylaxis per Criteria 1	
5	Injection site joint pain		
6	Injection site erythema		

Table 5 Illustration of Implementation of CQ Variables Using a CQ Lookup Table

In addition to using lookup tables, CQs can be programmatically determined based on provided criteria from a clinician, the SAP, or directly in the specifications for the CQ. As illustrated in Table 6, assume that the specifications stated, under the derivation for the CQ variable (CQ11NAM), that any AE would be marked as a term of interest in the analysis data if AEBODSYS = 'Infections and infestations' and AESER = 'Y'.

Row	USUBJID	AEBODSYS	AEDECOD	AESER	CQ11NAM
1	ABC-DEF-1234	Infections and infestations	Laryngitis	N	
2	ABC-DEF-1234	Infections and infestations	Sinusitis	Y	Serious Infections
3	ABC-DEF-1234	Nervous system disorders	Headache	N	
4	ABC-DEF-1234	Nervous system disorders	Migraine	Y	
5	ABC-DEF-1234	Skin and subcutaneous tissue disorders	Angioedema	N	
6	ABC-DEF-5678	Investigations	Liver function test abnormal	N	
7	ABC-DEF-5678	Infections and infestations	Nasopharyngitis	Y	Serious Infections
8	ABC-DEF-5678	Jaundice	Sinusitis	N	

Table 6 Illustration of Implementation of CQ Variable Determined Programmatically

Furthermore, CQs can be determined based on field captured on the CRF. Display 1, shows an example CRF for the Adverse Event form. On the form there is a field to indicate whether the AE is of special interest to the principal investigator. This field, or one similar to it, can be used to define the CQzzNAM variable logic found in the study specifications.

Form: Adverse Events

Has the subject experienced any adverse events during the study? Yes
 No

Adverse Event: _____

Is this an AE of special interest? Yes
 No

Display 1 Sample CRF of an Adverse Event Form

GOOD IMPLEMENTATION CONSIDERATIONS

With the availability for use with ADaM OCCDS v.1.0, all implementation methods from this paper can be implemented for creating the analysis data sets. With the many options, it's always good to keep certain scenarios in mind when implementation is set to start:

- Cannot store multiple codes in the same variable for either SMQs or CQs
- At the compound, or even therapeutic area level, you will want to keep the naming of the SMQ and CQ variables consistent to assist in potential integration at a later time
- CQs will require the client to be responsible for any MedDRA version updates/changes
- Be aware of the potential change in SMQs through different versions of MedDRA
 - All terms included in a SMQ will have a status so different versions may have a term inactivated over time, but once a term is added to a SMQ it will always remain with that SMQ
 - It is recommended that the use of SMQs is based on the same dictionary as the coded data as mismatches of SMQ and the coded data could result with unexpected flagged records

CONCLUSION

There are multiple benefits through the use of SMQs and CQs when implemented correctly. In planning ahead you can use them to define conditions early on in compound development, grouping terms when applicable for future use. With SMQs, the pre-defined lists are results of an encompassing effort so you can potentially gain insight to using SMQs that can work across multiple TAs and be part of a standardized/consistent implementation process. In the case where a SMQ may not cover all medical topics or safety issues that may be of interest, CQ utilization will be of great benefit to your analysis. As with all things in data standards, an ounce of prevention (thoughtful consideration) can be worth a pound of cure (re-work).

REFERENCES

Introductory Guide for Standardised MedDRA Queries (SMQs) Version 19.1, MSSO-DI-6226-19.1.0, 2016, ICH, Available at http://www.meddra.org/sites/default/files/guidance/file/smq_intguide_19_1_english.pdf

List of SMQ Topics for Development by CIOMS Working Group for SMQs (as of 1 September 2016)
Available at http://www.meddra.org/sites/default/files/page/documents/list_of_smq_topics_for_website_september_2016.pdf

CDISC ADaM Occurrence Data Structure (OCCDS) (Version 1.0) Available at www.cdisc.org

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.