PharmaSUG 2019 - Paper DS-261

Raising the Bar: CDASH implementation in a biometrics CRO

Julie Barenholtz, Cytel Inc.

ABSTRACT

Since its first publication in 2008, CDASH has dramatically improved reliability in clinical trials data. 67% of current CDASH can map directly to SDTM, providing increased traceability between data collection and data analysis. One advantage of having data management and statistical programming under the same roof is the creation of efficient processes that lead to timelier submissions and faster patient access to novel treatments. There are many reasons to implement a CDASH library: having a standard set of CDASH compliant CRFs can greatly reduce the time for CRF design and time to go-live; the CDASH CRF questions and their intentions are comprehensive and widely understood among the industry; and the variable naming is designed in a way that is understood by an SDTM statistical programmer, allowing for a more efficient annotation of the aCRF.pdf and accurate placement of data into the right SDTM domain. Through this paper, I aim to share the benefits of creating both a CDASH library and a strong partnership with statistical programming in order to standardize TLF shells and programs

INTRODUCTION

There are often questions about the benefits of implementing CDASH. After all, CDASH conventions alone are not what are submitted to the FDA. Wouldn't it be faster to just quickly build a database and let your programmers and biostatisticians sort out how to submit the information to the regulatory agencies? There are many reasons to implement a CDASH library at your institution, including better data collection, easier data review, and more efficient submissions. With the release of CDASH v2.0 in 2017, we saw 67% of CDASH map directly to SDTMIG 3.2 variables (including mapping). Additionally, 86% of CDASH v2.0 map directly with standard mappings (e.g. dates). CDASH, or Clinical Data Acquisition Standards Harmonization, provides a way to collect data across studies that creates standardized formats and structures and enables traceability from capture to submission. This paper will describe the importance and use of the CDASH library, CRF examples, and how they relate to SDTM, as well as the use of the library to implement annotation on the BlankCRF and efficient SDTM mapping.

CDASH USEFULNESS AND TRACEABILTY

One of the first things we must do as data managers is to create CRFs based on the protocol we are working with and then provide this to biostatistician and programmer for review. As we know from working with SDTM, we collect much of the same types of data across all protocols and sponsors–demographics, physical exam information, vital signs, adverse events, concomitant medications and procedures, medical history, laboratory data. Because much of what we collect doesn't vary much across projects, we are given the opportunity to standardize how we are asking these questions and what the programming on the back end that will ultimately be used by statistical programmers and biostatisticians looks like. By asking questions in the same way across sites, we increase the reliability of these data and efficiency for sites who are tasked with entering an increasingly larger set of data per protocol. By standardizing the variable names and code lists, we are able to build databases more quickly and clean data more efficiently and reliably.

PRE-CDASH

Before the implementation of CDASH, you may have had case report forms (CRFs) that looked like the following in Table 1, where question text and variable names differed across protocols, and sites, data manager, programmers, and biostatisticians saw a variety of questions and variables. Table 1 shows how questions such as Exam Date and Time on the Physical Exam CRF can be asked six different ways and with six different variable names. Multiply this by the thousands of questions and variables we collect, and it becomes extremely daunting to think about updating both data cleaning programs and SDTM programs in order to get the data into a condition where we can review and submit it.

Prompt	Variable Name
Exam Date	EXAMDATE EXAMDAT EXAMDT EXDATE EXDAT EXDT
Physical Exam Date	PEEXDATE PEEXDAT PEEXDT PEEXAMDT PEDATE
Date of Exam	DATEEXAM DATEXAM DTEXAM DATEEX DATEX DTEX
Date of Physical Exam	DATEPEEX DATPEEX DTPEEX DTPEEXAM DATEPE
Date	DATE DAT
Physical Date	PEDATE PEDAT PEDT
Exam Time	EXAMTIME EXAMTIM EXAMTM EXTIME EXTIM EXTM
Time of Exam	TIMEEXAM TIMEXAM TMEXAM TIMEEX TIMEX TMEX
Time	TIME TIM TM
Physical Exam Time	PEEXAMTM
Time of Physical Exam	TIMEPEEX
Physical Time	PETIME

Table 1. Physical Exam; Example question text and variable name

CDASH COMPLIANT TO SDTM AUTOMATION

By using CDASH standards, we greatly reduce the variability in how questions are asked and variable names are programmed. For the same two questions related to the physical exam date and time we have just two choices, reflected in Table 2. We can either use full question text or prompt text, and the variable name stays the same for either option. We eliminated nearly 40 possibilities by complying with CDASH. We have provided question text familiar to and understood by sites and a variable name that can be used across all projects to mean the same thing, eliminating re-work of data cleaning and SDTM programming. The example below allows statistical programmers and biostatisticians to convert these two fields quickly to the SDTM; DTC variable, --DTC or PEDTC.

Prompt	Variable Name
What was the date of the physical examination?	PEDAT
Exam Date	PEDAT
What was the time of the physical examination?	PETIM
Exam Time	PETIM

Table 2. Physical Exam; CDASH Compliant question text and variable name

CDASH TO SDTM EXAMPLE

Table 3 includes examples where CDASH and SDTM have identical variable names. Again, by programming with CDASH, we can make the work of the SDTM programmer more efficient. With the exception of the date variables, nearly all of the variables collected on the concomitant medication and adverse event CRFs map directly between CDASH and SDTM.

Prompt	CDASH Variable Name	SDTM Variable Name
(Concomitant) [Medication/Treatment/Therapy]	CMTRT	CMTRT
Indication	CMINDC	CMINDC
Dose	CMDSTXT	CMDOSTXT CMDOSE
(Dose) Unit	CMDOSU	CMDOSU
Route	CMROUTE	CMROUTE
Start Date	CMSTDAT	*CMSTDTC*
End Date	CMSTTIM	
Ongoing	CMONGO	CMENRF; CMENRTPT
End Date	CMENDAT	*CMENDTC*
End Time	CMENTIM	
Adverse Event	AETERM	AETERM
Start Date	AESTDAT	*AESTDTC*
Start Time	AESTTIM	
End Date	AEENDAT	*AEENDTC*
End Time	AEENTIM	
Severity	AESEV	AESEV
Serious	AESER	AESER
Death	AESDTH	AESDTH
Life Threatening	AESLIFE	AESLIFE
Hospitalization	AESHOSP	AESHOSP
Disability	AESDISAB	AESDISAB
Congenital Anomaly or Birth Defect	AESCONG	AESCONG
Involves Cancer	AESCAN	AESCAN
Other Serious (Important Medical Events)	AESMIE	AESMIE
Relationship	AEREL	AEREL
Action Taken	AEACN	AEACN
Outcome	AEOUT	AEOUT

Table 3. CDASH – SDTM direct mapping

USING CDASH TO DRIVE SDTM

In addition to naming variables per CDASH, care must also be taken to use the Controlled Terminology. This again follows the same CT used by SDTM and greatly improves the efficiency during mapping from CDASH to SDTM. The CDASH library can be provided to the statistical programming group to automate and semi-automate as much of the mapping as possible, and the metadata extract out of the EDC can be used to create automatic and semi-automatic mappings for SDTM. We can automatically populate nearly 86% of each SDTM domain by understanding the relationship between CDASH and SDTM. By using variable naming in the CDASH library that is compliant with the naming conventions, even when those variables do not yet exist, statistical programmers familiar with SDTM standards will be able to more quickly populate the SDTM domains required for submission. In addition to populating SDTM domains, we are also able to create the SDTM aCRF more efficiently, because using the library forms dictates the variable naming and enables the semi-automation of this process as well.

CONCLUSION

CDISC standards are now widely accepted and required by regulatory agencies. Getting a CDASH library in place early and in collaboration with your statistical programmers and biostatistician will greatly improve the way your sites collect data, how the data are cleaned, and how quickly data can be submitted to the regulatory agencies. CDASH to SDTM improves the reliability and traceability of clinical trial data collected. Our primary goal is always to bring lifesaving medications and other treatments to patients as quickly and safely as possible, and implementing CDASH to SDTM in your organization will help to do just that.

ACKNOWLEDGEMENTS

Many thanks to Angelo Tinazzi Senior Director, Statistical Programming for collaborating in the CDASH/SDTM project and for input on the statistical programming pieces of this paper.

REFERENCES

Clinical Data Interchange Standards Consortium (https://www.cdisc.org/)

K. Howard and S. Labout. 2017. “SDTM vs CDASH: Why You Need Both!” https://www.cdisc.org/sites/default/files/resource/CDISC_SDTM_and_CDASH_You_Need_Both.pdf

CONTACT INFORMATION

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

Name: Julie Barenholtz
Enterprise: Cytel Inc.
Work Phone: 617-528-7106
E-mail: Julie.Barenholtz@cytel.com