

PharmaSUG China 3<sup>rd</sup> Conference, 2014  
**Brief Introduction of Oncology Domains in SDTMIG,**  
**Version 3.2**  
Haishan Kadeerbai, PPD Inc., Beijing

## **ABSTRACT**

The final document of CDISC Study Data Tabulation Model Implementation Guide (SDTMIG) v3.2 was released in the end of 2013. All of previous annotations in SDTMIG 3.1.3, originally published in 2012, have been incorporated in SDTMIG v3.2 that includes many enhancements and improvements from earlier versions. This paper presents a brief introduction on the oncology domains (TU, TR and RS) in Findings Observation Class, which are extremely useful in oncology clinical trials, mainly focused on: 1) the purpose and usage of the domains; 2) relationship among the domains; 3) the domains' cross-data (domains) linkability and relationship.

## **INTRODUCTION**

Previous SDTMIG, v3.1.3, was presented as an annotated version of SDTMIG 3.1.2 and the inclusion of the TU/TR/RS Oncology domains to the Findings General Observation Class were developed as the Oncology Disease-specific Therapeutic Area Supplement, and now they are officially incorporated in the new SDTMIG v3.2 portfolio. The primary reason why CDISC includes these three domains is that assessment of the change in tumor burden is an important feature of the clinical evaluation of cancer therapeutics: both tumor shrinkage (objective response) and disease progression are useful endpoints in cancer clinical trials. The tumor package domains exclusively represent data collected in oncology clinical trials which would be applicable for representing data to support assessment criteria such as RECIST (solid tumors), Cheson2 (e.g. lymphoma), or, Hallek3(3) (chronic lymphocytic leukemia), but users should keep it in mind that the three domains were developed specifically with RECIST in mind at the very first beginning, though RECIST is simple, using RECIST is not. Some response criteria may require additional data for an assessment of response and that data should reside in the appropriate SDTM domain (e.g. LB for lab test results, MI for microscopic examination details). CDISC recommends that to use these three domains as a whole for they are closely related, of which approach is consistent with current SDTM philosophy though each domain has a distinct purpose.

## **PURPOSES**

**The Tumor Identification (TU)** - domain represents data that uniquely identifies a tumor for continual tracking purposes.

**The Tumor Response (TR)** - domain represents quantitative measurements and/or qualitative assessments of the tumors identified in the TU domain, and these measurements are usually taken at baseline, then, at each subsequent visit to support response evaluations.

**The Disease Response (RS)** - domain represents the response evaluation(s) determined from the data in TR, and data from other sources (in other SDTM domains, e.g. LB) might also be used in an assessment of response.

## **ASSUMPTIONS**

### **TU**

The tumors are identified by an investigator and/or independent assessor and classified according to the disease assessment criteria. A record in the TU domain may contain the following information: a unique tumor ID value (TULNKID); anatomical location of the tumor (TULOC); method used to identify the tumor (TUMETHOD); role of the individual identifying the tumor (TUEVAL); and timing information (VISITNUM, EPOCH, TUDTC); location of the tumor (TULOC) and additional anatomical location variables (TULAT, TUDIR, TUPORTOT) which allow for more detailed information to be collected that further clarifies the value of the TULOC variable.

Normally the initial identification of a tumor should be done once at baseline (to identify as Target and Non-Target tumors in terms of RECIST) and then subsequently at post-baseline visits (may identify as New tumors in addition of the aforementioned two types), and the identification information, including the location description, must not be repeated for every visit for this kind of data would not contribute any informative insights and may be redundant to the analysis of the efficacy of the trial treatments. Therefore, when include post-baseline records the most recommended scenarios are: 1) a new tumor emerges; 2) baseline-identified tumor subsequently splits into separate distinct tumors

and additional records should be created to distinctly identify the split tumors; 3) baseline-identified tumors subsequently merge together then an additional record should be included to distinctly identify the merged tumor; 4) a separate set of records might be identified and represented when re-baseline of Targets and Non-Targets is required (e.g. cross-over studies).

The approach that measurements of each distinct tumor must be captured individually is preferred to follow the design of current oncology domains, but if the measurements of split/merge tumors are reported as a summary under the 'parent' tumor at some cases, then it may not be possible to include a record in the TU domain. In this situation, the assessments of split and merge tumors would be represented only in the TR domain

## TR

A record in the TR domain contains the following information: a unique tumor ID value (TRLNKID); test and result (TRTEST, TRORRES); method used (TRMETHOD); role of the individual assessing the tumor (TREVAL); and timing information (VISITNUM, EPOCH, TRDTC). The TR domain does not include anatomical location information because this it is already represented in TU. When a tumor is too small to measure per assessment criteria and a default value can be recorded for that assessment (e.g. in studies using RECIST, a default value of 5mm is used and expected in TRSTERSC & TRSTRESN, in the calculation when determining response).

## RS

Data collected in RS is the essential information for the assessment of a tumor's response to a clinical trial drug, it usually would be summarized to present when cancer patients improve ("respond"), stay the same ("stable") or worsen ("progression") during treatments. A record in the RS domain contains the following information: a unique tumor ID value (RSLNKID); response assessment result (RSORRES); role of the individual assessing the tumor (RSEVAL); timing information (VISITNUM, EPOCH, RSDTC); and criteria used in the assessment of response (RSCAT). If the assessment of response was based on other evidence that was not defined as objective evidence in the criteria being used to evaluate response per protocol then "CLINICAL ASSESSMENT" as a value for RSCAT is used to represent the situation.

### Certain Commonalities

- --TESTCD/--TEST values for the three domains are published as Controlled Terminology. Corresponding data would be included in the domain only if those data points have been collected on a CRF or have been supplied by an external assessor as part of an electronic data transfer (eDT). It is not intended that the sponsor would create derived records to supply those values in these domains. Derived records/results should be provided in the analysis dataset.
- The Acceptance Flag variable (--ACPTFL) identifies those records that have been determined to be the accepted assessments/measurements by an independent assessor. This flag would be provided by an independent assessor and when multiple assessors (e.g. RADIOLOGIST 1, RADIOLOGIST 2, ADJUDICATOR) provide assessments or evaluations at the same time point or an overall evaluation. This flag should not be used by a sponsor for any other purpose i.e. it is not expected that the TUACPTFL flag would be populated by the sponsor. Instead that type of record selection should be handled in the analysis dataset.
- The Evaluator Specified variable (--EVALID) is also subject to Controlled Terminology and used in conjunction with --EVAL to provide additional detail of who is providing tumor identification information. --EVAL must also be populated when --EVALID is populated. (e.g.--EVAL="INDEPENDENT ASSESSOR" and --EVALID="RADIOLOGIST 1")

## RELATIONSHIP AND LINKABILITY

The three oncology domains are closely related, and it is strongly recommended to use them as a whole as per SDTMIG instructed. The mechanism of how these three are related/linked can be briefly summarized as: 1) TULNKID/TRLNKID pair is used to relate an identification record in TU domain to assessment records in the TR domain. It's more like a point to point linkage. All new tumors are to be represented in TU and TR domains correspondently; 2) TRLNKGRP/RSLNKGRP pair is used to relate records in the TR domain to a response assessment record in RS domain, usually link the record having RSTEST="Overall Response" in the RS to TR as a one to many linkage.

A RELREC relationship record can be created to describe the link between the three domains, mostly as a dataset to dataset link, and a dataset to dataset link would be the most appropriate linking mechanism, in which way a cross-data (domains) linkability can be built to complete the relation loop not only just among the three, but also link them to any other SDTM domains.

Utilizing one of the existing ID variables is not possible because all three (GRPID, REFID & SPID) may be used for other purposes per SDTM. The –LNKID/--LNKGRP variable is used for values that support a RELREC dataset to dataset relationship and to provide a unique code for each identified tumor, and for each response and associated tumor measurements/assessments, hence build a comprehensive relationship between the three oncology domains to ensure their organic integrity.

## CONCLUSION

Oncology domains (TU, TR and RS) in Findings Observation Class in SDTMIG v3.2 are extremely useful in oncology clinical trials. They are designed as a package with specially RECIST in mind. Apply strictly the philosophy and the instructions when conducting oncology studies can yield high efficiency, precision and productivity.

## REFERENCES

- CDISC, Study Data Tabulation Model v1.4 <http://www.cdisc.org/sdtm>
- SDTM Implementation Guide v3.1.3, v3.2 <http://www.cdisc.org/sdtm>
- Response Evaluation Criteria In Solid Tumors <http://www.eortc.org/investigators-area/recist>

## CONTACT INFORMATION

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

Name: Haishan Kadeerbai  
Enterprise: PPD Inc.  
Address: 8/F, Tower B, Central Point Plaza No 11, Dongzhimen South Ave. Dongcheng District  
City, State ZIP: Beijing, 100007  
Work Phone: +86 10-57636486  
Fax: +86 10-57636251  
E-mail: Haishan.Kadeerbai@ppdi.com  
Web: www.ppdi.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.