SDTM - Tips and Tricks Oncology Domains

Created By : Abhijit Sen (Cognizant)/Godfrey Machado (TCS)
Abhijit Sen – 8.5 years of experience in Statistical Programming. CDISC SME.
Godfrey Machado – Around 10 years of experience in Statistical Programming, CDISC SME
05-Dec-2015

Index

Table of Content
Context Setting: Overview
Type of Oncology Data and Analysis
Understand Solid Tumor Data collection per RECIST
RECIST Data Collection
Oncology Specific SDTM Domains
TU, TR, RS
Challenges in SDTM conversion
Benefits in SDTM conversion
CRF Data to SDTM
Adjudication

Context Setting: Overview

- Introduction:

- In the last decade the oncology therapeutic area has grown strongly and becomes one of the largest therapeutic areas within the clinical research field despite the complexity and risks with regards to uncertain trial endpoints.

- Clinical Data Interchange Standards Consortium (CDISC) assists in submitting tabulation data to the U.S. Food and Drug Administration (FDA). While it creates opportunities to standardize data structure, transforming various clinical data using CDISC SDTM poses significant challenges.

- Response data is one of the key efficacy measurements for oncology trials. There are generally two types of efficacy analysis for oncology trials that require response endpoint data: response analysis and time-to-event analysis.

- One of the advantages of the SDTM model is that it defines data structure and is not dependent on individual vendor's system. Therefore SDTM provides an opportunity to facilitate data exchange between vendor and sponsor using a single standard.

Objectives

- In the end of this presentation you will be able to:

- Understand the information capture on the tumor lesions and the disease response through Study Data Tabulation Model Implementation Guide (SDTMIG) v3.2.

- How the three domains are related but each has a distinct purpose.

- Real-life examples how the information of the standardized response criteria can be collected in each of the three finding domain classes.

Global Oncology Market Continues to Grow $35B Today and ~$100+B by 2020

Source: IMS Health, MIDAS, MAT Mar 2010.
**Oncology includes both solid & heme tumors and is defined by IMS as L1 (cytotoxics & targeted cytotoxics) and L2 (hormonals).*

Type of Oncology Data and Analysis

- Non-solid Tumor (leukemia and lymphoma)
  - Lab data analysis.

- Solid Tumor :
  - WHO Tumor Response Criteria
  - SWOG Criteria
  - RECIST (Response Evaluation Criteria in Solid Tumor)

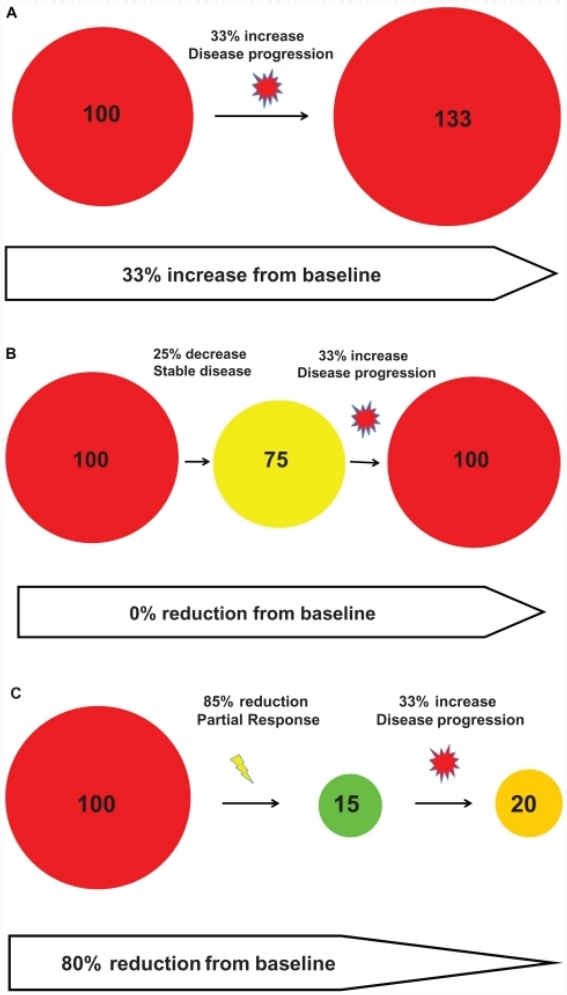
Type of Oncology Data and Analysis

- Non-solid Tumor (leukemia and lymphoma)
  - No difference from handling regular Lab data.

- Solid Tumor :
  - RECIST became the standard guideline. Data collection and analysis could be very challenging.

Understand Solid Tumor Data collection per RECIST

- At baseline, tumors are categorized as
  - Target Lesion / Measurable Lesion
  - Non-Target Lesion / Non-Measurable Lesion



RECIST Data Collection Instrument --- CRF

To make it easier to be understood, the following is a simplified version of RECIST CRF

Target lesions			***Give measurement of longest diameter***			
	Assessment/Date →		Baseline	Cycle _2_	Cycle _4_	Cycle _6_
#	Site	Method	YY-MMM-DD ___-_____-___	YY-MMM-DD ___-_____-___	YY-MMM-DD ___-_____-___	YY-MMM-DD ___-_____-___
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
	Overall Sum →					
Non-Target lesions	**Assess as present, absent, worsen or not done**					
	Overall response →		----------			

Solid Tumor Data Analysis Parameters per RECIST

- Best Overall Response
- The objective response rate (the proportion of CR or PR in terms of best overall response)
- Time to PD
- Time to PFS (progression-free survival)
- Time to Response
- Time to Treatment Failure

All these parameters depend on overall assessment at each cycle.

Oncology Specific SDTM Domains

- The oncology specific SDTM domains were first introduced in SDTMIG v3.1.3 in July 2012.
- The domains, TU, TR, RS are intended to represent data collected in clinical trials where tumors or lymph nodes are identified at baseline visits and then repeatedly measured or assessed at subsequent time points.
- The results of the measurements and assessments are used in the evaluation of the disease response.
- As such these domains 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)

New Variables!!!!

--LINKID

- – to link records between TU and TR
- – as GRPID, REFID and SPID are already needed

--ACPTFL

- – acceptance flag (independent assessor)

--EVALID

- – evaluator ID
- – to be used in combination with –EVAL

New Variables!!!!

--LAT

- – qualifier for anatomical location detailing laterality
- – e.g. "LEFT", "RIGHT", "BILATERAL“

--DIR

- – qualifier for anatomical location further detailing directionality
- – e.g. "UPPER", "ANTERIOR", "DORSAL", "PROXIMAL“

--DSTRBN

- – qualifier for anatomical location or specimen further detailing the distribution
- – e.g. "ENTIRE", "SINGLE", "SEGMENT", "MANY"

Oncology specific AE

CTCAE grade

*AETOXGR,C2,$AETOXGR

☐ 1 Mild AE

☐ 2 Moderate AE

☐ 3 Severe AE

☐ 4 Life threatening or Disabling AE

☐ 5 Death related to AE

AE change date#

AEC01DAT,C10

year mmm dd

CTCAE grade change#

AEC01TOX,C2,$AETOXGR

☐ 1 Mild AE

☐ 2 Moderate AE

☐ 3 Severe AE

☐ 4 Life threatening or Disabling AE

☐ 5 Death related to AE

Oncology specific AE

Possible Solution!!!

General recommendations

EVAL - This column can be left Null when the Investigator provides the complete set of data in the domain. However the column should contain no Null values when data from one or more independent assessors is included meaning that the rows attributed to the Investigator should contain a value of INVESTIGATOR.

EVALID - The Evaluator Specified variable is used in conjunction with TUEVAL to provide an additional level of detail. When multiple assessors play the role identified in TUEVAL, values of TUEVALID will attribute a row of data to a particular assessor.

TRTESTCD / TRTEST values for this domain are published as Controlled Terminology. The sponsor should not derive results for any test if the result was not collected.

While running **OPEN CDISC VALIDATOR** tool on your SDTM data, make sure you select the correct version of SDTM (3.1.3 onwards) and CT in order to avoid error in your report.

Tumor Identification (TU)

The TU domain represents data that uniquely identifies tumors (i.e. malignant tumors and other sites of disease, e.g. lymph nodes). The tumors are identified by an investigator and/or independent assessor and classified according to the disease assessment criteria. In RECIST terms this equates to the identification of Target, Non-Target or New tumors. A record in the TU domain contains the following information: a unique tumor ID value; anatomical location of the tumor; method used to identify the tumor; role of the individual identifying the tumor; and timing information.

TUGRPID	TULNKID	TUTESTCD	TUTEST	TUORRES	TULOC	TUMETHOD	VISIT	TUDTC
A1	A1	TUMIDENT	Tumor Identification	ASSESSABLE TUMOR LESION	ADENOPATHY CERVICAL	SPIRAL CT	SCREENING	2012-11-01
M1	M1	TUMIDENT	Tumor Identification	MEASURABLE TUMOR LESION	ADENOPATHY INGUINAL	SPIRAL CT	SCREENING	2012-11-01
M2	M2	TUMIDENT	Tumor Identification	MEASURABLE TUMOR LESION	ADENOPATHY PARATRACHEAL	SPIRAL CT	SCREENING	2012-11-01
N1	N1	TUMIDENT	Tumor Identification	NEW	ADENOPATHY MESENTERIC	SPIRAL CT	VISIT 3	2013-06-11

Tips and Tricks - Tumor Identification (TU)

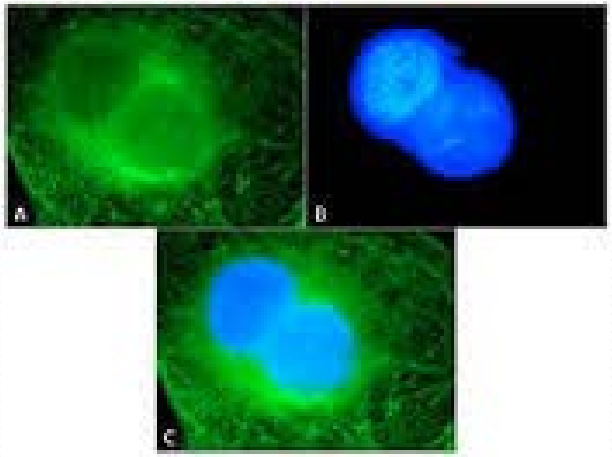
The initial identification of a tumor is done once, usually at baseline (e.g. the identification of Target and Non-Target tumors). A post-baseline records might be included in the TU domain when:

- A new tumor may emerge at any time during a study.
- If a tumor identified at baseline subsequently splits into separate distinct tumors
- If one or more tumors identified at baseline subsequently merge together
- A re-baseline of Targets and Non-Targets is required (e.g. a cross-over study)

Splitting of Tumors

TU

Merging of Tumors



Tumor Response (TR)

The TR domain represents quantitative measurements and/or qualitative assessments of the tumors i.e. malignant tumors and other sites of disease, e.g. lymph nodes) identified in the TU domain. These measurements are usually taken at baseline and then at each subsequent assessment to support response evaluations. A record in the TR domain contains the following information: a unique tumor ID value; test and result; method used; role of the individual a

TUGRPID	TULNKID	TUTESTCD	TUTEST	TUORRES	TULOC	TUMETHOD	VISIT	TUDTC
A1	A1	TUMIDENT	Tumor Identification	ASSESSABLE TUMOR LESION	ADENOPATHY CERVICAL	SPIRAL CT	SCREENING	2012-11-01

TRLNKID	TRTESTCD	TRTEST	TRORRES	TRMETHOD	VISIT	TRDTC
A1	LESSTAT	Lesion Status	NO CHANGE	SPIRAL CT	VISIT 1	2013-01-25
A1	LESSTAT	Lesion Status	NOT EVALUABLE	SPIRAL CT	VISIT 2	2013-03-26
A1	LESSTAT	Lesion Status	DECREASED	SPIRAL CT	VISIT 3	2013-05-13

The TR domain does not include anatomical location information on each measurement record because this would be a duplication of information already represented in TU. This duplication of data was a deciding factor in multi-domain approach to representing this data.

Tips and Tricks - Tumor Response (TR)

The TR domain represents quantitative measurements and/or qualitative assessments of the tumors (i.e. malignant tumors and other sites of disease, e.g. lymph nodes) identified in the TU domain. These measurements are usually taken at baseline and then at each subsequent assessment to support response evaluations.

- TRLNKID is used to relate records in the TR domain to an identification record in TU domain.
- TRLNKGRP is used to relate records in the TR domain to a response assessment record in RS domain. The organization of data across the TR and RS domains requires a RELREC relationship to link the related data rows.
- TRTESTCD / TRTEST values for this domain are published as Controlled Terminology. The sponsor should not derive results for any test (e.g. “Percent Change From Nadir in Sum of Diameter”) if the result was not collected.
- When a tumor has split or merged, assessments will be recorded for the new records created in the TU domain.

TR

Disease Response (RS)

The RS domain represents the response evaluation(s) determined from the data in TR. Data from other sources (In other SDTM domains) might also be used in an assessment of response.

TRLNKGRP	TRLNKID	TRTESTCD	TRTEST	TRORRES	TRORRESU	TRMETHOD	TRACPTFL	VISIT
R2-V1	R2-101	LDIAM	Longest Diameter	104.49	mm	PET	Y	SCREENING
R2-V1	R2-101	PPD	Product Perpendicular Diameters	7245.34	mm2	PET	Y	SCREENING
R2-V2	R2-101	LDIAM	Longest Diameter	77.75	mm	SPIRAL CT	Y	VISIT 1
R2-V2	R2-101	PPD	Product Perpendicular Diameters	2506.66	mm2	SPIRAL CT	Y	VISIT 1
R2-V2	R2-501	TUMSTATE	Tumor State	PRESENT		SPIRAL CT	Y	VISIT 1
R2-V2	R2-801	TUMSTATE	Tumor State	PRESENT		SPIRAL CT	Y	VISIT 1

RSLNKGRP	RSTEST	RSCAT	RSORRES	RSEVAL	RSEVALID	RSACPTFL	VISIT
R1-V2	Overall Response	CHESON 2007	SD	RADIOLOGIST	RADIOLOGIST 1	N	VISIT 1
R2-V2	Overall Response	CHESON 2007	PD	RADIOLOGIST	RADIOLOGIST 2	Y	VISIT 1

Disease Response (RS) cntd….

For some RSTEST's, RSLNKGRP is empty. Best Overall Response is an evaluation of response determined on patient level and independent of the visit. For this reason RSLNKGRP and VISIT are empty for this RSTEST. For each response grouped on type of tumor RSLNKGRP is also empty as RSLNKGRP is used to link to all of the measurements and assessments in the TR domain per time point

RSLNKGRP	RSTEST	RSCAT	RSORRES	RSSTAT	RSREASND	RSEVAL	VISIT
	Assessable Lesion Response	CHESON 2007	SD			RADIOLOGIST	VISIT 1
	Measurable Lesion Response	CHESON 2007	PR			RADIOLOGIST	VISIT 1
	Liver Response	CHESON 2007	CR			RADIOLOGIST	VISIT 1
	Spleen Response	CHESON 2007	CR			RADIOLOGIST	VISIT 1
	Best Overall Response	CHESON 2007	PR			RADIOLOGIST	

Tips and Tricks – Disease Response (RS)

- RSTESTCD / RSTEST values for this domain are published as Controlled Terminology.
- RSCAT is used to identify the criteria used in the assessment of response and a version number if appropriate. “CLINICAL ASSESSMENT” is used to represent the situation when 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.
- The Evaluator Specified variable (RSEVALID) is used in conjunction with RSEVAL to provide additional detail of who is providing the response assessments. For example RSEVAL=”INDEPENDENT ASSESSOR” and RSEVALID=”RADIOLOGIST 1”. The RSEVALID variable is to Controlled Terminology. RSEVAL must also be populated when RSEVALID is populated.
- The Acceptance Flag variable (RSACPTFL) identifies those records that have been determined to be the accepted assessments/measurements by an independent assessor.

RS

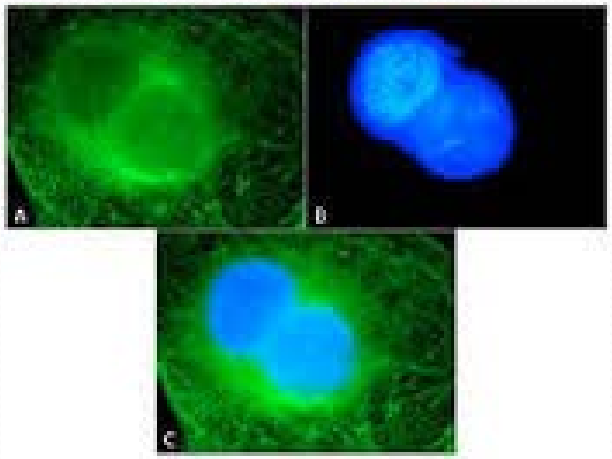
Challenges in SDTM conversion

Main Challenges:

- Dynamic Oncology data (Splitting/Merging of Tumors)
- Adjudication

Splitting of Tumors

Merging of Tumors



Example: TU Domain

TUSEQ	TUGRPID	TUREFID	TUSPID	TULNKID	TUTESTC	TUTEST	TUORRES	TUSTRESC
4		IMG-00001	TRG-SCR-04	T04	TUMIDENT	Tumor Identification	TARGET	TARGET
5		IMG-00001	NTG-SCR-01	NT01	TUMIDENT	Tumor Identification	NON-TARGET	NON-TARGET
6		IMG-00001	NTG-SCR-02	NT02	TUMIDENT	Tumor Identification	NON-TARGET	NON-TARGET
7	T04	IMG-00002	TRG-W16-07	T04.1	TUSPLIT	Tumor Split	TARGET	TARGET
8	T04	IMG-00002	TRG-W16-08	T04.2	TUSPLIT	Tumor Split	TARGET	TARGET
9	T02/T03	IMG-00003	TRG-W24-09	T02/T03	TUMERGE	Tumor Merged	TARGET	TARGET
10		IMG-00004	NWT-W32-01	NEW01	TUMIDENT	Tumor Identification	NEW	NEW

Example: TR Domain

D	E	F	G	H	I	J	K	P	Q
TRSEQ	TRGRPID	TRREFID	TRLNKGRP	TRLNKID	TRTESTCD	TRTEST	TRORRES	TRSTAT	TRREASND
11	TARGET		R-A2	T04	LDIAM	Longest Diameter	65		
12	TARGET		R-A2		SUMLDIAM	Sum of Longest Diameter	125		
13	NON-TARGET		R-A2	NT01	TUMSTATE	Tumor State	PRESENT		
14	NON-TARGET		R-A2	NT02	TUMSTATE	Tumor State	ABSENT		
15	TARGET		R-A3	T01	LDIAM	Longest Diameter	TOO SMALL TO MEASURE		
16	TARGET		R-A3	T02	LDIAM	Longest Diameter	12		
17	TARGET		R-A3	T03	LDIAM	Longest Diameter	15		
18	TARGET		R-A3	T04	LDIAM	Longest Diameter		NOT DONE	TUMOR SPLIT OR DIVIDED
19	TARGET		R-A3	T04.1	LDIAM	Longest Diameter	30		
20	TARGET		R-A3	T04.2	LDIAM	Longest Diameter	10		
21	TARGET		R-A3		SUMLDIAM	Sum of Longest Diameter	72		
22	NON-TARGET		R-A3	NT01	TUMSTATE	Tumor State	PRESENT		
23	NON-TARGET		R-A3	NT02	TUMSTATE	Tumor State	ABSENT		
24	TARGET		R-A4	T01	LDIAM	Longest Diameter	0		
25	TARGET		R-A4	T02	LDIAM	Longest Diameter		NOT DONE	TUMOR MERGED OR COALESCED
26	TARGET		R-A4	T03	LDIAM	Longest Diameter		NOT DONE	TUMOR MERGED OR COALESCED
27	TARGET		R-A4	T02/T03	LDIAM	Longest Diameter	20		
28	TARGET		R-A4	T04	LDIAM	Longest Diameter		NOT DONE	TUMOR SPLIT OR DIVIDED
29	TARGET		R-A4	T04.1	LDIAM	Longest Diameter	30		
30	TARGET		R-A4	T04.2	LDIAM	Longest Diameter	6		
31	TARGET		R-A4		SUMLDIAM	Sum of Longest Diameter	56		

Example: RS

D	E	F	G	H	I	J	K	L
SSEQ	RSGRPI	RSLNKGR	RSLNKID	RSTESTCD	RSTEST	RSCAT	RSORRES	RSSTRESC
1				TRGRESP	Target Response	RECIST 1.0	SD	SD
2				NTRGRESP	Non-target Response	RECIST 1.0	SD	SD
3		R-A2		OVRLRESP	Overall Response	RECIST 1.0	SD	SD
4				TRGRESP	Target Response	RECIST 1.0	PR	PR
5				NTRGRESP	Non-target Response	RECIST 1.0	SD	SD
6		R-A3		OVRLRESP	Overall Response	RECIST 1.0	PR	PR
7				TRGRESP	Target Response	RECIST 1.0	PR	PR
8				NTRGRESP	Non-target Response	RECIST 1.0	SD	SD
9		R-A4		OVRLRESP	Overall Response	RECIST 1.0	PR	PR
10				TRGRESP	Target Response	RECIST 1.0	PD	PD
11				NTRGRESP	Non-target Response	RECIST 1.0	PD	PD
							Increased shortness of	

Adjudication

Take RECIST idea to another level - adjudication

It is becoming more and more common that Companies use adjudicators to do tumor assessments independently from trial investigators.

In practice, there are 2 adjudicators, if theirs assessments are not consistent, then 3rd adjudicator will step in and provide his/her option that which adjudicator he/she would agree with.

The challenge for Biostat & Programming is :

1. To compare the assessments from the first 2 adjudicators.
2. If they are different, then compare them with the 3rd adjudicator's assessments.
3. The agreed assessments will be used for analysis.

TU example

TUGRPID	TULNKID	TUTESTCD	TUTEST	TUORRES	TULOC	TUMETHOD	TUEVAL	VISIT
A1	A1	TUMIDENT	Tumor Identification	ASSESSABLE TUMOR LESION	ADENOPATHY CERVICAL	SPIRAL CT	INVESTIGATOR	SCREENING
M2	M2	TUMIDENT	Tumor Identification	MEASURABLE TUMOR LESION	ADENOPATHY PARATRACHEAL	SPIRAL CT	INVESTIGATOR	SCREENING
R1-101	R1-101	TUMIDENT	Tumor Identification	MEASURABLE TUMOR LESION	ADENOPATHY INGUINAL	PET	RADIOLOGIST	SCREENING
R1-501	R1-501	TUMIDENT	Tumor Identification	ASSESSABLE TUMOR LESION	ADENOPATHY INGUINAL	PET	RADIOLOGIST	SCREENING

TUREFID	TULNKID	TUTESTCD	TUORRES	TULOC	TUMETHOD	TUEVAL	TUEVALID	TUACPTFL	VISIT
27595	R1-101	TUMIDENT	MEASURABLE TUMOR LESION	ADENOPATHY INGUINAL	PET	RADIOLOGIST	RADIOLOGIST 1	N	SCREENING
27595	R1-102	TUMIDENT	MEASURABLE TUMOR LESION	ADENOPATHY ABDOMINAL	PET	RADIOLOGIST	RADIOLOGIST 1	N	SCREENING
27595	R1-501	TUMIDENT	ASSESSABLE TUMOR LESION	ADENOPATHY INGUINAL	PET	RADIOLOGIST	RADIOLOGIST 1	N	SCREENING
27595	R1-502	TUMIDENT	ASSESSABLE TUMOR LESION	ADENOPATHY COMMON ILIAC	PET	RADIOLOGIST	RADIOLOGIST 1	N	SCREENING
31436	R1-801	TUMIDENT	NEW	OTHER	SPIRAL CT	RADIOLOGIST	RADIOLOGIST 1	N	VISIT 2
27595	R2-101	TUMIDENT	MEASURABLE TUMOR LESION	ADENOPATHY ABDOMINAL	PET	RADIOLOGIST	RADIOLOGIST 2	Y	SCREENING
27595	R2-102	TUMIDENT	MEASURABLE TUMOR LESION	ADENOPATHY MEDIASTINAL	PET	RADIOLOGIST	RADIOLOGIST 2	Y	SCREENING
27595	R2-501	TUMIDENT	ASSESSABLE TUMOR LESION	ADENOPATHY MEDIASTINAL	PET	RADIOLOGIST	RADIOLOGIST 2	Y	SCREENING
32693	R2-801	TUMIDENT	NEW	SUBCUTANEOUS MASS	SPIRAL CT	RADIOLOGIST	RADIOLOGIST 2	Y	VISIT 1

	TRLNKGRP	TRLNKID	TRTESTCD	TRTEST	TRORRES	TRORRESU	TRMETHOD	TRACPTFL	VISIT
M	R2-V1	R2-101	LDIAM	Longest Diameter	104.49	mm	PET	Y	SCREENING
M	R2-V1	R2-101	PPD	Product Perpendicular Diameters	7245.34	mm2	PET	Y	SCREENING
M	R2-V2	R2-101	LDIAM	Longest Diameter	77.75	mm	SPIRAL CT	Y	VISIT 1
M	R2-V2	R2-101	PPD	Product Perpendicular Diameters	2506.68	mm2	SPIRAL CT	Y	VISIT 1
A	R2-V2	R2-501	TUMSTATE	Tumor State	PRESENT		SPIRAL CT	Y	VISIT 1

Benefits in SDTM conversion

Mapping Benefits:

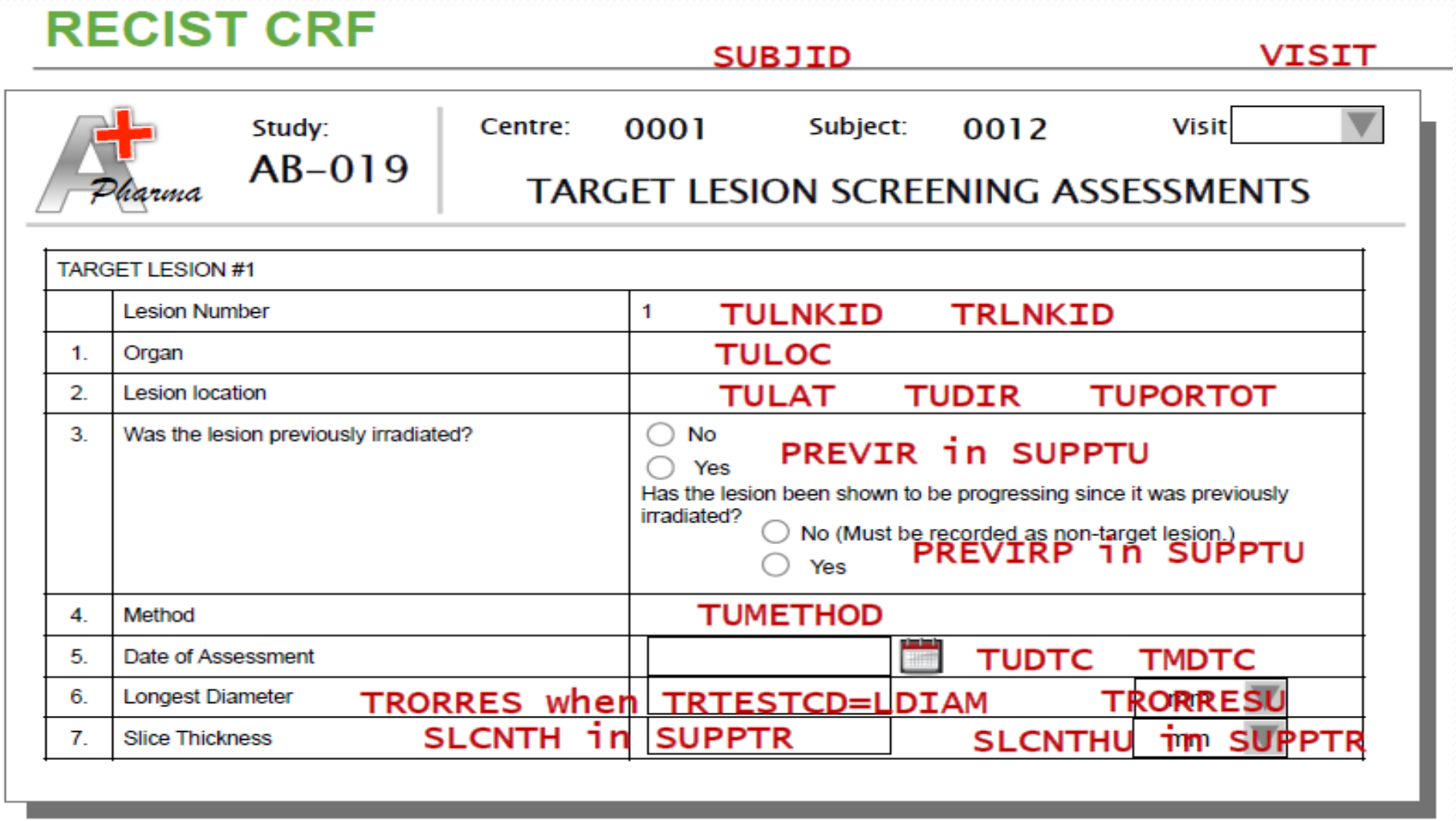
- This proposed domains allows sponsors to integrate response data easily from all sources: external and site collected investigator data.

- There is minimal post processing required for further analysis since the format is nearly analysis ready and very similar to the final ADaM dataset format. In this model, the traceability is achieved by augmenting the analysis data with the rows of data from which the derived overall response is determined and adding variables to indicate the domain, variable and sequence number of the source SDTM data.

- It satisfies normalization principle of the relational databases. This RS domain focuses on the basic SDTM structure and avoids the use of supplemental qualifiers in an attempt to streamline the external data transfer from vendor to sponsor.

- This proposal of a standard response domain eliminates the need for sponsors to remap data from each vendor and provides a transfer standard to be adopted by industry.

CRF Data to SDTM



CRF Data to SDTM

RECIST CRF

SUBJID VISIT

Study: AB–019 | Centre: 0001 | Subject: 0012 | Visit

CLINICAL VISIT RESPONSE

RESPONSE SUMMARY		
	Target Lesion Response Assessment	○ Complete Response ○ Partial Response ○ Stable Disease ○ Progressive Disease ○ Not Evaluable RSORRES when RSTESTCD=TRGRESP
	Non-Target Lesion Response Assessment	○ Complete Response ○ Non-CR / Non-PD ○ Progressive Disease ○ Not Evaluable ○ Not Applicable RSORRES when RSTESTCD=NTRGRESP
	Were there any new lesions since the last assessment?	○ No ○ Yes TRORRES when TRTESTCD=TUMSTATE
	Investigator Overall Response using RECIST 1.0	○ Complete Response ○ Partial Response ○ Stable Disease ○ Progressive Disease ○ Not Evaluable RSORRES when RSTESTCD=OVRLRESP

References:

- New response evaluation criteria in solid tumours:

Revised RECIST guideline (version 1.1)

- SDTMIG 3.2
- Implementation of Oncology Specific SDTM domains (Jacintha Eben, SGS Life Science Services, Mechelen, Belgium)
- SDTM examples for Oncology use cases.

THANK
YOU