



Looking for SDTM migration specialist



- CD13

London – 12-15/10/2014

Angelo Tinazzi
Cytel Inc., Wilmington Del. USA
Succursale de Meyrin – Geneva – Switzerland
angelo.tinazzi@cytel.com

The information contained in this presentation is based on personal research of the author and does not necessarily represent Cytel Inc.

CDISC Standards

CDISC Standards

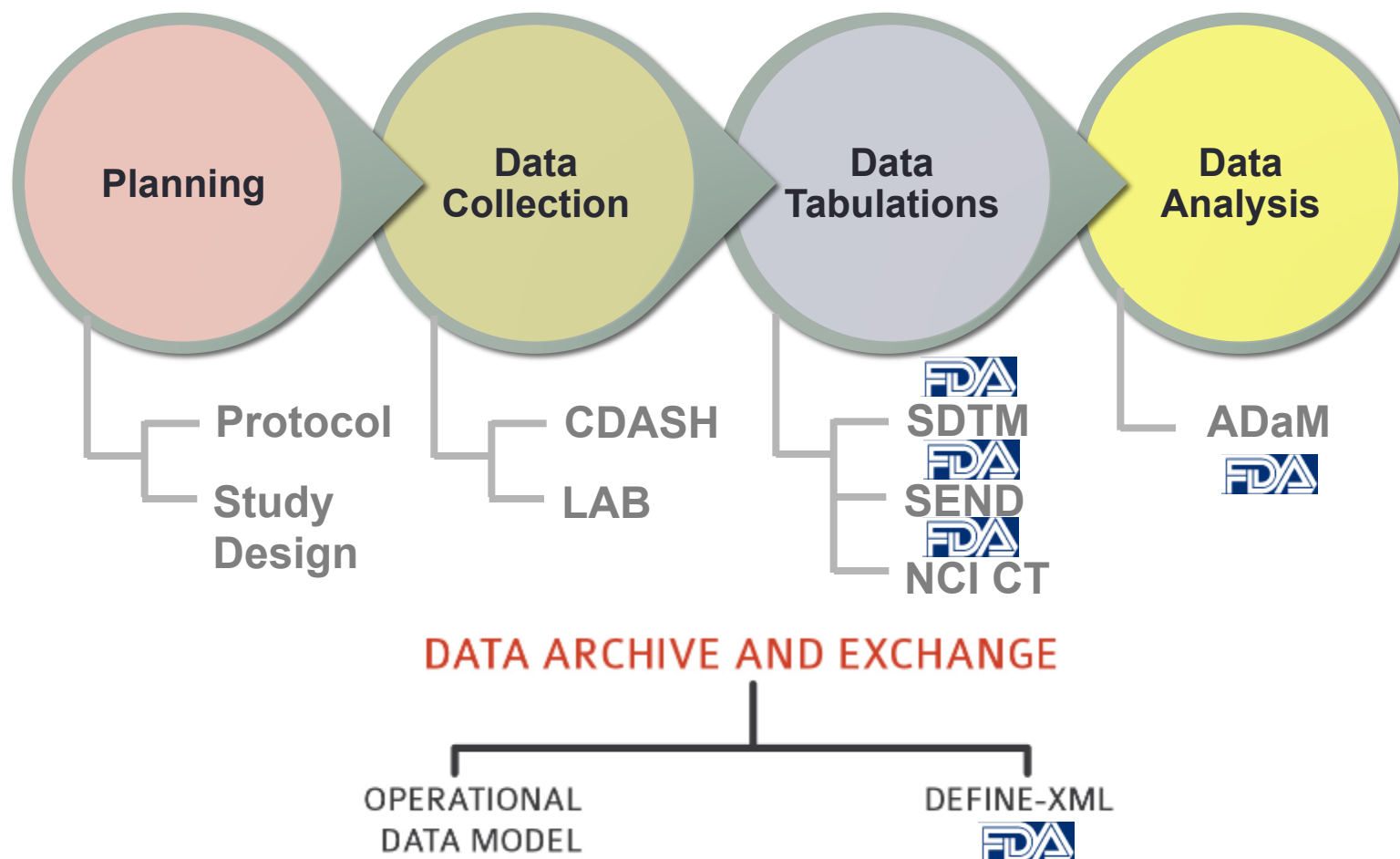
Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions



CDISC Standards

Key References

Document / Guidance

Most recent SDTM released standard guidance:

- CDISC Study Data Tabulation Model (SDTM) v1.4
- Study Data Tabulation Model Implementation Guide (SDTMIG) v3.2
- Associated Persons Implementation Guide (SDTMIG-AP) v1.0

Released Therapeutic Area Standards:

- Diabetes, Alzheimer, Asthma, Multiple Sclerosis, Pain, Parkinson Disease, Polycystic Kidney Disease, Tuberculosis and Virology
- Oncology for Tumor Response domains (integrated into SDTMIG 3.1.3)

CDISC/NCI-EVE Standard Controlled Terminology

FDA Guidance and Technical Documents:

- CDER and CBER: Guidance for Industry - Providing Regulatory Submissions in Electronic Format - Standardized Study Data (Draft)
- Study Data Specifications (soon replaced by Study Data Technical Conformance Guide)
- CDER Common Data Standards Issues

CDISC Standards

[Legacy Conversion](#)

[Challenges](#)

[Case 1. Gap Analysis](#)

[Case 2: An Onco Study](#)

[Conclusions](#)

Legacy Conversion

Any conversion that involves source data being captured in a given format (neither CDASH nor SDTM have been used) which requires a significant transformation effort¹



Before Transformation



After Transformation

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

¹ SDTM Domain Development on the Critical Path – Strategies for Success» DC Izard, PharmaSUG 2012

Legacy Conversion

Steps to follow

- Gap Analysis
- Understanding Source Datasets
- SDTM Modelling
 - Annotated CRF
 - Additional Migration Specifications
- Migration
- Finalize, Validate and Document
 - OpenCDISC Interpretation
 - define.xml and Reviewer Guide

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

«Pruning» Datasets and Grouping/Linking to SDTM Domains



Conclusions

Legacy Conversion Challenges

«Pruning» Datasets and Grouping/Linking to SDTM Domains

SDTM	Raw Datasets			
AE	AE	AEAER		
CM	CM	CMP	PRAML	BPP
DM	DEMO	SCEL	DS	
DS	DEMO	DS	RANDOM EOT	
EX	INFA	INFB		
IE	IE (only violated criteria)			
MH	MH	CH	PSC	DIAG
VS	VS	PMBSA		
YG	NCNC	NPTN	SCT	
ZG	CG	MARA	LVEF	NLMU
ZR	PRAM	CD	RA	

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Legacy Conversion Challenges

Many Datasets into one: the example of Disposition (DS)

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

	DSSEQ	DSTERM	DSDECOD	DSCAT	DSSCAT	EPOCH	DSSTDTC
1	1	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE			2011-08-09
2	2	RANDOMIZED	RANDOMIZED	PROTOCOL MILESTONE			2011-08-15
3	3	TREATMENT FAILURE/RELAPSE	LACK OF EFFICACY	DISPOSITION EVENT	END OF TREATMENT	TREATMENT	2011-08-19
4	4	DEATH, REASON - UNKNOWN CAUSE OF ...	DEATH	DISPOSITION EVENT	END OF STUDY	FOLLOW-UP	2013-01-07

DM **Demography** **DSCAT = PROTOCOL MILESTONE**

Consent Date:

Date of Birth:

	PT	CPEVENT	DSSTDATF	BRTHDATF
1	101001	SCREENING	20110809	19760630
2	101002	SCREENING	20110907	19481121
3	101003	SCREENING	20111101	19450510

DS **Randomization** **DSCAT = PROTOCOL MILESTONE** **DSTERM** **DSDECOD = RANDOMIZED**

Date of Randomization:

	PT	CPEVENT	RDDATF
1	101001	SCREENING	20110815
2	101002	SCREENING	20110912
3	101003	SCREENING	20111107

DS **End of Treatment** **DSCAT = DISPOSITION EVENT** **DSSCAT = END OF TREATMENT**

Primary Reason for subject discontinuing treatment (Check ONE only):

☐ **DSDECOD = DEATH** Death

☐ **DSDECOD = COMPLETED** Maximum allowed treatment is completed **DSTERM**

☐ **DSDECOD = LACK** Failure to achieve a CR, CRp, or CRI after 1 or 2 attempts (induction 1, induction 2) **DSTERM**

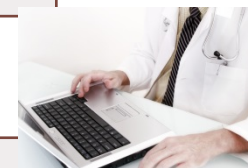
	PT	CPEVENT	DSSTDATF	DSDECODF
11	101015	END OF STUDY	20130114	DEATH
12	101016	END OF STUDY	20130501	DEATH
13	102001	END OF STUDY	20140411	LOST TO FOLLOW UP

Legacy Conversion Challenges

Migrating/Harmonizing Clinical Terminology

Mapping to CDISC (NCI) Controlled Terminology - e.g. Concomitant Medications Route

Source	CDISC/NCI CT
INFUSION	PARENTERAL Administration by injection, infusion, or implantation
MOUTH/THROAT	ORAL Administration to or by way of the mouth
TOPICAL-EYE	OPHTHALMIC
TOPICAL PO	TOPICAL



CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

184 unique terms translated to **25** unique terms in CDISC CT

Mapping to CDISC (NCI) Controlled Terminology e.g. Laboratory Unit

CDISC/NCI CT	Source
10 ⁹ /L	X 10 ⁹ /L X1000/MM3 X10 ³ /MM3
10 ⁶ /L	X10 ⁶ /UL X10 ³ /ML
umol/L	MICROMOL/L

141 unique terms translated to **16** unique terms in CDISC CT

Case 1: Gap Analysis

- A small biotech with limited operational staff
- Several CROs historically involved in the management of their trials
- Missing Trials Documents Inventory
- CDISC SDTM partially implemented
- Lack of understanding

!!! Out of Control !!!



- Cytel asked to oversee CROs and to conduct ISS/ISE (and to fill the gaps)

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Case 1: Gap Analysis

- Initiated prior to commencing migration activities
- How to perform gap analysis
 - Itemization and evaluation of files to support migration activities
 - Document inventory
 - Study documents
 - CDISC Standards
 - Company Standards /
Company Implementation Guidance
 - Validate sample CRF fields versus source data
 - External data requirements e.g. central labs or local labs
 - Comparison of protocol amendments/versions against CRF versions
 - Clarifies the scope and challenges of migration activities
 - Identifies differences in data collection formats



CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Case 1: Gap Analysis

Documents/Data Inventory

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Protocol	Phase	Ongoing/ Closed?	Study Subjects	Country	Raw Data	SDTM	Analysis Data	protocol	SAP	aCRF	Blank CRF	CSR
TRIAL 01	Ia	Closed	CF with Pa	Netherlands	present	To migrate	Non-ADaM	v2.0.1 / 2006	v1.1 / 2007	v1.7.4 / 2007	NA	v1.1 / 201301
TRIAL 02	Ib/Ila	Closed	CF with Pa	Hungary, Serbia	Waiting- will get from another vendor	To migrate	Non-ADaM	v1.1 / 2007	v2.0 / 2010	v1.1 / 2007	NA	v1.1 / 201404
TRIAL 03	Ib/Ila	Closed	CF with Pa	Belgium	present	To migrate	Non-ADaM	v1.4 / 2008	v1.2 / 2010	v2.2 / 2008	NA	v1.1 / 201304
TRIAL 04	III	Ongoing	CF with Pa	North America, Europe, Australia, New Zealand	present	v3.1.3	Available	v1.4 / 2009	v.2.1 / 2011	v1.5 / 2010	Available	Not Applicable

Case 1: Gap Analysis

Documents/Data Inventory

- Assessment of existing SDTM migration
- First raw datasets 'pruning'
 - aCRF vs Raw datasets
 - Quantify SDTM gaps
- Status of medical coding for CMs and AEs
- Lab Conversions
- External Data

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Case 2: A Complex Oncology Study

- **Several Data Challenges**
 - An oncology study with hundreds of patients
 - A lot of free-text to handle including date
- **Integration of several external data**
 - Normal Ranges of Local Labs and conversion to SI unit
 - Independent Review of key Efficacy data
- **Periodic regular refresh**
- **Compliant with ISS/ISE conducted by another CRO**
 - Coding info update
 - Different mapping decisions
 - Integration of several external data

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Case 2: A Complex Oncology Study

Apply CDISC to Oncology Studies – Common Questions/Debates

Reporting Cycles and Visits in Oncology Studies



Associate Director, Statistical Programming at [Redacted]
Clinical

In most of the oncology studies I have worked with, there is both a cycle and a visit identifier associated with each subject contact. In the past, following the guidance in the protocol, I have combined these in VISIT in SDTM. (example of Cycle 1 visit 1, Cycle1 visit 2, Cycle 2 visit 1, etc.) This method has worked and supported the analytics for those studies.

I am trying to map Dose Limiting Toxicity(DLT) data into SDTM. Can anyone help me in which domain it will be mapped to or will it go into SUPPAE?



SAS Clinical Programmer at [Redacted] Healthcare

TNM Classification

-> SDTM IG V3.1.3 Final

Posted 2013-05-20 8:00 AM (#8180)
Subject: TNM Classification

What is the preferred way of storing the TNM Classification in an oncology study?

Date of progression After Treatment (Oncology



Scientific Programmer at [Redacted]

I am working on an oncology study where sponsor has a form for collecting Oncology Drug Biologics which is mapped to CM domain. In this form there is an item to capture Date of Progression After Treatment to which site can enter Req/UNK-Req/UNK-Req/UNK for Day, Month and Year. I have following questions:

1) Why this Date of Progression after treatment is captured and what is its significance. Does it help in differentiating between true progression vs pseudoprogression

I have prior cancer therapy I'll map to SDTM CM domain



Data Manager [Redacted] of Medicine
Top Contributor

but my data have Best Response for that prior therapy. Since CM domain does not have an ORRES variable, any suggestions on what CM variable could be used for the best response to prior therapy?



CDISC - SDTM Experts



CDISC Analyst (SDTM)



CDISC Advocates

Search | Statistics

[CDISC Public Discussion Forums](#)

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Case 2: A Complex Oncology Study

Where to store anti-cancer therapies? Is it a CM?

CMTRT	CMDECOD	CMCAT	CMSCAT	CMDOSTXT
PACKED RED BLOOD CELLS	PACKED RED BLOO.	BLOOD AND PLATELET TRANSFUSION		2
PLATELETS (RANDOM DONOR)	PLATELETS (RANDC	BLOOD AND PLATELET TRANSFUSION		4
PLATELETS (RANDOM DONOR)	PLATELETS (RANDC	BLOOD AND PLATELET TRANSFUSION		4
HYSTERECTOMY		PREVIOUS NON-AML CANCER TREATMENTS	SURGERY	
REMOVAL OF SKIN CANCER		PREVIOUS NON-AML CANCER TREATMENTS	SURGERY	
RIGHT HEMICOLECTOMY		PREVIOUS NON-AML CANCER TREATMENTS	SURGERY	
ACYCLOVIR	ACYCLOVIR/0058730	PRIOR/CONCOMITANT MEDICATION		
		PRIOR AML THERAPY	INDUCTION	200
		PRIOR AML THERAPY	INDUCTION	60
		PRIOR AML THERAPY	INDUCTION	50

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

- Decision to store in CM with appropriate CMCAT/CMSCAT prior anti-cancer therapies and YG sponsor domain to store post study treatment anti-cancer therapies



The new PR domain available in IG 3.2 may be used to store Surgery or other procedures

Case 2: A Complex Oncology Study

Tumor Progression of Previous Therapies

Subsequent Non- Protocol AML Treatments ☐ None Given						
Type of Therapy see legend below	Therapy or Treatment	Date Started	Number of Cycles	Best Response see legend below	Date of Best Response	Date of Relapse
		Date Ended dd mmm yyyy				
Treatment 1						
☐			☐	☐		
Other specify:						☐ NA
		☐ Ongoing				

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

- TU/TR/RS are standard domains in SDTM (version $\geq 3.1.3$)
- The three domains match requirements for solid tumors (RECIST criteria) with the concept of tumor lesion
- Are the three domains meant to be used only for efficacy?
E.g. on-treatment assessment only
- Can it be applied to other non-solid tumors?
- **Decision to create a sponsor domain (ZR) that ‘simulate’ the standard SDTM.RS domain**

Case 2: A Complex Oncology Study

Prior Therapies Tumor Progression: two different approaches

Sponsor A

CM	DOMAIN	CMSEQ	CMSPID	CMTRT	CMCAT	CMSCAT
	CM	7004	1	CAZEPARINE	PRIOR AML THERAPY	INDUCTION
	CM	8005	1	DAUNORUBICIN	PRIOR AML THERAPY	INDUCTION
	CM	9006	1	MIDOSTAURIN VS PLACEBO	PRIOR AML THERAPY	INDUCTION

ZR	DOMAIN	ZRSEQ	ZRSPID	ZRTESTCD	ZRTEST	ZRCAT	ZRRORRES
	ZR	1001	1	BESTRESP	Best Overall Response	PRIOR THERAPY RESPONSE ASSESSMENT	TF - TREATMENT FAILURE
	ZR	1002	1	BESTRESP	Best Overall Response	PRIOR THERAPY RESPONSE ASSESSMENT	TF - TREATMENT FAILURE
	ZR	1003	1	BESTRESP	Best Overall Response	PRIOR THERAPY RESPONSE ASSESSMENT	TF - TREATMENT FAILURE

	RDOMAIN	IDVAR	IDVARVAL	RELID
RELREC	CM	CMSEQ	7004	1
	ZR	ZRSEQ	1001	1
	CM	CMSEQ	8005	2
	ZR	ZRSEQ	1002	2
	CM	CMSEQ	9006	3
	ZR	ZRSEQ	1003	3

The link requires also USUBJID (here hidden)

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Sponsor B

CM	DOMAIN	CMSEQ	CMSPID	CMTRT	CMCAT	CMSCAT
	CM	8	1	KAPECITABINE	PRIOR ANTI-CANCER TREATMENT INCL. TREA	DRUG THERAPY
	CM	9	1	OXALIPLATINA	PRIOR ANTI-CANCER TREATMENT INCL. TREA	DRUG THERAPY
	CM	10	1	RADIO THERAPY	PRIOR ANTI-CANCER TREATMENT INCL. TREA	RADIO THERAPY

SUPPCM	RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
	CM	CMSEQ	8	BRESP	Best Response	PROGRESSIVE DISEASE

Case 2: A Complex Oncology Study

The Study Reviewer Guide: The ideal place where to clarify potential source of misinterpretation



■ Randomization Stratification Factors in SUPPDM

Key analysis data points include:

- Randomization strata are in SUPPDM
 - Age: QNAM= **AGEGROUP**
 - Disease Status: QNAM= **DISSTAT**
 - Geographical Region: QNAM= **LOCATION**

■ Where to retrieve key efficacy information

- Primary Efficacy Endpoint: death subjects are in DM where DTHFL=Y, with date of death **DM.DMDTC** and cause of death in SUPPZS where QNAM=DEATH. For subject alive, last follow-date should be retrieved from ZS; for statistical analysis purpose last date patient was seen alive should be calculated using any date available in the SDTM datasets, except for the following variables:
 - ZS.ZSSTRESC=DEATH, in this case the date (ZSDTC) may be after date of death as this date represents the date when the investigator was informed of the death. For the same reason is not recommended to use date in SV dataset
 - DM.DMDTC

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions

Conclusions



- Migrating to SDTM is a complicated task and it is not just a matter of renaming variables
- ADaM is the right place for «derivation», but actually migrating to SDTM implies the «shake» of original source data
- Non-SDTM users or beginners may find difficult to find what they are looking for
- The migration requires knowledge of several different area
- **The SDTM migration specialist should be a must for all organizations**

CDISC Standards

Legacy Conversion

Challenges

Case 1. Gap Analysis

Case 2: An Onco Study

Conclusions



Thank you for your time!

Angelo Tinazzi – Associate Director – Statistical Programming
angelo.tinazzi@cytel.com

Case 2: A Complex Oncology Study

Mapping Details – New Domains

- Check that none of the existing published domains will fit the need
 - Use existing models as a prototype or choose the appropriate general observation class
 - How to determine where data belong in the SDTM
-  IG v3.1.3 section 8.6
- Assign a 2–letter domain code. Possible convention:
 - Z- for findings
 - Y- for interventions
 - X- for events

Case 2: A Complex Oncology Study

Mapping Details – New Domains

Other non-protocol cancer therapies (YG)

IDs
2.2.4

Variables which identify the study, the subject, the domain and the seq nr. of record

Timing
2.2.5

Variables which describe the timing of an observation (e.g. date collection)

STUDYID	USUBJID	DOMAIN	YGSEQ	YGSPID	YGSTDTC	YGENDTC
XXXX	YYYYYY	YG	1	1	2011-01-12	2011-02-20
XXXX	YYYYYY	YG	2	1	2011-01-12	2011-02-25

Interventions
Qualifiers
2.2.1

Select and include the relevant Qualifier variables from the identified general observation

YGTRT	YGCAT	YGSCAT
FTC PLUS SORAFENIB	SUBSEQUENT NON-PROTOCOL AML TREATMENTS	CHEMOTHERAPY
CLAG-M THERAPY	SUBSEQUENT NON-PROTOCOL AML TREATMENTS	CHEMOTHERAPY