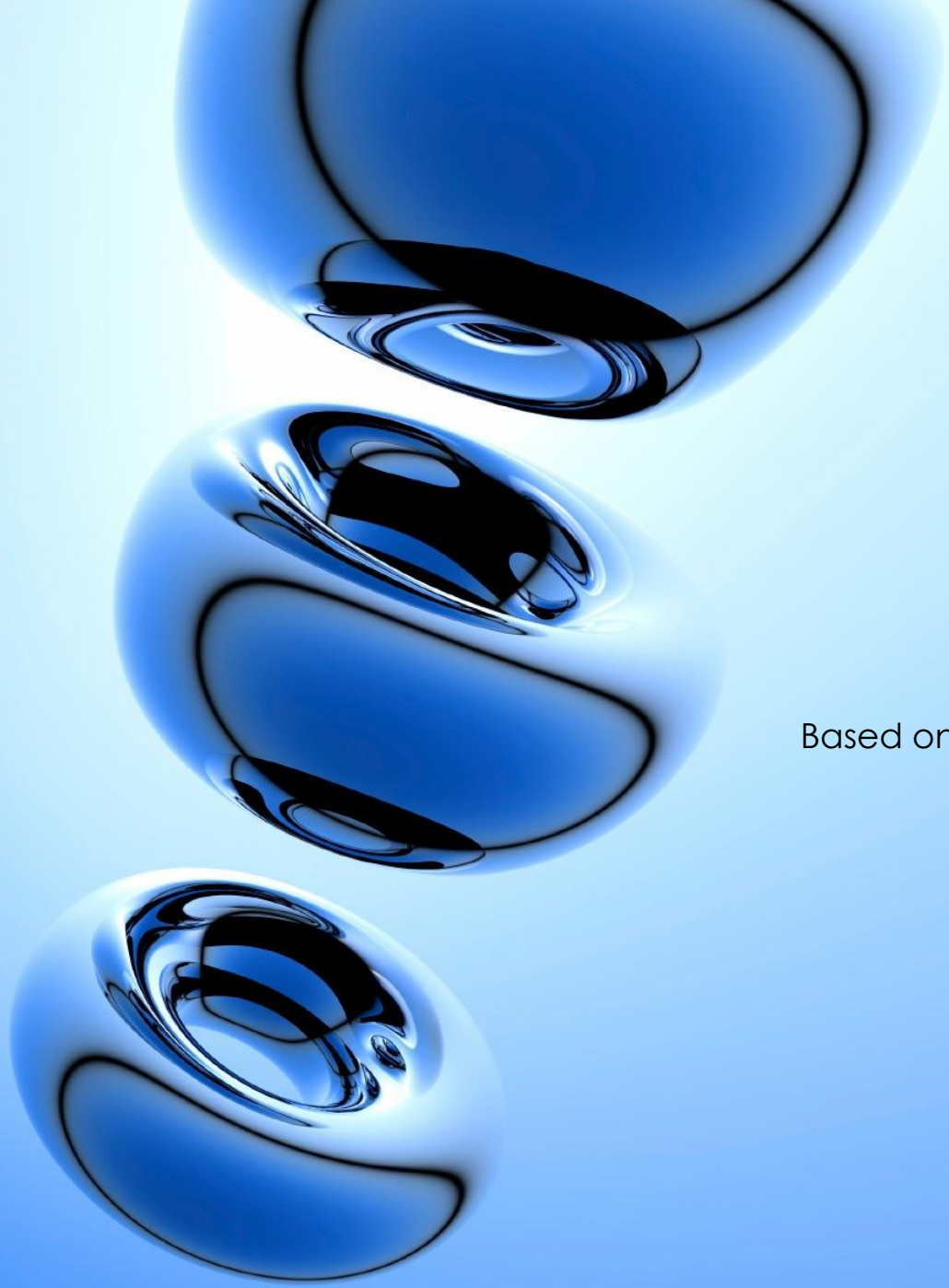


Cytel

The CDISC Stupidario

(The CDISC "Nonsense")

Based on a poster presented at PhUSE-EU 2018 PP26

Angelo Tinazzi, Cytel Inc.

GUF CDISC - Genève / 04 December 2018

Content

- Introduction…What is a “Stupidario”
- One “Stupidario” for each CDISC Submission deliverable
- Conclusions

Introduction

What is a “Stupidario”

Cytel

Introduction
What is a "Stupidario"?

— Può tornarsene tranquillamente a casa, signora: quel virus era nel computer delle diagnosi.

<< Madame you can safely go home: that virus was on computer storing all diagnoses and not on you>>

Introduction
Examples of CDISC "Stupidario"?

My Stats friend: ***Angelo can you please let me know what is the CDISC standard for representing summary of demographics in output table?***

CDISC is only for data standards. There are no industry standard for output templates (yet). The PhUSE initiative has released some white papers with some recommendations

Introduction
Examples of CDISC "Stupidario"?

Questioning a CRO about using an “outdated” version of the Controlled Terminology: “***This is XXX Inc. standard, the development of SDTM IG 3.1.3 has been done in 2013”***

There are actually no requirements to have Ig and CDISC CT from the same period and therefore the CRO, despite working with SDTM Ig 3.1.3, could have implemented a more recent version of the CDISC CT e.g. one of the 2018 CDISC CT

A “Stupidario” in every CDISC Submission deliverable

acrf.pdf

SDTM

ADaM

define.xml

Improper or insufficient documentation

A "Stupidario" in every CDISC Submission deliverable aCRF

From the CDISC Metadata Submission Guideline (MSG) for SDTMIG

*Annotated CRFs included in the eCTD should be **bookmarked 2 ways** (dual bookmarking): bookmarks by **time-points**, often analogous to planned visits in the study, and **bookmarks by CRF topics** or forms. SDTM domains do not necessarily have a 1-to-1 relationship with CRF topics or forms, nor is the reverse true.*

For example, in the annotated CRF, both DM and SC are collected on the Demography panel, while SC data are collected from the Enrolment Form and the Demography pages

A “Stupidario” in every CDISC Submission deliverable aCRF

- Domain
  - AE
  - ACEND
  - ADDINFO
  - ADPROC
  - CM
  - CMAPMUNE
  - CRAPRMPT

- Domain
  - Adverse Events
  - ACEND Questionnaire
  - Additional Info
  - Ancillary Procedures
  - Concomitant Medications
  - Treatment Adninistration
  - CRA Confirmation

A “Stupidario” in every CDISC Submission deliverable aCRF

- By Domains
  - CE - Clinical Events
    - Page 39
    - Page 40
    - Page 41
    - Page 42
    - Page 43
    - Page 44

- By Domains
  - CE - Clinical Events
    - Visit 3 Months
    - Visit 6 Months
    - Visit 9 Months
    - Visit 1 Years
    - Visit 2 Years
    - Visit 5 Years

A “Stupidario” in every CDISC Submission deliverable aCRF

Domain = VS

Domain = DM

Screening Visit 1 (Day -90) [V1]

Demographics [DM]

VS = Vital signs

DM = Demographics

Screening Visit 1 (Day -90) [V1]

Demographics [DM]

Check also CDISC German UN Proposal
PhUSE EU 2018 Poster “Guideline for submission ready aCRF” – PP02

A "Stupidario" in every CDISC Submission deliverable
SDTM

Conformance Issue (P21 Message)	Justification Provided in the cSDRG
Invalid value for --TEST variable	*Many instances of --TEST >40 characters. --TEST values are directly assigned from the labels taken from the Case Report Form to have clear understanding of the test code and therefore text was not changed*

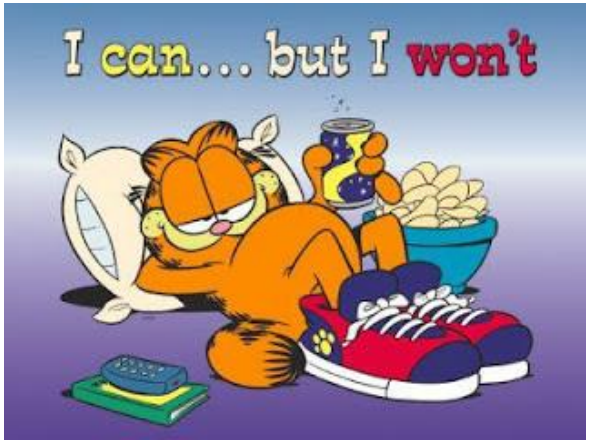
This is a wrong implementation!

--TEST should have been abbreviated and full text specified in the cSDRG

A "Stupidario" in every CDISC Submission deliverable
SDTM

Conformance Issue (P21 Message)	Justification Provided in the cSDRG
Define.xml/CDISC dataset Description mismatch	*LB, IE, QS, FA label is incorrect in XPT*
Variable is in wrong order within domain	*QSCAT and QSSCAT are incorrectly placed*

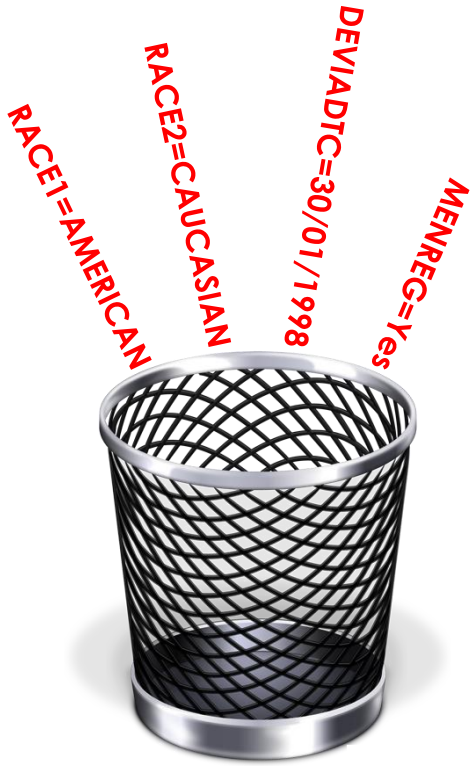
Why don't you fix it !?!?



A "Stupidario" in every CDISC Submission deliverable
SDTM

SUPP is not a simple "***Trash Can***"

Information stored in SUPP "***deserve***" the same "treatment" as those stored in the parent domain



Apply CDISC CT e.g.

- ***YN*** instead of Yes/No
- ***RACE*** for multiple Races

Apply other standards e.g. ***ISO for date***

How many ADaM datasets do I need to create?

No wrong/correct answer. It is analysis driven

What ADaM datasets are required for a submission (FDA)?

ADSL and ADaM datasets containing primary/secondary analysis (FDA SDTCG) **….. at least** (Angelo Tinazzi)

Should there be an ADaM dataset for every SDTM domain?

No

What is ADIE for – ADaM for Inclusion / Exclusion Criteria?

ADIE what? *An ADaM dataset for violated Inclusion and Exclusion Criteria.* **This is really not needed!**

A "Stupidario" in every CDISC Submission deliverable ADaM

*I'm getting the following error message from P21 "**Inconsistent value for AVALC**". This is because in my study for one parameter I have values such as **'<2' (in AVALC) that have been imputed to a numeric value of '2' (in AVAL)**. This was done according to the Statistical Analysis Plan. **However I have also numeric results reported exactly as '2'** and I get this error even if I set my AVALC equal to Null when my result does not contain any sign*

AVISIT	PARAM	AVAL	AVALC
Baseline	Glucose (mg/dl)	2	<2
Visit 1	Glucose (mg/dl)	2	2

From ADaM Ig 1.1

Within a given parameter, if there exists a row on which both AVALC and AVAL are populated, then there must be a one-to-one mapping between AVALC and AVAL on all rows on which both variables are populated

AVISIT	PARAM	AVAL	AVALC	LBSTRESC
Baseline	Glucose (mg/dl)	2	<2	<2
Visit 1	Glucose (mg/dl)	2	2	2

A "Stupidario" in every CDISC Submission deliverable define.xml

ARM	Description of Planned Arm		text	14	["Enrolled", "Screen Failure"] <Arm>	Assigned	Taken from IVRS dataset RANDOM.TRT
ACTARMCD	Actual Arm Code		text	8	["ENROLLED" = "Enrolled", "SCRNFAIL" = "Screen Failure"] <Arm (Code)>	Derived	Same as ARMCD
ACTARM	Description of Actual Arm		text	14	["Enrolled", "Screen Failure"] <Arm>	Assigned	Assigned from TA.ARM based on ACTARMCD.
COUNTRY	Country		text	3	ISO 3166	Assigned	Derived from SITEINFO.CTRY

CMTRT	Reported Name of Drug, Med, or Therapy	5	text	150		CRF Pages 8 9 13 18 19 20 21 22 23 24 45 46 47 48 63 65	Taken from raw datasets CMTRT and CMBASE

RANDOM.TRT what?
SITEINFO.CTRY what?
CMTRT and CMBASE what?

These are mapping specifications and they should be not included in the define.xml

A "Stupidario" in every CDISC Submission deliverable define.xml

Something Missing Here

ATPTREF	Analysis Timepoint Reference	text	40		Derived:
ADY	Analysis Relative Day	integer	8		Derived:
ATPT	Analysis Timepoint	text	40		Derived:

TRTA	Actual Treatment	text	14	TRT	Derived: Set to "SCREEN FAILURE" for all screen failures (check if ADSL.ACTARM = "SCREEN FAILURE") and set to "NOT TREATED" for all patients not treated (check if ADSL.ACTARM ="NOT TREATED"). For all other patients the following rules apply: If study is not in ("[redacted]","[redacted]-E") then set to ADTREAT.TRTA. Else If study ="[redacted]" then do: If SR.EPOCH in("SCREENING","BLINDED TREATMENT") set to ADTREAT.TRTA where ADTREAT.APHASE="BLINDED TREATMENT", else if SR.EPOCH in("FOLLOW-UP","OPEN LABEL TREATMENT","OPEN LABEL TREATMENT EXT") set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL TREATMENT", else if SR.EPOCH is missing and SR.SRDTC is not missing then do: [if SR.SRDTC <= ADTREAT.AENDT where ADTREAT.APHASE="BLINDED TREATMENT" then set to ADTREAT.TRTA where ADTREAT.APHASE="BLINDED TREATMENT", else set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL TREATMENT"]. Else if study = "[redacted]" then do: If SR.EPOCH in("SCREENING","OPEN LABEL FIRST TREATMENT") set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL FIRST TREATMENT", else if SR.EPOCH in("FOLLOW-UP","OPEN LABEL SECOND TREATMENT") set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL SECOND TREATMENT", else if SR.EPOCH is missing and SR.SRDTC is not missing then do: [if SR.SRDTC < ADTREAT.TRTSDT where ADTREAT.APHASE="OPEN LABEL SECOND TREATMENT" then set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL FIRST TREATMENT" else set to ADTREAT.TRTA where ADTREAT.APHASE="OPEN LABEL SECOND TREATMENT"]

AVISIT	Analysis Visit	text	25	[redacted]_AVISIT	Derived: See SAP 4.8.5 for the time frames Calculate the time frames (AVISIT) checking whether the related SR.SRDTC is in the frame of the period starting at Date of first exposure to treatment (ADSL.TRTSDT). [redacted]AVISIT Derivation Algorithm

Check also PhUSE EU 2018 "Do's and Don'ts of Define.xml" - SA04

A "Stupidario" in every CDISC Submission deliverable
Improper or insufficient documentation (reviewer guide)

Conformance Issue (P21 Message)	Justification Provided in the cSDRG
NULL value in SEX variable marked as Required	Data Issue: Sex is collected in the raw data

Does it mean the sex was collected but for some subject the information was not available in the original data?

A "Stupidario" in every CDISC Submission deliverable
Improper or insufficient documentation (reviewer guide)

Conformance Issue (P21 Message)	Justification Provided in the cSDRG
NULL value in AEDECOD variable marked as Required	Terms were not coded in the database

Incomplete coding might be the object of a rejection to PMDA and major concern of the FDA. You need a strong rationale and therefore the explanation should have provided the reason why the term was not coded

Conclusions

Quality Matters

cSDRG Conformance section made by the sponsor

SD1118: Neither DSSTDTC, DSDTC nor DSSTDY are populated. The randomization file did not provide date information to populate DSDTC or DSSTDTC and consequently DSSTDY

FDA feedback (mock-submission)

"NOT RANDOMIZED" is not a Disposition Event. Info about a missing event is not an event

AD0058: *DT is not a numeric variable in ADAE dataset

FDA feedback (mock-submission)

- Explanation in Reviewer's Guide is not valid
- It's a violation of ADaM standard

Conclusions

- The efficacy and safety of your drug are of course what matter, but **lack of traceability**, **poor or insufficient documentation** might **trigger questions and concerns from the reviewer**
- You may think these are minor issues because they do not ultimately impact any results. However, **you are risking your credibility with the FDA reviewer**, who may conclude that your package is not of good quality

Angelo Tinazzi – Director – Standards, Systems, CDISC Consulting (CDISC E3C Member)

angelo.tinazzi@cytel.com

Cytel, Shaping the Future of Drug Development

References

- C. Paul, S. Sturm, "**Guideline for submission ready aCRF**", PhUSE-EU 2018, PP02
- D. Roulstone, "**Do's and Don'ts of Define.xml**", PhUSE-EU 2018, SA04
- A. Tinazzi, "**The « CDISC Stupidario » (the CDISC Nonsense)**", PhUSE-EU 2018, PP26
- V. Debbeti, "**How to Prepare High-quality Metadata for Submission**", PhUSE-US 2018, SI12
- PhUSE WG, "**Best Practices for Metadata Documentation (Define-XML vs reviewer's guide)**"
- PhUSE WG, "**Define-XML v2.0 Completion Guidelines & Style Sheet Recommendations**"

Abstract

In the last 5-10 years I have been exposed to several studies requiring the use of the CDISC standards, either as programmer study lead or as CDISC SME (Subject Matter Expert) reviewing both internal (Cytel) or external packages (delivered by Pharma or other CROs), where I also regularly provide answers to questions/doubts.

With this presentation I would like to go through the main CDISC "Nonsense" from my experience. This can range from "nonsense" questions to a complete misunderstanding of the CDISC Ig(s); some of this "nonsense" has also emerged from the CDISC packages I have reviewed including CDISC documentation such as the reviewer guide.

The main focus of the presentation will be the SDTM and ADaM standards.

Backup Slides

Where is my Traceability

- The sponsor team derived the mean of the three ECG measurements (triplicates); up to here nothing wrong – this was correctly derived in the ADaM
- However, I recommended the sponsor keeping also the three original records from which the derived parameter was created
- The answer from the sponsor was a bit "unsympathetic"

  "The original records were not retained in ADEG because they are in SDTM.EG"

Although there is no 'obligations' from the ADaM Ig to keep all records from SDTM when creating ADaM datasets, it is a good attitude to keep these records when these records are the source of records (parameters) derived in ADAM

My ADaM PARAM should be enough

- Why you need AVALU?
- We need it to store the unit of PARAMThe answer from the sponsor was a bit “unsympathetic”

This is not really needed unless you are doing an ECG analysis based on the subject position. Again this is clear from the ADaM Ig section 3.3.4:

> ***“PARAM must include all descriptive and qualifying information relevant to the analysis purpose of the parameter”***

this means, in few words, that in most of the cases concatenating the parameter name and its unit should be enough (this is what usually goes into the statistical output).

cSDRG to be improved

- Issue: “Missing FADY variable, when FADTC variable is present
- Explanation: “Variable not used”

Although –DY variable is permissible and sponsor could omit it, the FDA Study Technical Conformance Guide requires –DY variable to be included when –DTC is included in the data. Simply saying “Variable not used” does not matter!

cSDRG to be improved

- Issue (AE): “Permissible variable with missing value for all Records”
- Explanation: “No data has been collected”

This is about seriousness criteria. It would have been better to clearly specify in the explanation for which variables this issue concern and probably either say that no serious AE with that specific criteria did occur (and in that in case you can also omit the criteria variable) or mention to the reviewer that the study CRF was not collecting such a detail (if that was the case)

cSDRG to be improved

- Issue: “Inconsistent value for Standard Units”
- Explanation: “Data Issue: We have not been able to convert these to standard units”

More details about laboratory parameter and unit concerned should have been mentioned