
CDASH in Data Management

In search of the simplified process promised

PhUSE, Eschborn, 06-May-2009

Stefan Beimel





Agenda

Short introduction to CDASH

Comparison CDASH - SDTM

Implementation at Merz

‘ The Tool ‘

Open questions



What is CDASH



Clinical Data Acquisition Standards Harmonization

Final 1.0 01-Oct-2008 [www.cdisc.org/standards/cdash/]

Which data are collected in clinical trials

Highly recommended fields, e.g. AE Term, ECG status (done/ not done)

Fields not recommended for collection (in the draft only!), e.g. AE ongoing flag,
Total daily dose of Concomitant Medication

How data are collected

Best practices for creation and review of CRF

CRF completion instruction for every variable

No graphical layout given

How data are stored

Domains / variables are similar to SDTM

More convenient for Data Management

⇒ **Serves as the interface between Data Management and Statistics**

Differences CDASH - SDTM



Advantages for data management

No derivations	(Complex) derivations required
Additional variables possible	Additional entries in SUPP domains
Contains all CRF entries	Contains actual findings only

Other differences

Recommendation only	Regulatory requirement
16 basic domains	30 domains, incl. study level info



Why does Merz implement CDASH

To make communication with CROs and other partners easier

- currently CROs should deliver Oracle Clinical like structure which is usually difficult

To deliver the data structure to statistics independently of the source system

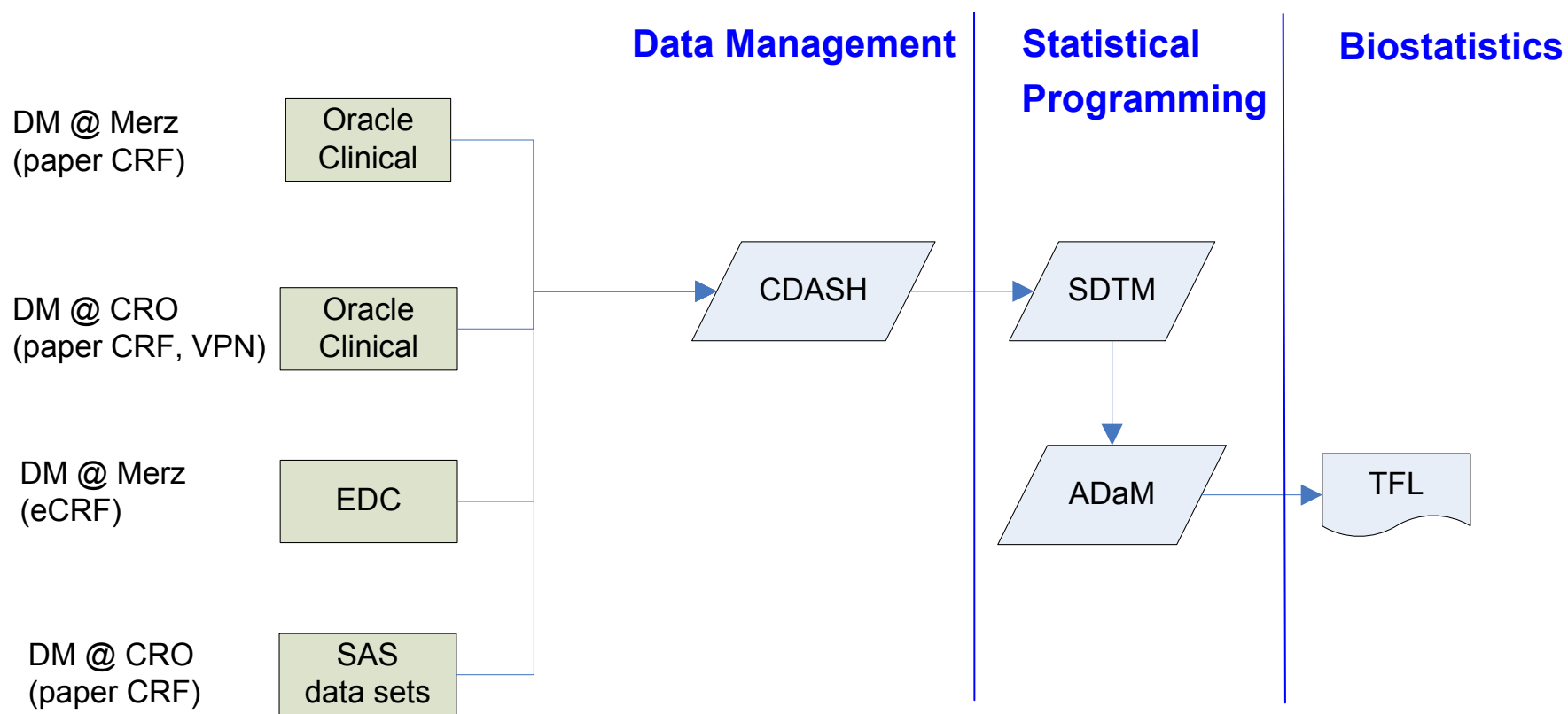
- currently 3 different sources for data at Merz
 - Merz' Oracle Clinical (inhouse studies / VPN studies)
 - CROs Clinical Data Management System (SAS data sets)
 - EDC system

To make subsequent processes easier

- CDASH is similar to SDTM, i.e. conversion to SDTM will be easier than from any other data structure



Planned process with CDASH / SDTM ...





Implementation of CDASH

As CDASH is originally designed to be a CRF, the ideal implementation strategy would be to design the data entry screens in CDASH structure.

However, there are

- **problems with CDASHs data structure in itself**
- **problems with CDMS*: Different output data structure**

*** Clinical Data Management System**



Problems with CDASHs data structure in itself

Problem I.a: Inconsistent key structure within one domain.

Domain AE: **AEYN** (per subject) and **AETERM** (per event)
VS: **VSDATE** (per measurement) and **VSORRES** (per parameter)

Problem I.b: Different code lists per variable

Domain SC*: SCTEST = **Childbearing potential**, SCORRES = **yes | no**
SCTEST = **Eye Color**, SCORRES = **blue | green | red**

Problem I.c: Common timing variable VISDAT

Collected usually in a separate module, expected in each domain by CDASH.

Problem I.d: Domain DS* on different modules

Collection of 'important' events like Informed Consent, Randomization, Discontinuation

* SC: Subject Characteristics, DS: Disposition Events



Problems with CDMS: Different output data structure

Problem II.a: Additional auxiliary/ key variables

Oracle Clinical: Qualifying Value, Repeat SeqNo, DCM name, ...
EDC system: ...

Problem II.b: Ability (or inability) to store similar data in different modules in the same data set.

Blood Pressure and Body Temperature at Screening vs. **Blood Pressure** only at Visit 2
Date of examination at Screening vs. **Date and Time** at Visit 2

Problem II.c: Ability (or inability) to extract formatted and unformatted values

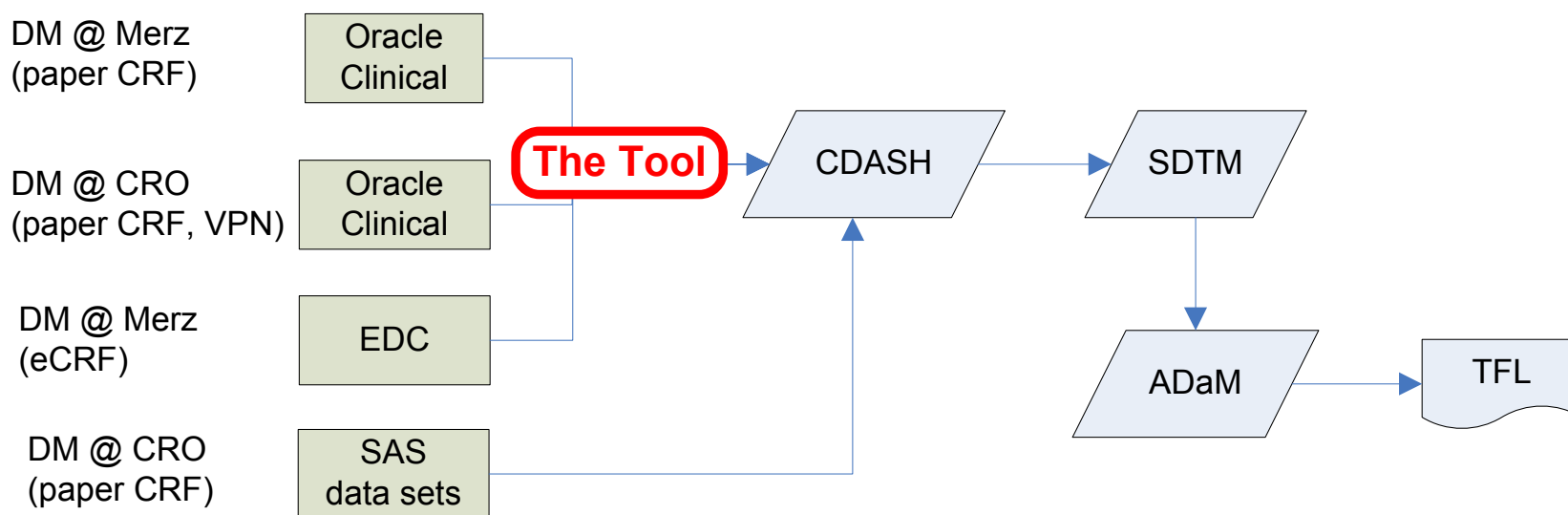
Problem II.d: Different representation of dates

YYYYMMDD vs. DD-MMM-YYYY

Solution

A tool that

- converts the original data into CDASH structure
- solves problems I.a to II.d
- requires no (SAS) programming
- is easy to use



Example conversion for domain IE

Original data

Ini.xls

CDASH like data

	SUBJID	VISITNUM	DCISHTNM	EXCR1	EXCR2
1	S1001	1	M009	2	2
2	S1001	1	M010	.	.
3	S1001	2	P002	.	.
4	S1001	3	M012	.	2
5	S1002	1	M009	2	2
6	S1002	1	M010	.	.
7	S1002	2	P002	.	.

	SUBJID	VISITNUM	INCR1	INCR2
1	S1001	1	1	1
2	S1001	2	.	.
3	S1001	3	.	.
4	S1002	1	1	1
5	S1002	2	.	.



	SUBJID	IETESTCD	IETEST	VISIT	IEORRES
1	S1001	E01	Any clinically	SCREENING	no
2	S1001	E02	Any disease or	SCREENING	no
3	S1001	E02	Any disease or	DAY 1	no
4	S1001	I01	Healthy male a	SCREENING	yes
5	S1001	I02	White origin	SCREENING	yes
6	S1002	E01	Any clinically	SCREENING	no
7	S1002	E02	Any disease or	SCREENING	no
8	S1002	I01	Healthy male a	SCREENING	yes
9	S1002	I02	White origin	SCREENING	yes

Example of initialization file



Study	MRZ XXX												
DB instance	OCPROD												
View type	current												
Oracle Clinical		Rule	CDASH				Additional Key I				Additional Key II		
Data Set	Variable	D = delete C = copy/ rename TR = transp. by Rep# K=add key for Copy	Domain	Variable TR: prefix	Label	Type	Name	Value "constant", _NAME_, _LABEL_	Label	Type	Name	Value "constant", _NAME_, _LABEL_	Label
LABEL			AE		Adverse Events (CDASH)								
LABEL			CM		Concomitant Medication (CDASH)								
LABEL			IE		In- / Exclusion Criteria (CDASH)								
CIV	STUDYID	C											
CIV	SITEID	C											
CIV	SUBJID	C											
EXCL	QUALIFYV	D											
EXCL	REPEATSN	D											
EXCL	VISIT	K											
EXCL	VISITNUM	K											
EXCL	EXCR1	C	IE	IEORRES	I/E Criterion Original Result	\$32	IETESTCD	"E01"	Incl/Excl Criterion Short Name	\$8	IETEST	"Any clinically	Incl/Excl Criterion
EXCL	EXCR2	C	IE	IEORRES	I/E Criterion Original Result	\$32	IETESTCD	"E02"	Incl/Excl Criterion Short Name	\$8	IETEST	"Any disease	Incl/Excl Criterion
INCL	QUALIFYV	D											
INCL	REPEATSN	D											
INCL	VISIT	K											
INCL	VISITNUM	K											
INCL	INCR1	C	IE	IEORRES	I/E Criterion Original Result	\$32	IETESTCD	"I01"	Incl/Excl Criterion Short Name	\$8	IETEST	"Healthy male	Incl/Excl Criterion
INCL	INCR2	C	IE	IEORRES	I/E Criterion Original Result	\$32	IETESTCD	"I02"	Incl/Excl Criterion Short Name	\$8	IETEST	"White origin"	Incl/Excl Criterion



The Tool can ...

Map variables to new variables / rename a variable

EXCL.EXCR1, EXCL.EXCR2, INCL.INCR1, INCL.INCR2 → IE.IEORRES,

Create new key variables

IETEST, IETESTCD

Delete variables

QUALIFYV, REPEATSN

Map the formatted or unformatted value or both

By specifying *_NUM in the original variable name (not in the example)

Copy original data sets not mentioned at all

To avoid the need to specify every single variable in the INI file

With these functions most of the problems can be solved.



The Tool cannot ...

Map conditionally

- If in the original data, the date of lab sampling for **safety analysis** (target domain LB) and **genotyping** (target domain SC) is collected in the same module and variable, they cannot be separated
- Deletion of fulfilled in-/ exclusion criteria

Map one variable to more than one domain/ variable

It is not possible to duplicate VISDAT (problem I.c)

Apply any functions

It is not possible to convert dates into the CDASH recommended format (problem II.d)



Open questions

How to handle data that are (currently) not described in CDASH?

In case a CRO's CDMS is used: How to define an appropriate data structure, how to send this to a CRO and how to QC the incoming data?

What happens, if CDASH recommendations do not fit the current data acquisition in the company?



Thank you for your attention