SI04

Optimization of the traceability when applying an ADaM Parallel Conversion Method

Roxane Debrus

Business & Decision Life Sciences

Agenda

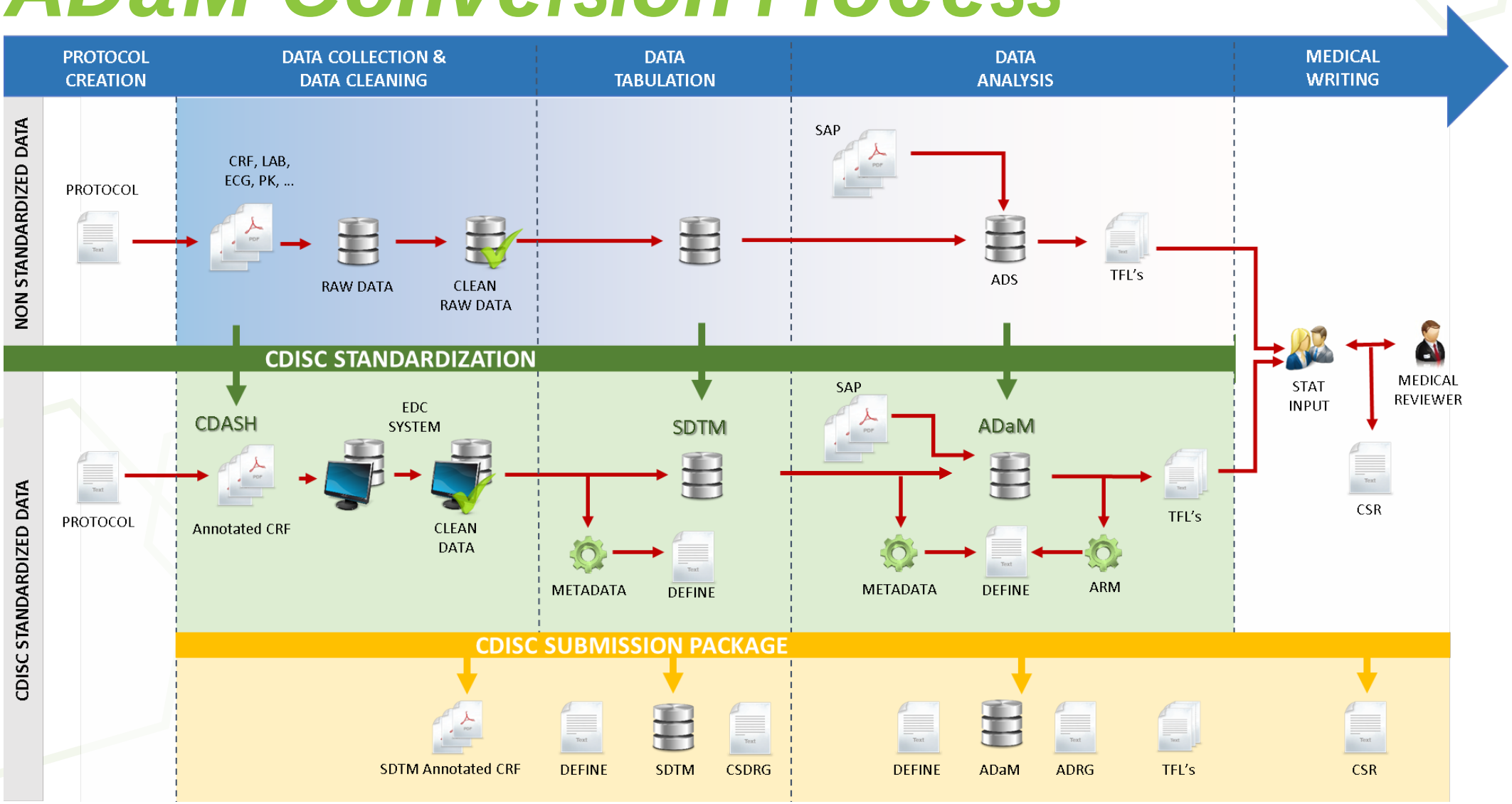
ADaM Conversion Process

%LIB_QC_contents_html

%adam_sdtm_compa

Conclusion

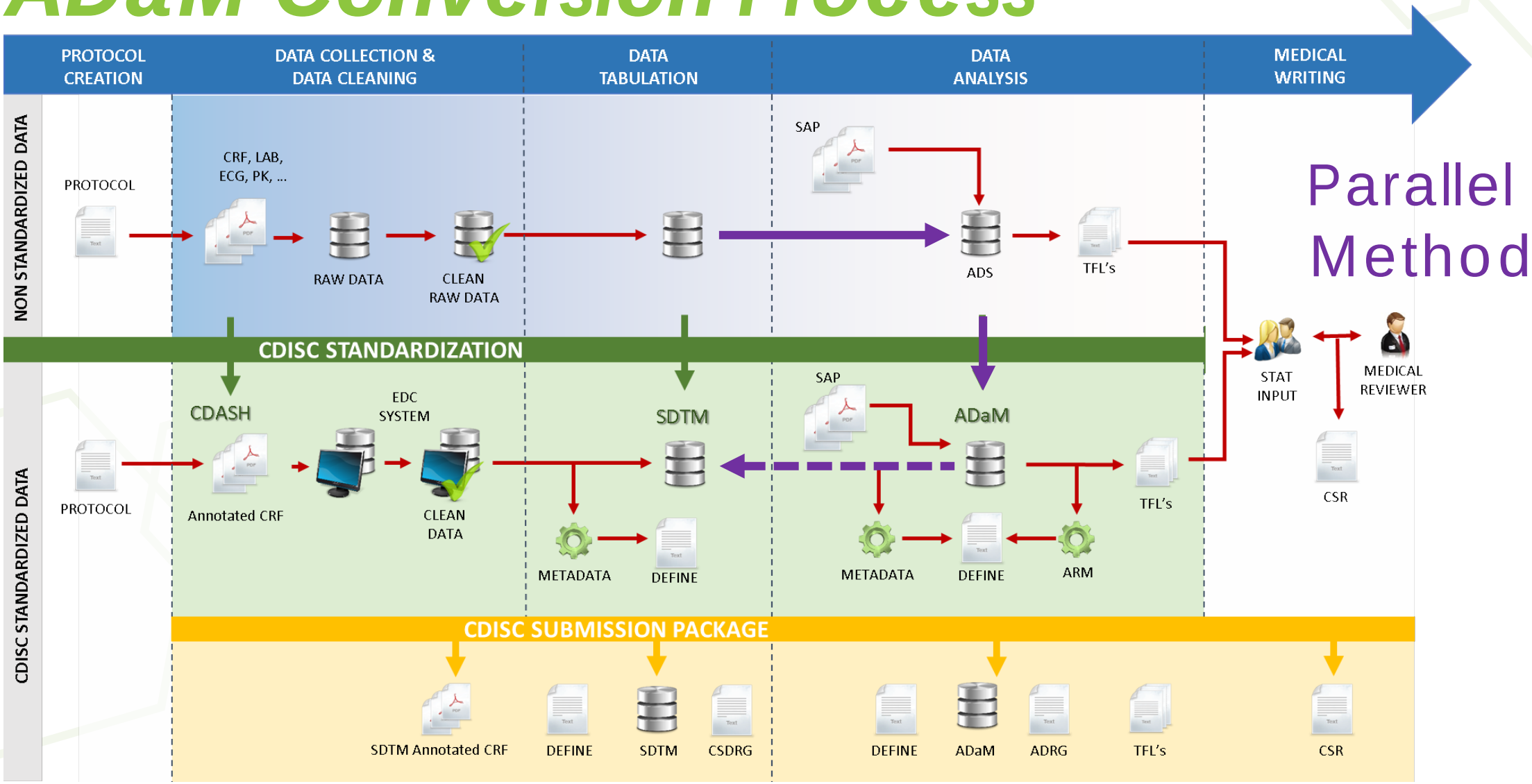
ADaM Conversion Process



Business & Decision Life Sciences

ROI² return on investment through return on information

ADaM Conversion Process



Business & Decision Life Sciences

ROI² return on investment through return on information

ADaM Conversion Process

- ***Why traceability ?***
  - *Understand the analysis datasets structure*
  - *Identify the data collected and algorithms used to derive variables*
  - *Understand how analysis was performed*
  - *Relate counts from CSR's TFL to the underlying data*

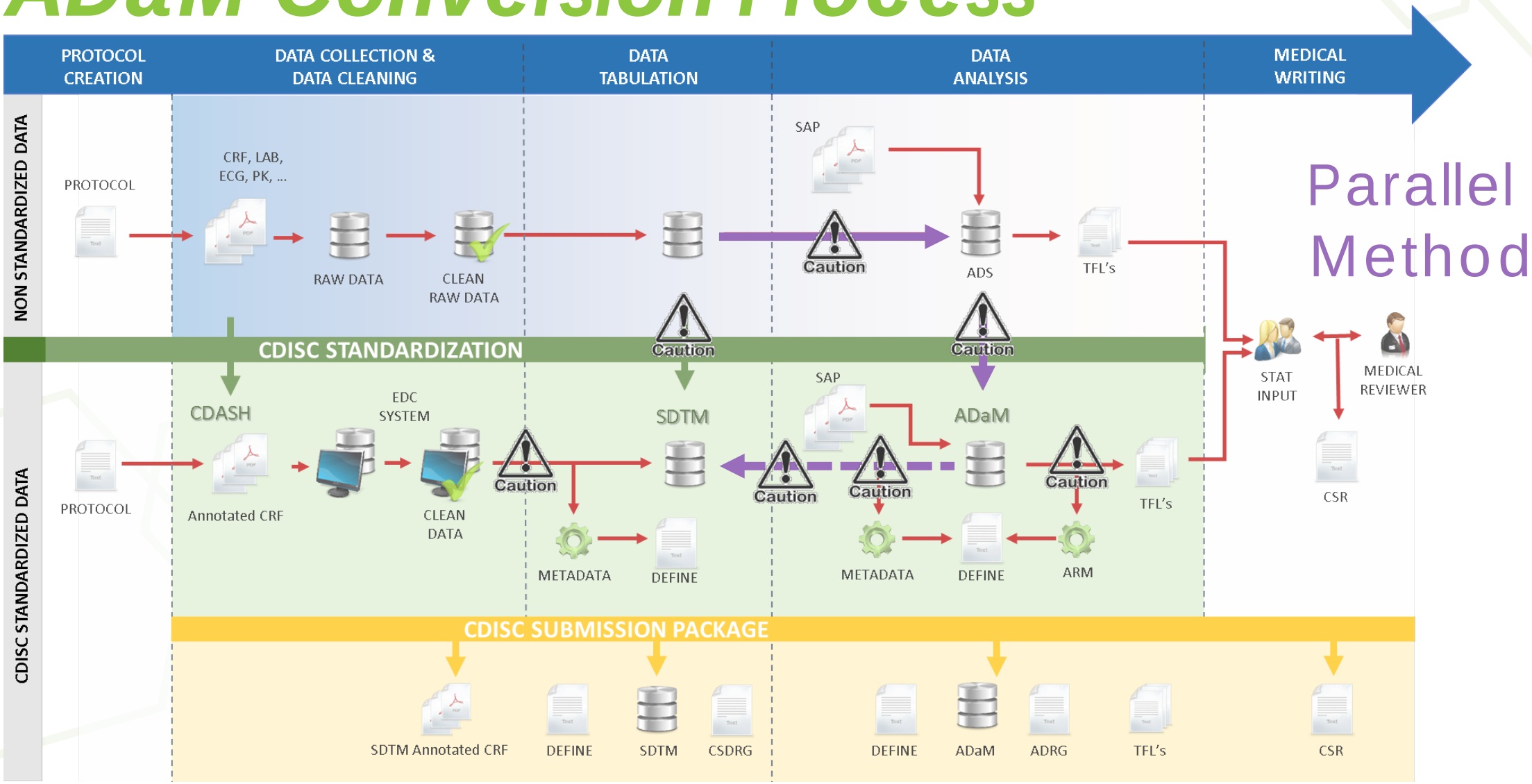
“Based on reviewer experience, establishing traceability is one of the most problematic issues associated with any data conversion”

➔ *8 Specific issues identified*

Reference : FDA, “Technical Specifications Document: Study Data Technical Conformance Guide”, Version 4.1, Issued March 31, 2018

FDA

ADaM Conversion Process



Agenda

ADaM Conversion Process

%LIB_QC_contents_html

%adam_sdtm_compa

Conclusion

%LIB_QC_contents_html

- ***Characteristics***
  - *Reads any type of libraries (CDISC compliant and Legacy)*
  - *Report a description of each dataset and its content*
  - *Produces a set of HTML files*

- ***Outputs***
  - *Global library description*
  - *Detailed description for each dataset*
    - Main characteristics:
      variable name, label, type and length
    - Content description:
      N distinct, N missing, Mean, Min, Max

%LIB_QC_contents_html

M:\CRO\PRESENTATIONS\BDLS PRESENTATIONS\3. Events\201811 PhUSE Frankfurt\Presentations\SI04\CON - Internet Explorer

M:\CRO\PRESENTATIONS\BDLS PRESENTATI | M:\CRO\PRESENTATIONS\B...

AE : Adverse Events
CE : Clinical Events
CM : Concomitant Medications
CO : Comments
CP : Compliance
DC : Pharmacodynamic Concentrations
DM : Demographics
DS : Disposition
EX : Exposure
FAAE : Findings About Adverse Events
FU : Follow-Up
HC : Health Care
IE : Inclusion/Exclusion Criteria Not Met
LB : Laboratory Tests Results
MH : Medical History
ML : Meal Data
PC : Pharmacokinetic Concentrations
PP : Pharmacokinetic Parameters
QS : Questionnaires
RELREC : Related Records
SC : Subject Characteristics
SE : Subject Elements
SU : Substance Use
SUPPAE : Supplemental Qualifiers for AE
SUPPCE : Supplemental Qualifiers for CE
SUPPCM : Supplemental Qualifiers for CM
SUPPDC : Supplemental Qualifiers for DC
SUPPDM : Supplemental Qualifiers for DM
SUPPDS : Supplemental Qualifiers for DS
SUPPHC : Supplemental Qualifiers for HC
SUPPLB : Supplemental Qualifiers for LB
SUPPMH : Supplemental Qualifiers for MH

SDTM.DM - Data Set Contents - Internet Explorer

M:\CRO\PRESENTATIONS\BDLS PRESENTATIONS | SDTM.DM - Data Set Contents

SDTM.DM

Demographics
Created: 19MAR18:17:21:28
obs : 730
Keys : STUDYID, USUBJID

#	Variable	Label	Type	Length	+	Contents
1	STUDYID	Study Identifier	CHAR	9	+	1 distinct values, 0 missing values, all shown
2	DOMAIN	Domain Abbreviation	CHAR	2	+	1 distinct values, 0 missing values, all shown
3	USUBJID	Unique Subject Identifier	CHAR	16	+	730 distinct values, 0 missing values, subset shown
4	SUBJID	Subject Identifier for the Study	CHAR	4	+	730 distinct values, 0 missing values, subset shown
5	RFSTDTC	Subject Reference Start Date/Time	CHAR	10	+	234 distinct values, 0 missing values, subset shown
6	RFENDTC	Subject Reference End Date/Time	CHAR	10	+	305 distinct values, 3 missing values, subset shown
7	RFXSTDTC	Date/Time of First Study Treatment	CHAR	10	+	234 distinct values, 0 missing values, subset shown
8	RFXENDTC	Date/Time of Last Study Treatment	CHAR	10	+	305 distinct values, 3 missing values, subset shown
9	RFICDTC	Date/Time of Informed Consent	CHAR	10	+	228 distinct values, 0 missing values, subset shown
10	RFPENDTC	Date/Time of End of Participation	CHAR	10	+	253 distinct values, 0 missing values, subset shown
11	DTHDTC	Date/Time of Death	CHAR	10	+	2 distinct values, 729 missing values, all shown
12	DTHFL	Subject Death Flag	CHAR	1	+	2 distinct values, 729 missing values, all shown
13	SITEID	Study Site Identifier	CHAR	8	+	124 distinct values, 0 missing values, subset shown
14	INVNAM	Investigator Name	CHAR	17	+	124 distinct values, 0 missing values, subset shown
15	BRTHDTC	Date/Time of Birth	CHAR	4	+	59 distinct values, 0 missing values, all shown
16	AGE	Age	NUM	8	+	59 distinct values, 0 missing values, all shown Mean: 45.049 Minimum: 18 Maximum: 77 Missing values: 0
17	AGEU	Age Units	CHAR	5	+	1 distinct values, 0 missing values, all shown
18	SEX	Sex	CHAR	1	+	2 distinct values, 0 missing values, all shown
19	RACE	Race	CHAR	41	+	5 distinct values, 0 missing values, all shown
20	ETHNIC	Ethnicity	CHAR	22	+	2 distinct values, 0 missing values, all shown
21	ARMCD	Planned Arm Code	CHAR	1	+	3 distinct values, 0 missing values, all shown
22	ARM	Description of Planned Arm	CHAR	14	+	3 distinct values, 0 missing values, all shown
23	ACTARMCD	Actual Arm Code	CHAR	1	+	3 distinct values, 0 missing values, all shown
24	ACTARM	Description of Actual Arm	CHAR	14	+	3 distinct values, 0 missing values, all shown
25	COUNTRY	Country	CHAR	3	+	17 distinct values, 0 missing values, all shown

%LIB_QC_contents_html

Categorical Variables

18	SEX	Sex	CHAR	1	– 2 distinct values, 0 missing values, all shown "F" (389x) "M" (341x)
19	RACE	Race	CHAR	41	– 5 distinct values, 0 missing values, all shown "ASIAN" (23x) "BLACK OR AFRICAN AMERICAN" (17x) "MULTIPLE" (2x) "NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER" (1x) "WHITE" (687x)

Discrete Variables

11	REGION1	Region Cat1	CHAR	40	– 4 distinct values, 0 missing values, all shown "Asia" (17x) "Europe" (399x) "North America" (282x) "Pacific" (32x)
12	REGION1N	Region Cat1 (N)	NUM	8	– 4 distinct values, 0 missing values, all shown Mean: 1.632 Minimum: 1 Maximum: 5 Missing values: 0 1 (399x) 2 (282x) 4 (17x) 5 (32x)

Business & Decision Life Sciences

ROI² return on investment through return on information

%LIB_QC_contents_html

Continuous Variables

38	DIAGT	Time since diagnosis	NUM	8	– 709 distinct values, 0 missing values, subset shown Mean: 22.507 Minimum: 1.1362080767 Maximum: 67.071868583 Missing values: 0

```
1.1362080767   1.1553730322   1.28678987   1.5331964408
1.6783025325   1.8535249829   1.8973305955   2.3572895277
2.3819301848   2.4640657084   2.7707049966   2.8501026694
2.8692676249   3.1567419576   3.2114989733
...
50.611909651   51.23340178   51.707049966   52.232717317
52.706365503   53.11430527   54.015058179   54.828199863
54.904859685   55.263518138   55.619438741   56.539356605
60.613278576   63.646817248   67.071868583
```

Empty variable

25	AESCONG	Congenital Anomaly or Birth Defect	CHAR	1	– All values are missing

Business & Decision Life Sciences

ROI² return on investment through return on information

%LIB_QC_contents_html

Practical Example 1

SDTM.DM - Data Set Contents - Internet Explorer

SDTM.DM

Demographics
Created: 19MAR18:17:21:28
obs : 730
Keys : STUDYID, USUBJID

#	Variable	Label	Type	Length	+	Contents
1	STUDYID	Study Identifier	CHAR	9	+	1 distinct values, 0 missing values, all shown
2	DOMAIN	Domain Abbreviation	CHAR	2	+	1 distinct values, 0 missing values, all shown
3	USUBJID	Unique Subject Identifier	CHAR	16	+	730 distinct values, 0 missing values, subset shown
8	RFXENDTC	Date/Time of Last Study Treatment	CHAR	10	+	305 distinct values, 3 missing values, subset shown
9	RFICDTC	Date/Time of Informed Consent	CHAR	10	+	228 distinct values, 0 missing values, subset shown
10	RFPENDTC	Date/Time of End of Participation				
11	DTHDTC	Date/Time of Death				
12	DTHFL	Subject Death Flag				
13	SITEID	Study Site Identifier				
14	INVNAM	Investigator Name	CHAR	17	+	124 distinct values, 0 missing values, subset shown
15	BRTHDTC	Date/Time of Birth	CHAR	4	+	59 distinct values, 0 missing values, all shown
16	AGE	Age	NUM	8	+	59 distinct values, 0 missing values, all shown Mean: 45.049 Minimum: 18 Maximum: 77 Missing values: 0
17	AGEU	Age Units	CHAR	5	+	1 distinct values, 0 missing values, all shown
18	SEX	Sex	CHAR	1	–	2 distinct values, 0 missing values, all shown "F" (389x) "M" (341x)
19	RACE	Race	CHAR	41	–	5 distinct values, 0 missing values, all shown "ASIAN" (23x) "BLACK OR AFRICAN AMERICAN" (17x) "MULTIPLE" (2x) "NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER" (1x)
24	ACTARM	Description of Actual Arm	CHAR	14	+	3 distinct values, 0 missing values, all shown
25	COUNTRY	Country	CHAR	3	+	17 distinct values, 0 missing values, all shown

ADAM.ADSL - Data Set Contents - Internet Explorer

ADAM.ADSL

Subject Level Analysis Dataset
Created: 23APR18:15:29:45
obs : 730
Keys : STUDYID, USUBJID

#	Variable	Label	Type	Lengtl	+	Contents
1	STUDYID	Study Identifier	CHAR	9	+	1 distinct values, 0 missing values, all shown
2	USUBJID	Unique Subject Identifier	CHAR	16	+	730 distinct values, 0 missing values, subset shown
3	SUBJID	Subject Identifier for the Study	CHAR	4	+	730 distinct values, 0 missing values, subset shown
8	COUNTRDC	Country Decode	CHAR	40	+	17 distinct values, 0 missing values, all shown
9	REGION	Region	CHAR	40	+	3 distinct values, 0 missing values, all shown
13	REGION2	Region Cat2	CHAR	40	+	2 distinct values, 0 missing values, all shown
14	REGION2N	Region Cat2 (N)	NUM	8	+	2 distinct values, 0 missing values, all shown Mean: 1.716 Minimum: 1 Maximum: 2 Missing values: 0
15	REGION3	Region Cat3	CHAR	40	+	3 distinct values, 0 missing values, all shown
16	REGION3N	Region Cat3 (N)	NUM	8	+	3 distinct values, 0 missing values, all shown Mean: 1.456 Minimum: 1 Maximum: 4 Missing values: 0
17	AGE	Age	NUM	8	+	59 distinct values, 0 missing values, all shown Mean: 45.049 Minimum: 18 Maximum: 77 Missing values: 0
18	AGEU	Age Units	CHAR	5	+	1 distinct values, 0 missing values, all shown
19	AGEGR1	Age Group 1	CHAR	20	+	4 distinct values, 0 missing values, all shown
20	AGEGR1N	Age Group 1 (N)	NUM	8	+	4 distinct values, 0 missing values, all shown Mean: 1.484 Minimum: 1 Maximum: 4 Missing values: 0
25	SEX	Sex	CHAR	1	–	2 distinct values, 0 missing values, all shown "F" (389x) "M" (341x)
26	SEXN	Sex (N)	NUM	8	+	2 distinct values, 0 missing values, all shown Mean: 1.533 Minimum: 1 Maximum: 2 Missing values: 0
27	SEXDC	Sex Decode	CHAR	8	+	2 distinct values, 0 missing values, all shown
28	RACE	Race	CHAR	41	–	5 distinct values, 0 missing values, all shown "ASIAN" (23x)

16 AGE Age NUM 8 + 59 distinct values, 0 missing values, all shown
Mean: 45.049 Minimum: 18 Maximum: 77 Missing values: 0

17 AGE Age NUM 8 + 59 distinct values, 0 missing values, all shown
Mean: 45.049 Minimum: 18 Maximum: 77 Missing values: 0

18 SEX Sex CHAR 1 – 2 distinct values, 0 missing values, all shown
"F" (389x)
"M" (341x)

25 SEX Sex CHAR 1 – 2 distinct values, 0 missing values, all shown
"F" (389x)
"M" (341x)

Business & Decision Life Sciences

ROI² return on investment through return on information

%LIB_QC_contents_html

Practical Example 2

According to the ADaM IG, population flags shouldn't have a missing value and should be always equal to "Y" or "N".

94	**SAFFL**	Safety Population Flag	CHAR	1	– 1 distinct values, 23 missing values, all shown **MISSING** (23x) "Y" (700x) "N" (7x)

94	**SAFFL**	Safety Population Flag	CHAR	2	– 3 distinct values, 0 missing values, all shown "Y" (700x) "N" (7x) "NA" (23x)

Business & Decision Life Sciences

ROI² return on investment through return on information

%LIB_QC_contents_html

- ***Advantages***

  Easy to produce, *Easy to integrate,*
  Easy to share, *Easy to open,* *Easy to read*

- ***How to call it***

```
%LIB_QC_contents_html(SDTM);
%LIB_QC_contents_html(W:\PROJECT-X\STUDY-0001\SDTM);

Option(s):  %LIB_QC_contents_html(SDTM, select=DM AE);
            %LIB_QC_contents_html(SDTM, APPLY_FORMATS=Yes);
```

➔ *Options can be combined*

Business & Decision Life Sciences
ROI² return on investment through return on information

%LIB_QC_contents_html

- ***Limitations***
  - *No possible "filter" or "subset by"*
  - *Free text variables*
  - *No annotations possible*

- ***Improvements***
  - *Develop new options for*
    - Selection
    - Display
    - Output format
  - *Compare collected values with expected values*

Agenda

ADaM Conversion Process

%LIB_QC_contents_html

%adam_sdtm_compa

Conclusion

%adam_sdtm_compa

- ***Characteristics***
  - *Compare SDTM vs. ADaM data and metadata*
  - *Based on 8 major CDISC principles*

- ***Outputs***
  - *Reports a description of all findings*
  - *Produces one XML workbook*
    - Summary
    - SDTM Overview
    - ADaM Overview
    - Exception Metadata
    - Exception Data
    - Log Overview

%adam_sdtm_compa

- *8 major CDISC principles*
  - *Check #101:* *A variable present in ADaM is indicated as a copy of a variable present in SDTM but the variables do not have identical metadata attributes*
  - *Check #102:* *A variable present in ADaM is indicated as a copy of a SDTM variable but the SDTM variable does not exist in SDTM variable metadata*
  - *Check #103:* *A variable present in ADaM is indicated as a derived variable but the same SDTM variable name exist in SDTM variable metadata*

%adam_sdtm_compa

- *8 major CDISC principles*

  - *Check #104:* *SDTM.EX is present and neither ADSL.TRTSDT or ADSL.TRTSDTM are present*

  - *Check #105:* *The variable referenced in the computational algorithm of the ADaM variable does not exist in SDTM or ADaM variable metadata*

%adam_sdtm_compa

- *8 major CDISC principles*

  - *Check #201:* *A variable present in ADaM is indicated as a copy of a variable present in SDTM but the values are not identical*
  - *Check #202:* *The values of USUBJID are not present in SDTM.DM*
  - *Check #203:* *SRCDOM has a value that is not a SDTM dataset, an ADaM dataset name or null*

%adam_sdtm_compa

- ***Advantages***

  Easy to produce, *Easy to integrate,* *Easy to archive*

  Easy to share, *Easy to open,* *Easy to read*

- ***How to call it***

```
%adam_sdtm_compa (
    studyid= STUDY-0001,
    sdtm_path= W:\PROJECT-X\STUDY-0001\SDTM\Metadata,
    sdtm_data= W:\PROJECT-X\STUDY-0001\SDTM,
    adam_path= W:\PROJECT-X\STUDY-0001\ADAM\Metadata,
    adam_data= W:\PROJECT-X\STUDY-0001\ADAM
);
```

%adam_sdtm_compa

- ***Limitations***
  - *False positive (#101, #105)*
  - *Executable only on BDLS's conversion*

- ***Improvements***
  - *Refinement of existing checks*
  - *Addition of new checks*
  - *Develop the macro to extract metadata from the Define.xml*

Agenda

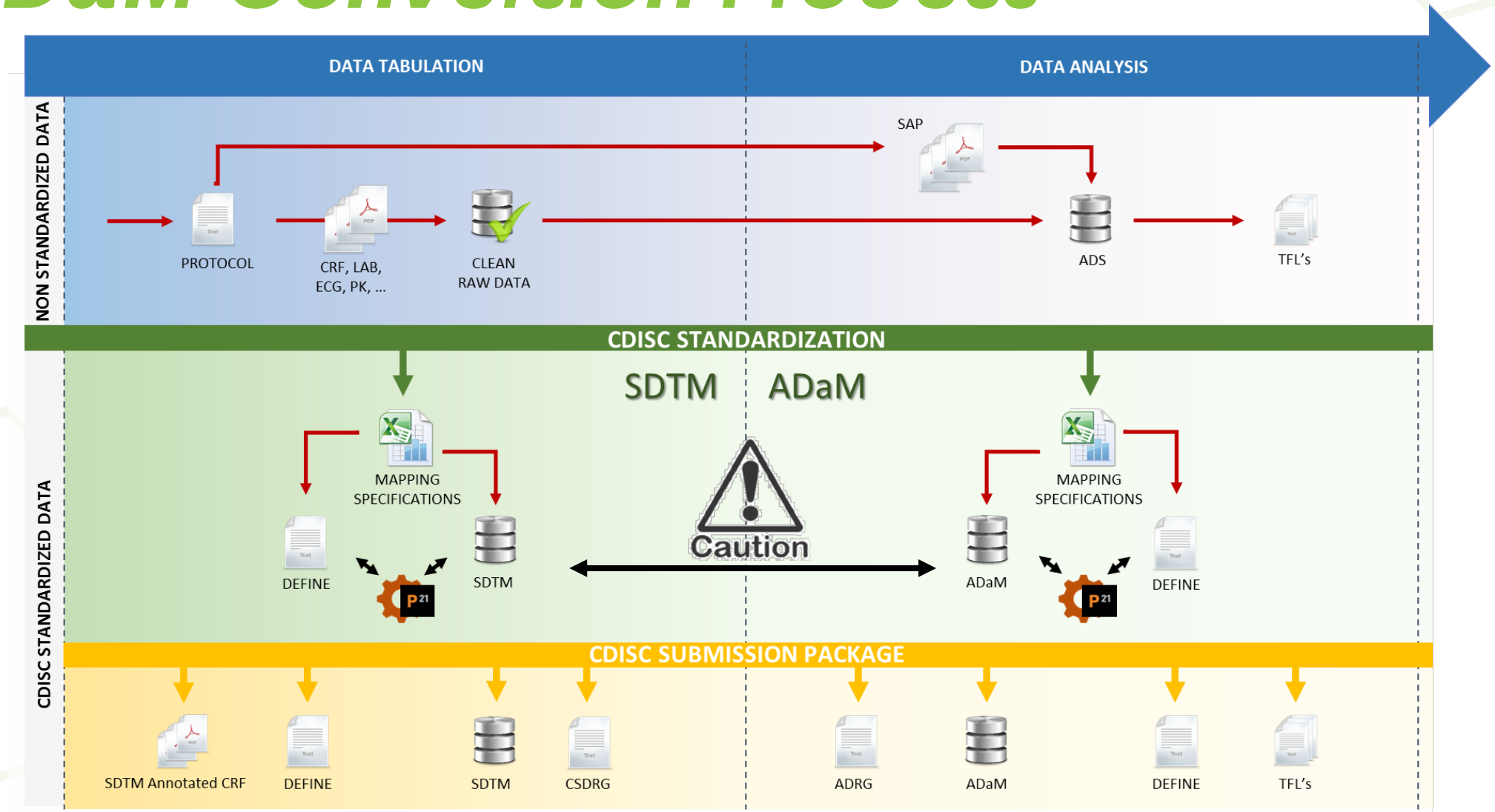
ADaM Conversion Process

%LIB_QC_contents_html

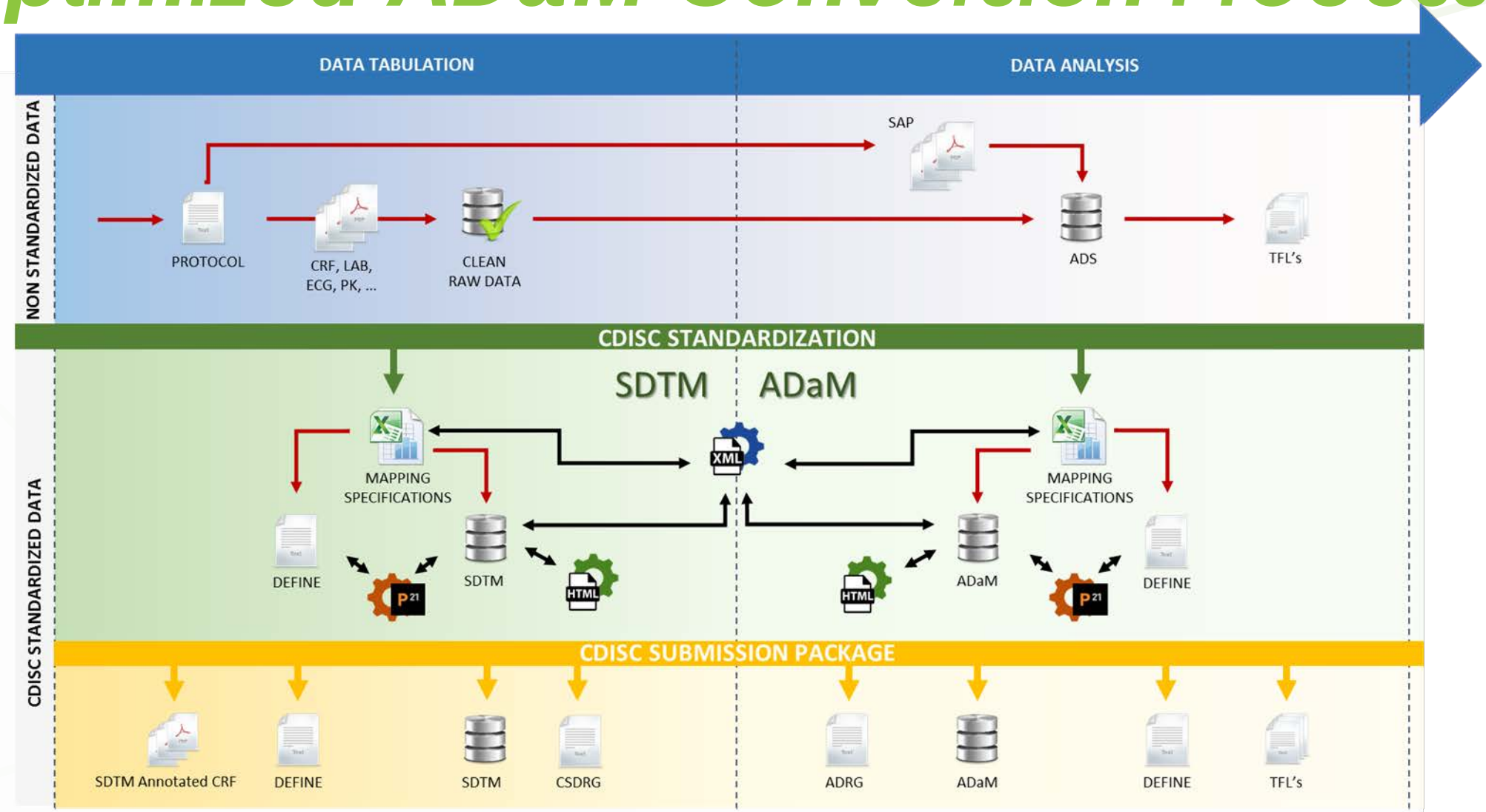
%adam_sdtm_compa

Conclusion

ADaM Conversion Process

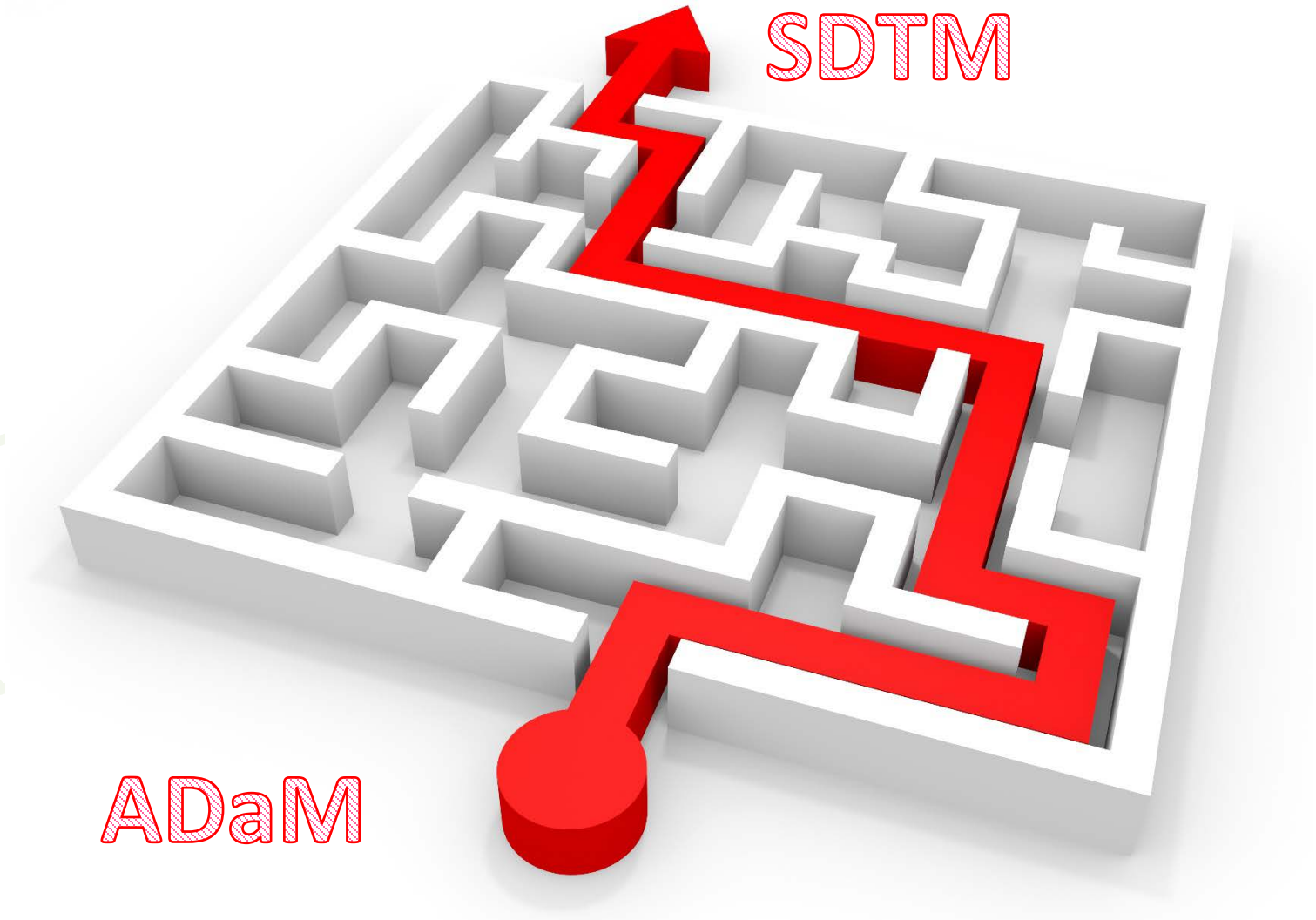


Optimized ADaM Conversion Process



Business & Decision Life Sciences

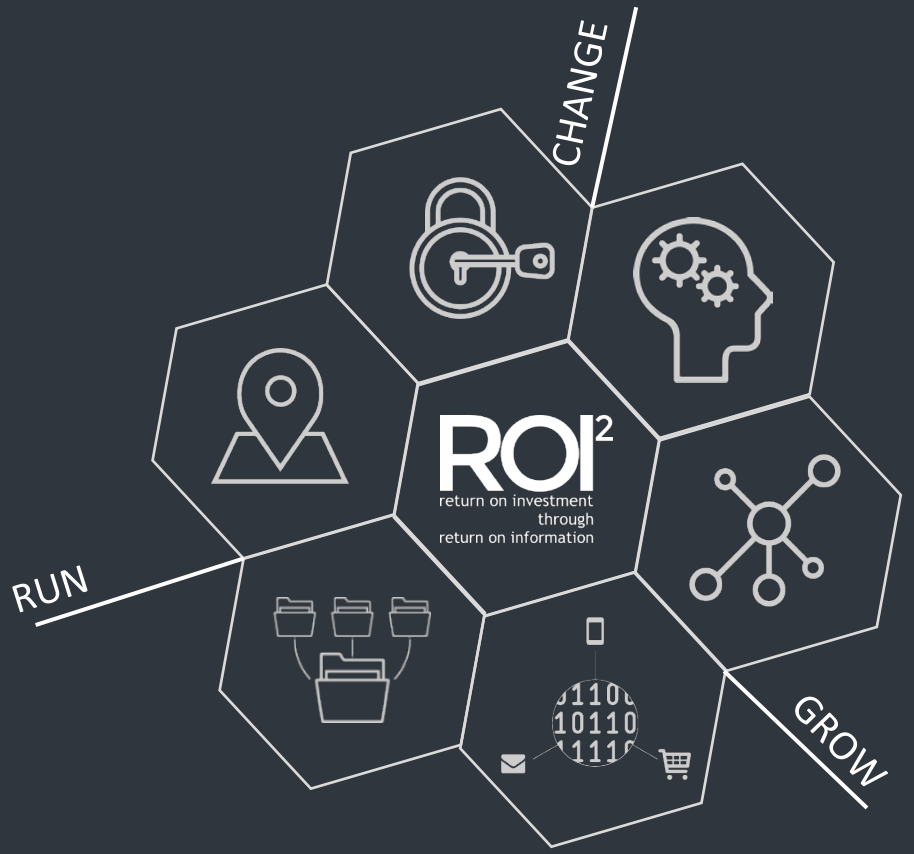
Optimized ADaM Conversion Process



Optimized ADaM Conversion Process

- ***Added Value ?***
  - *Traceability is one of the main goals of ADaM generation*
  - *Identified by FDA as one of the most problematic issues during submission*
  - *Optimization is possible with automated checks (SAS Macro)*
  - *Can be included into actual SOP*

➔ *Undoubtedly a return on investment*



ROI²
return on investment
through
return on information

Come to meet us at Phuse Booth #36

Roxane DEBRUS

Biostatistician, Project Manager
roxane.debrus@businessdecision.be
+32 475 95 35 00

Business & Decision
Life Sciences