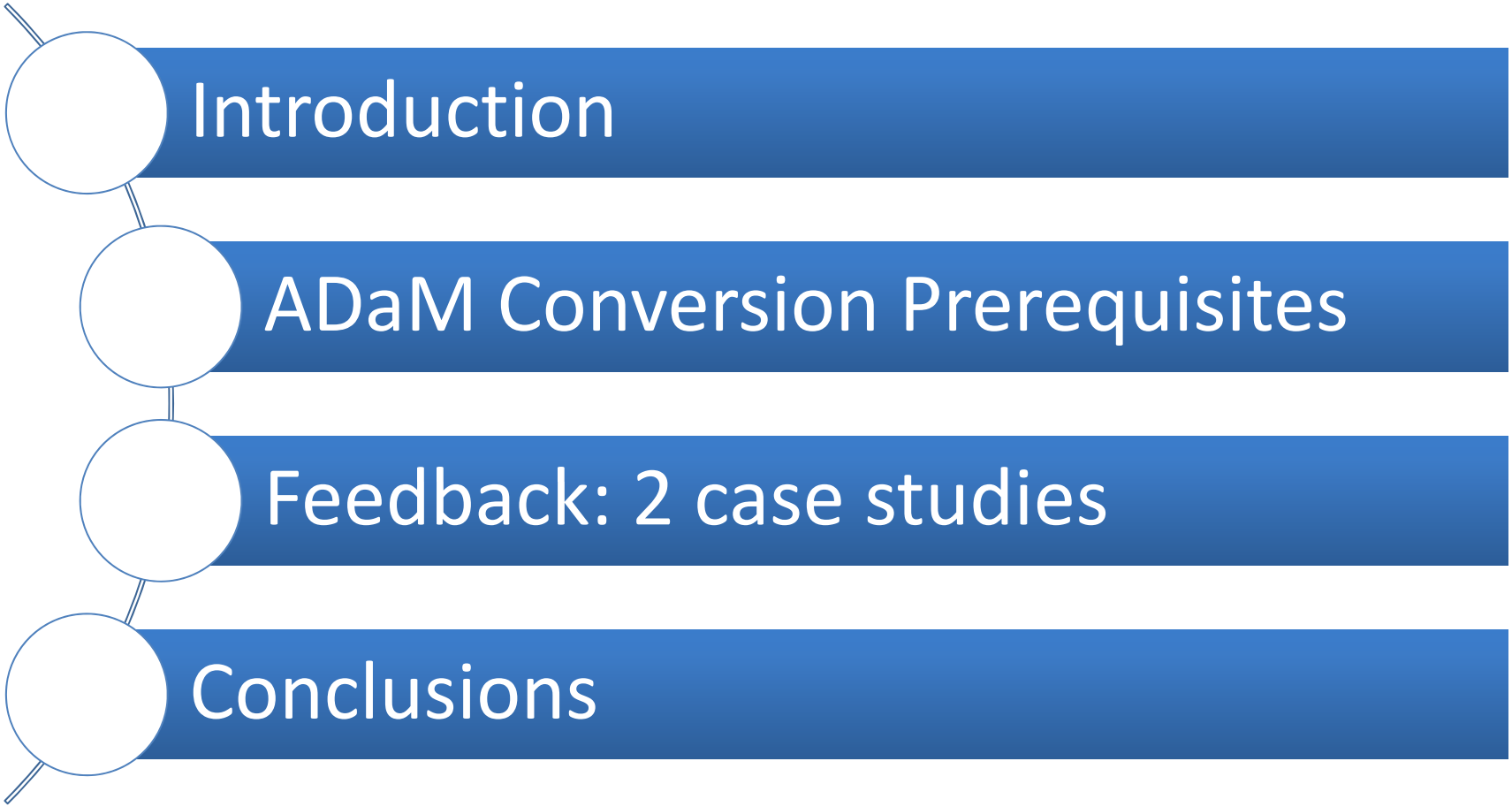
Business & Decision Life Sciences

ADaM Conversions: The Good, The Bad and the Ugly

Jessica Minkue Mi Edou

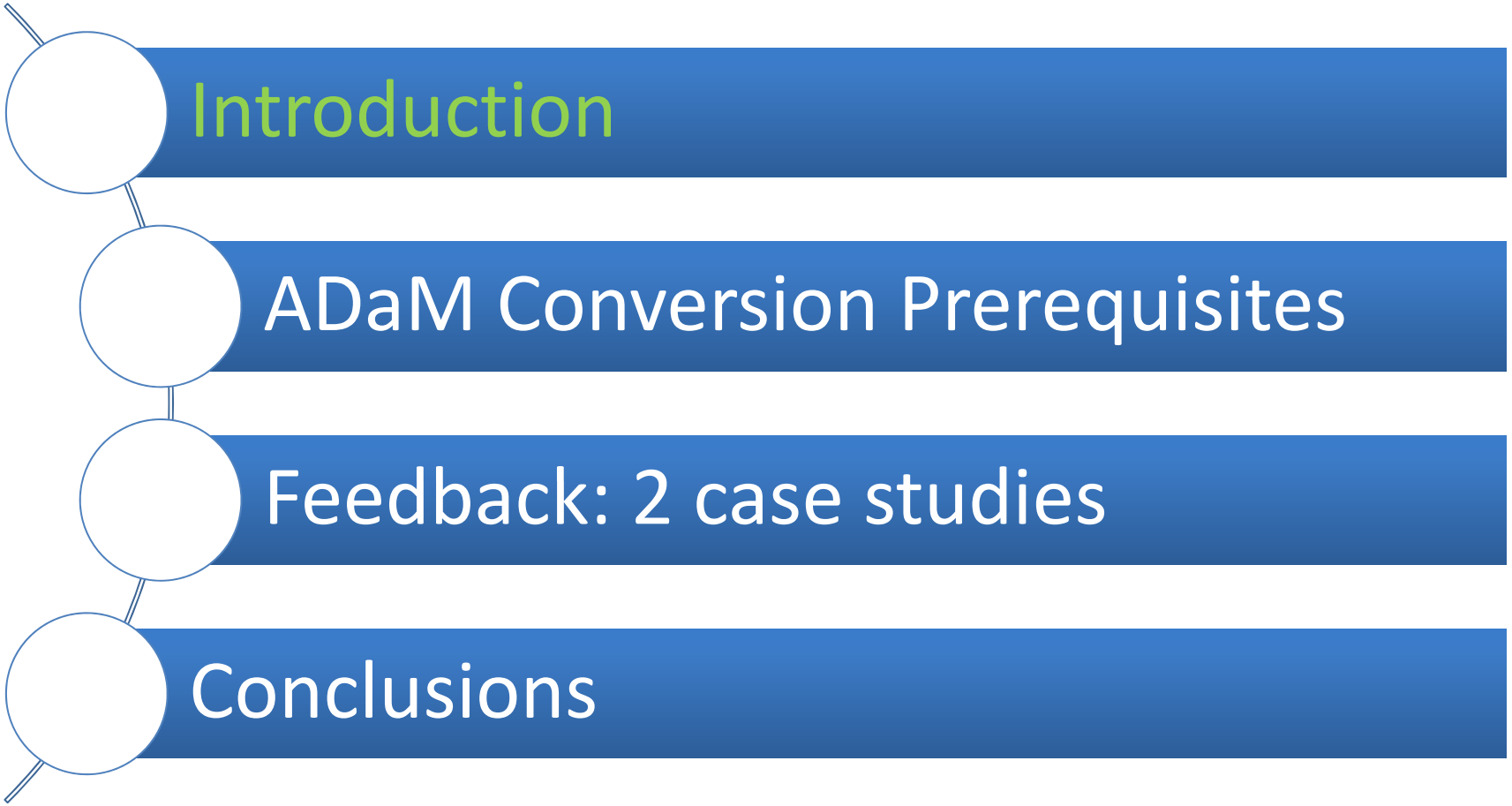
28 September 2015

ADaM Conversions:
The Good, the Bad and the Ugly



Business & Decision Life Sciences

ADaM Conversions:
The Good, the Bad and the Ugly



Business & Decision Life Sciences

Introduction

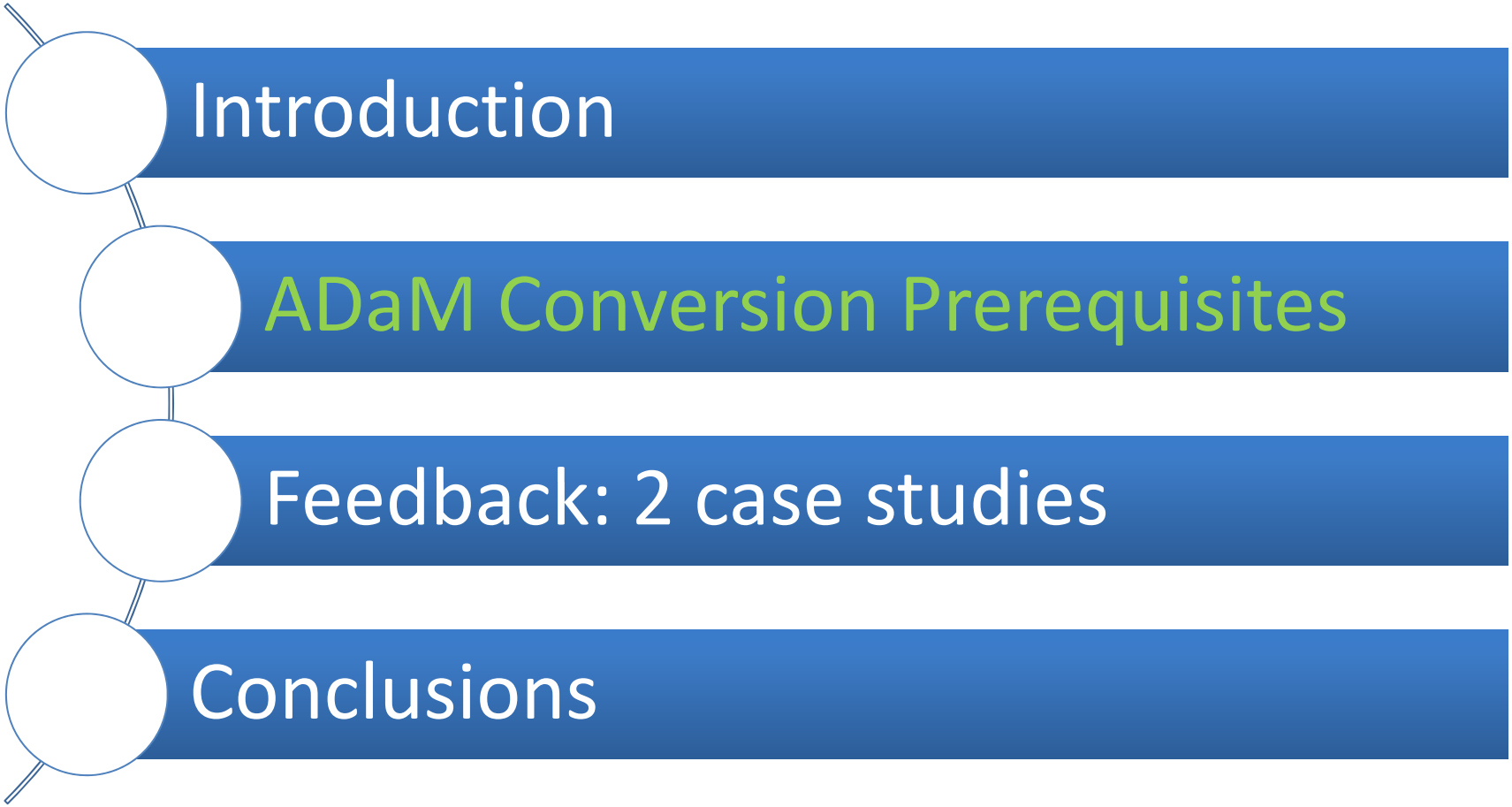
Business & Decision Life Sciences' experiences in ADaM conversion projects applying:

- Linear approach for a legacy study
- Parallel approach for Pooled datasets

Objective

Describing the good, bad and ugly experiences gained during the execution of these projects

ADaM Conversions:
The Good, the Bad and the Ugly



Business & Decision Life Sciences

ADaM Conversions Prequisites: Traceability

Good understanding of the relationship between

- Analysis results
- Analysis datasets
- SDTM domains

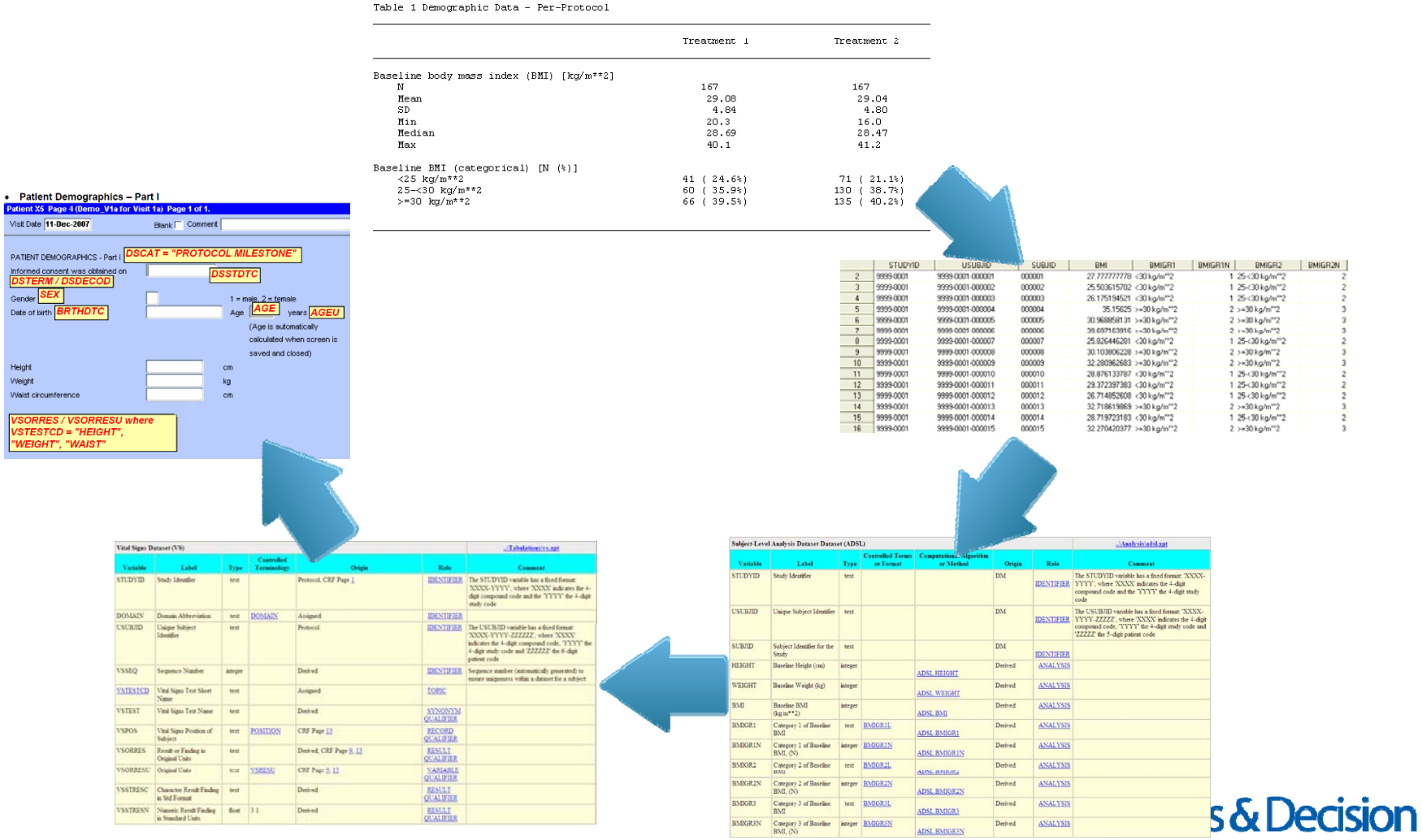
The **core** of any ADaM conversion is the SDTM/ADaM **traceability**, this can be found at two levels:

- **Metadata traceability**: finding a relationship between an analysis result and analysis dataset(s), or a relationship of the analysis variable to its source dataset(s) and variable(s)

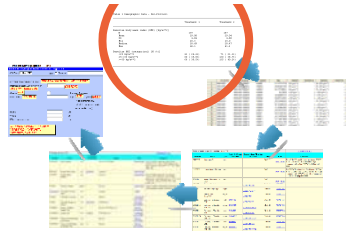
- **Data point traceability**: finding the predecessor record(s)

Business & Decision Life Sciences

ADaM Conversions Prerequisites:
Traceability



ADaM Conversions Prerequisites: Traceability



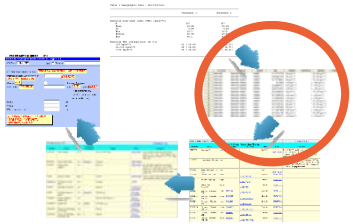
- Analysis Results

Table 1 Demographic Data - Per-Protocol

	Treatment 1	Treatment 2
Baseline body mass index (BMI) [kg/m**2]		
N	167	167
Mean	29.08	29.04
SD	4.84	4.80
Min	20.3	16.0
Median	28.69	28.47
Max	40.1	41.2
Baseline BMI (categorical) [N (%)]		
<25 kg/m**2	41 (24.6%)	71 (21.1%)
25-<30 kg/m**2	60 (35.9%)	130 (38.7%)
>=30 kg/m**2	66 (39.5%)	135 (40.2%)

Business & Decision Life Sciences

ADaM Conversions Prerequisites: Traceability

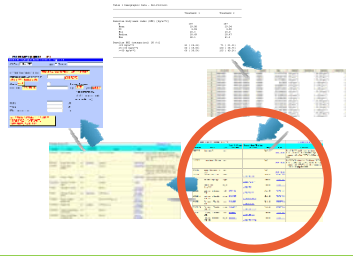


- ADaM Analysis Dataset

	STUDYID	USUBJID	SUBJID	BMI	BMIGR1	BMIGR1N	BMIGR2	BMIGR2N
2	9999-0001	9999-0001-000001	000001	27.777777778	<30 kg/m**2	1	25-<30 kg/m**2	2
3	9999-0001	9999-0001-000002	000002	25.503615702	<30 kg/m**2	1	25-<30 kg/m**2	2
4	9999-0001	9999-0001-000003	000003	26.175194521	<30 kg/m**2	1	25-<30 kg/m**2	2
5	9999-0001	9999-0001-000004	000004	35.15625	>=30 kg/m**2	2	>=30 kg/m**2	3
6	9999-0001	9999-0001-000005	000005	30.968858131	>=30 kg/m**2	2	>=30 kg/m**2	3
7	9999-0001	9999-0001-000006	000006	39.697163916	>=30 kg/m**2	2	>=30 kg/m**2	3
8	9999-0001	9999-0001-000007	000007	25.826446281	<30 kg/m**2	1	25-<30 kg/m**2	2
9	9999-0001	9999-0001-000008	000008	30.103806228	>=30 kg/m**2	2	>=30 kg/m**2	3
10	9999-0001	9999-0001-000009	000009	32.280962683	>=30 kg/m**2	2	>=30 kg/m**2	3
11	9999-0001	9999-0001-000010	000010	28.876133787	<30 kg/m**2	1	25-<30 kg/m**2	2
12	9999-0001	9999-0001-000011	000011	29.372397383	<30 kg/m**2	1	25-<30 kg/m**2	2
13	9999-0001	9999-0001-000012	000012	26.714852608	<30 kg/m**2	1	25-<30 kg/m**2	2
14	9999-0001	9999-0001-000013	000013	32.718619869	>=30 kg/m**2	2	>=30 kg/m**2	3
15	9999-0001	9999-0001-000014	000014	28.719723183	<30 kg/m**2	1	25-<30 kg/m**2	2
16	9999-0001	9999-0001-000015	000015	32.270420377	>=30 kg/m**2	2	>=30 kg/m**2	3

Business & Decision Life Sciences

ADaM Conversions Prerequisites: Traceability



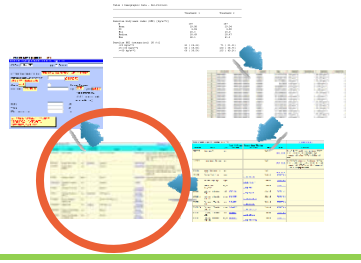
- ADaM define.xml

Computational Algorithms (ADSL.BMI)	
Reference Name	**Computation Method**
ADSL.BMI	Continuous variable, calculated using ADSL.WEIGHT/(ADSL.HEIGHT*0,01)**2 value at visit 3 if visit 3 data not available, the last data collected before randomisation
Computational Algorithms (ADSL.HEIGHT)	
Reference Name	**Computation Method**
ADSL.HEIGHT	equal to VS.VSSTRESN when VS.VSTESTCD="HEIGHT"
Computational Algorithms (ADSL.WEIGHT)	
Reference Name	**Computation Method**
ADSL.WEIGHT	equal to VS.VSSTRESN when VS.VSTESTCD="WEIGHT" and VS.VISITNUM=30 if visit 3 data not available, the last data collected before randomisation

				ADSL.HEIGHT			
WEIGHT	Baseline Weight (kg)	integer		ADSL.WEIGHT	Derived	ANALYSIS	
BMI	Baseline BMI (kg/m**2)	integer		ADSL.BMI	Derived	ANALYSIS	
BMIGR1	Category 1 of Baseline BMI	text	BMIGR1L	ADSL.BMIGR1	Derived	ANALYSIS	
BMIGR1N	Category 1 of Baseline BMI, (N)	integer	BMIGR1N	ADSL.BMIGR1N	Derived	ANALYSIS	
BMIGR2	Category 2 of Baseline BMI	text	BMIGR2L	ADSL.BMIGR2	Derived	ANALYSIS	
BMIGR2N	Category 2 of Baseline BMI, (N)	integer	BMIGR2N	ADSL.BMIGR2N	Derived	ANALYSIS	
BMIGR3	Category 3 of Baseline BMI	text	BMIGR3L	ADSL.BMIGR3	Derived	ANALYSIS	
BMIGR3N	Category 3 of Baseline BMI, (N)	integer	BMIGR3N	ADSL.BMIGR3N	Derived	ANALYSIS	

Business & Decision Life Sciences

ADaM Conversions Prerequisites: Traceability



- SDTM define.xml and aCRF

Value Level Metadata (ValueList.VS.VSTESTCD)

Source Variable	Value	Label	Type	Controlled Terminology	Origin	Role	Comment
VSTESTCD	DIABP	DIASTOLIC BLOOD PRESSURE	text		CRF Page 13		
VSTESTCD	HEIGHT	HEIGHT	text		CRF Page 9		
VSTESTCD	PULSE	PULSE RATE	text		CRF Page 13		
VSTESTCD	SYSBP	SYSTOLIC BLOOD PRESSURE	text		CRF Page 13		
VSTESTCD	WAIST	WAIST CIRCUMFERENCE	text		CRF Page 9		
VSTESTCD	WEIGHT	WEIGHT	text		CRF Page 9		

			patient code
VSSEQ	Sequence Number	ER	Sequence number (automatically generated) to ensure uniqueness within a dataset for a subject
VSTESTCD	Vital Signs Test Short Name		
VSTEST	Vital Signs Test Name	M	
VSPOS	Vital Signs Position of Subject		
VSORRES	Result or Finding in Original Units		
VSORRESU	Original Units		
VSSTRESC	Character Result/Findi in Std Format		
VSSTRESN	Numeric Result/Finding in Standard Units		

Patient Demographics – Part I

Patient X5 Page 4 (Demo_V1a for Visit 1a) Page 1 of 1.

Visit Date 11-Dec-2007 Blank ☐ Comment

PATIENT DEMOGRAPHICS - Part I — DSCAT = "PROTOCOL MILESTONE"

Informed consent was obtained on — DSSTDTC

DSTERM / DSDECOD

Gender — SEX

1 = male, 2 = female

Date of birth — BRTHDTC

Age — AGE years AGEU

(Age is automatically calculated when screen is saved and closed)

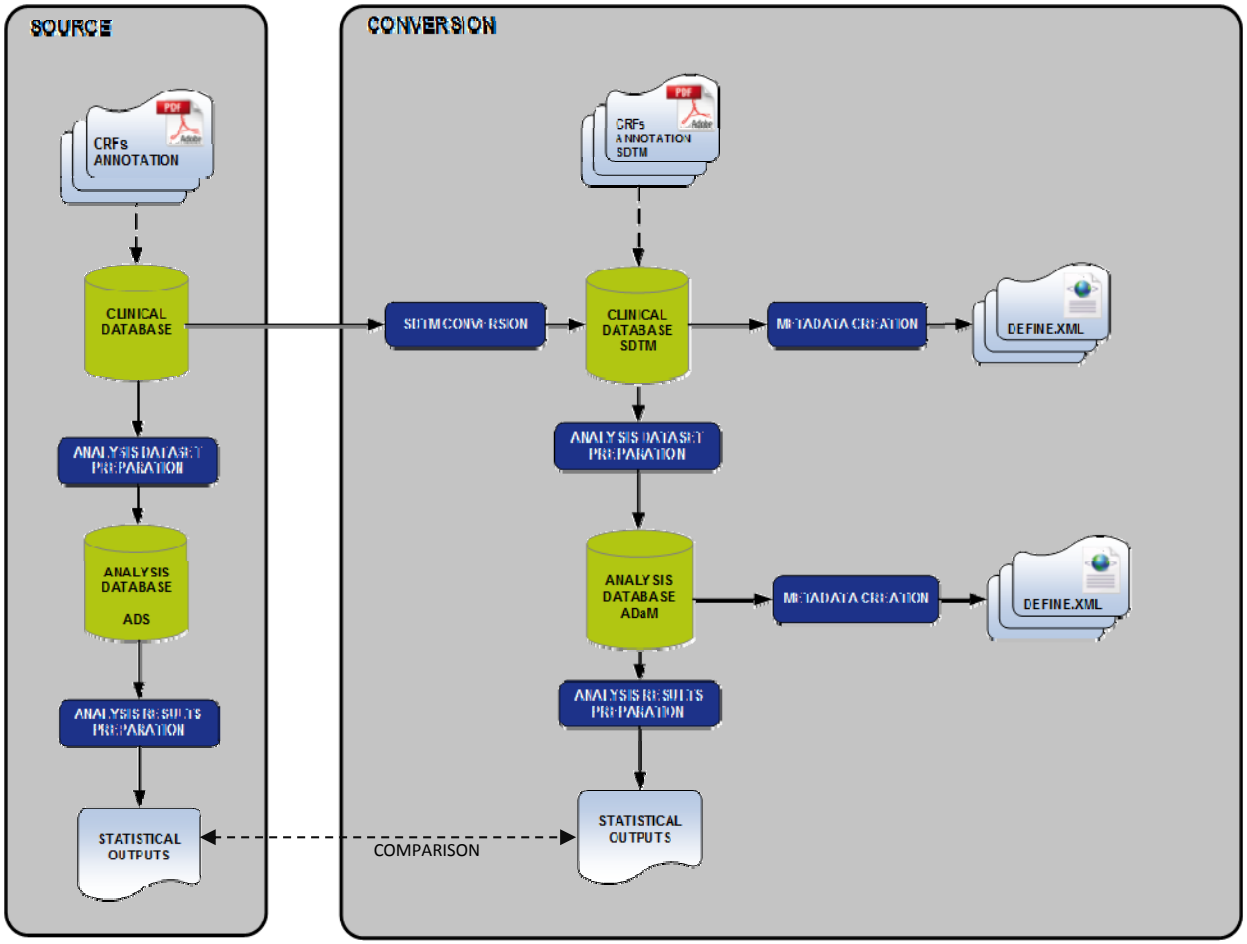
Height cm

Weight kg

Waist circumference cm

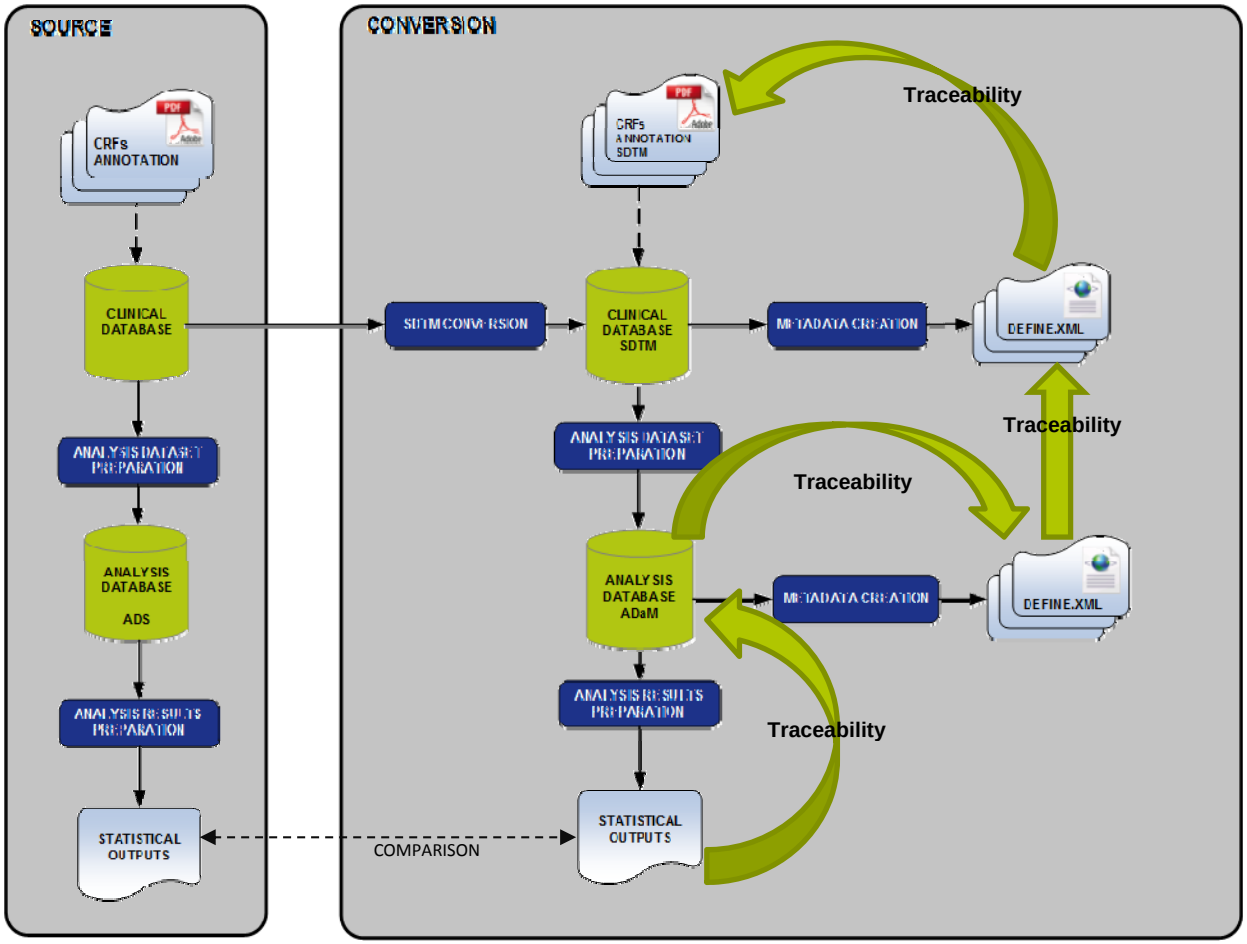
VSORRES / VSORRESU where VSTESTCD = "HEIGHT", "WEIGHT", "WAIST"

ADaM Conversions Prerequisites:
Linear Approach



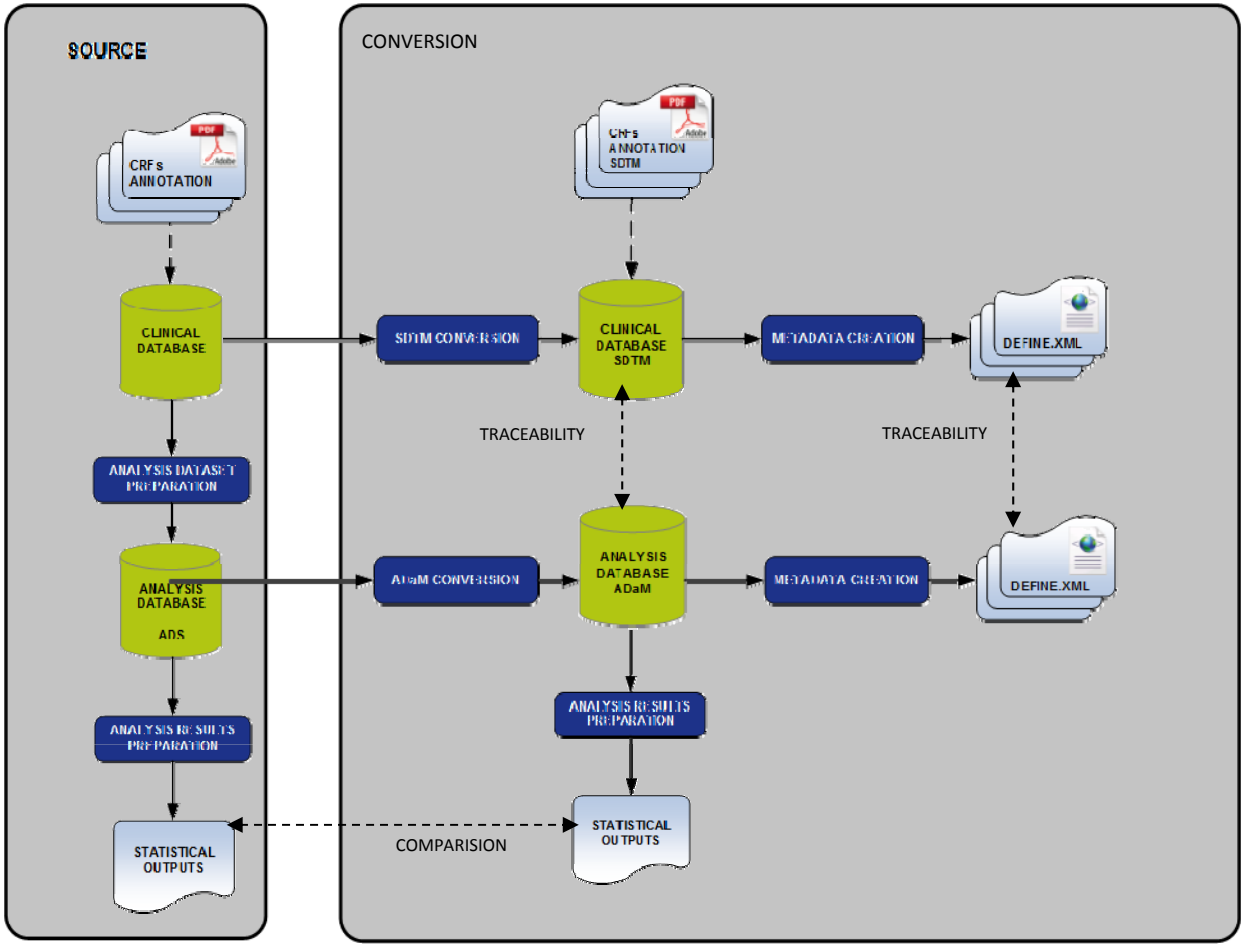
Business & Decision Life Sciences

ADaM Conversions Prerequisites:
Linear Approach



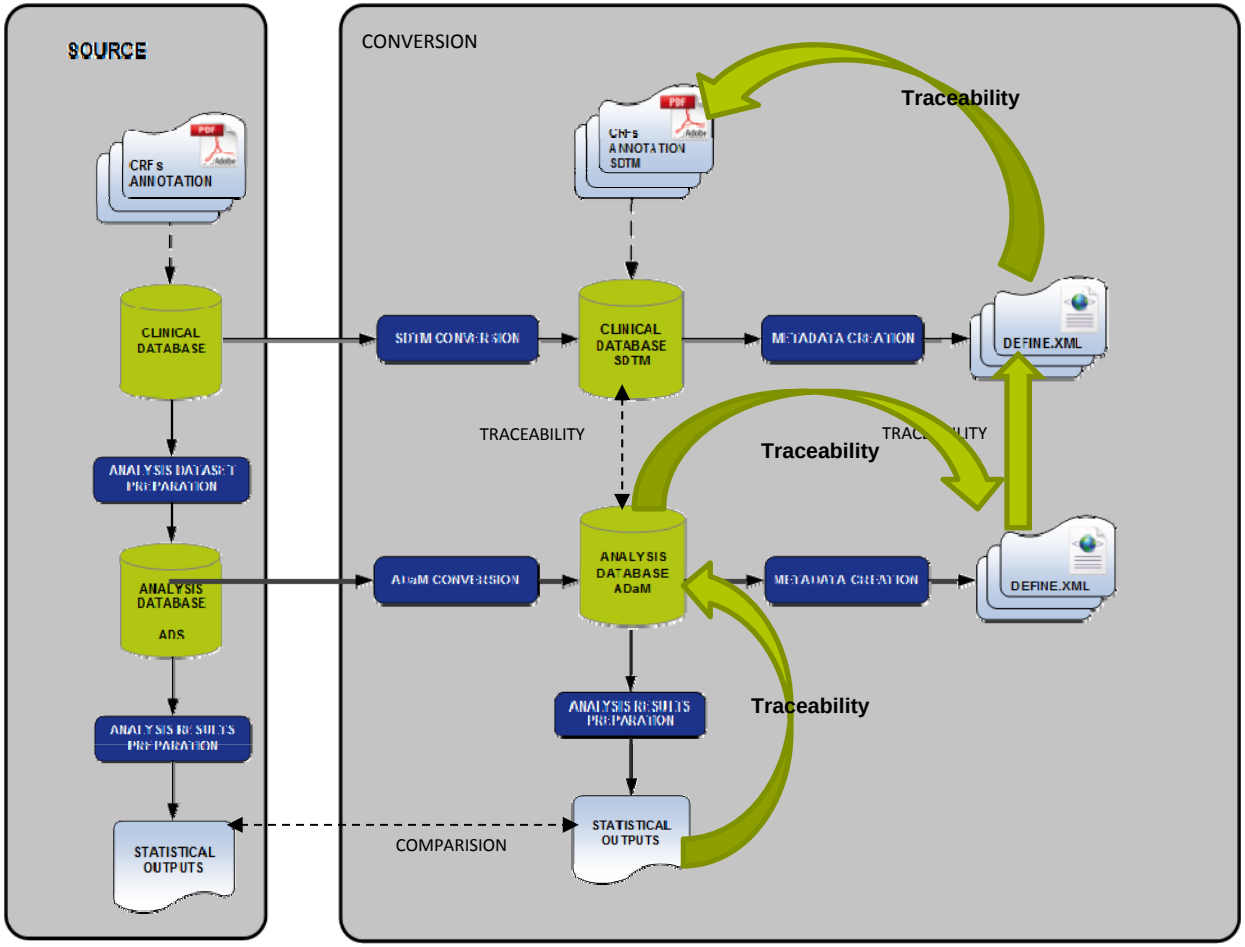
Business & Decision Life Sciences

ADaM Conversions Prerequisites:
Parallel Approach



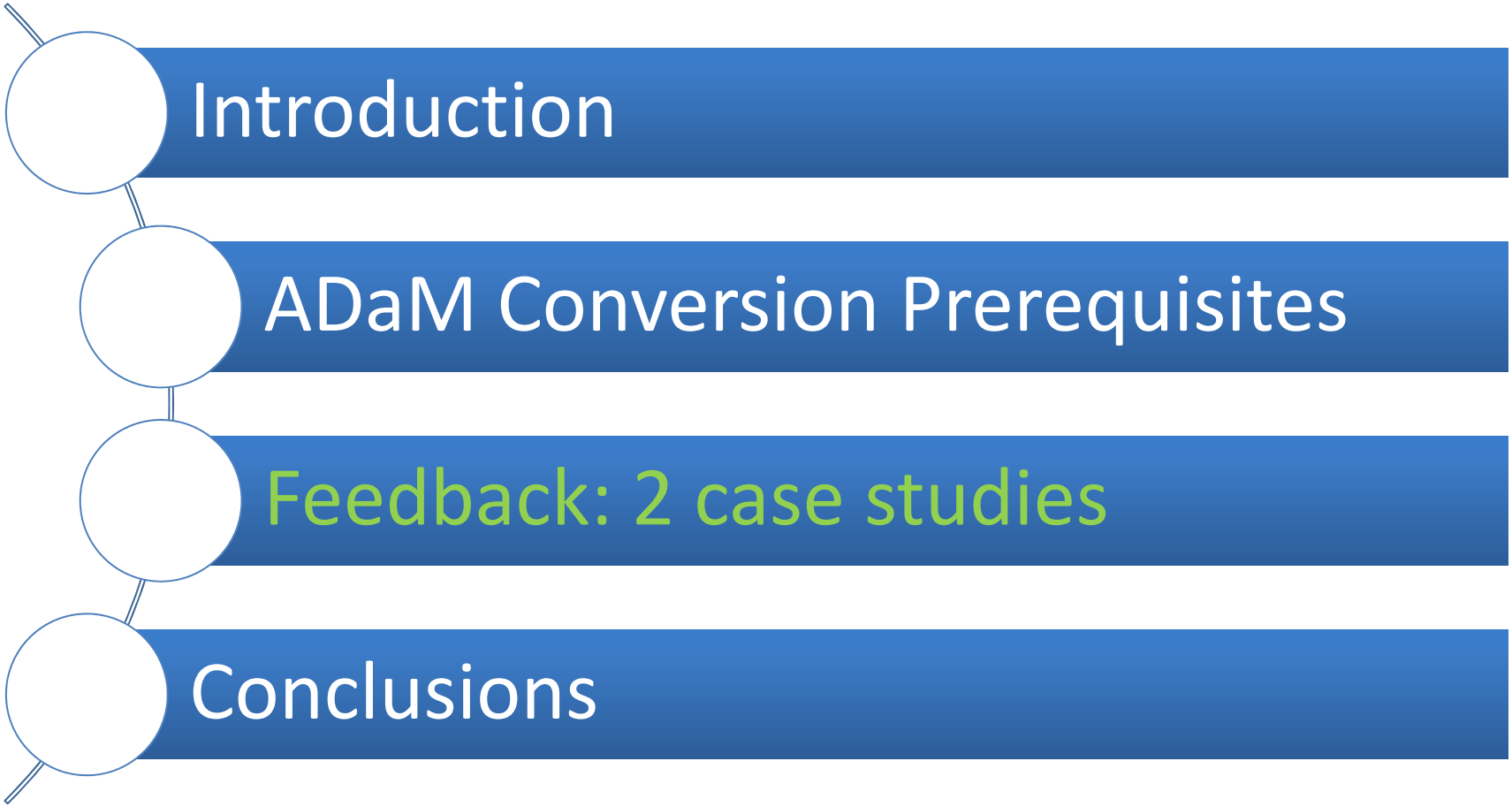
Business & Decision Life Sciences

ADaM Conversions Prerequisites:
Parallel Approach



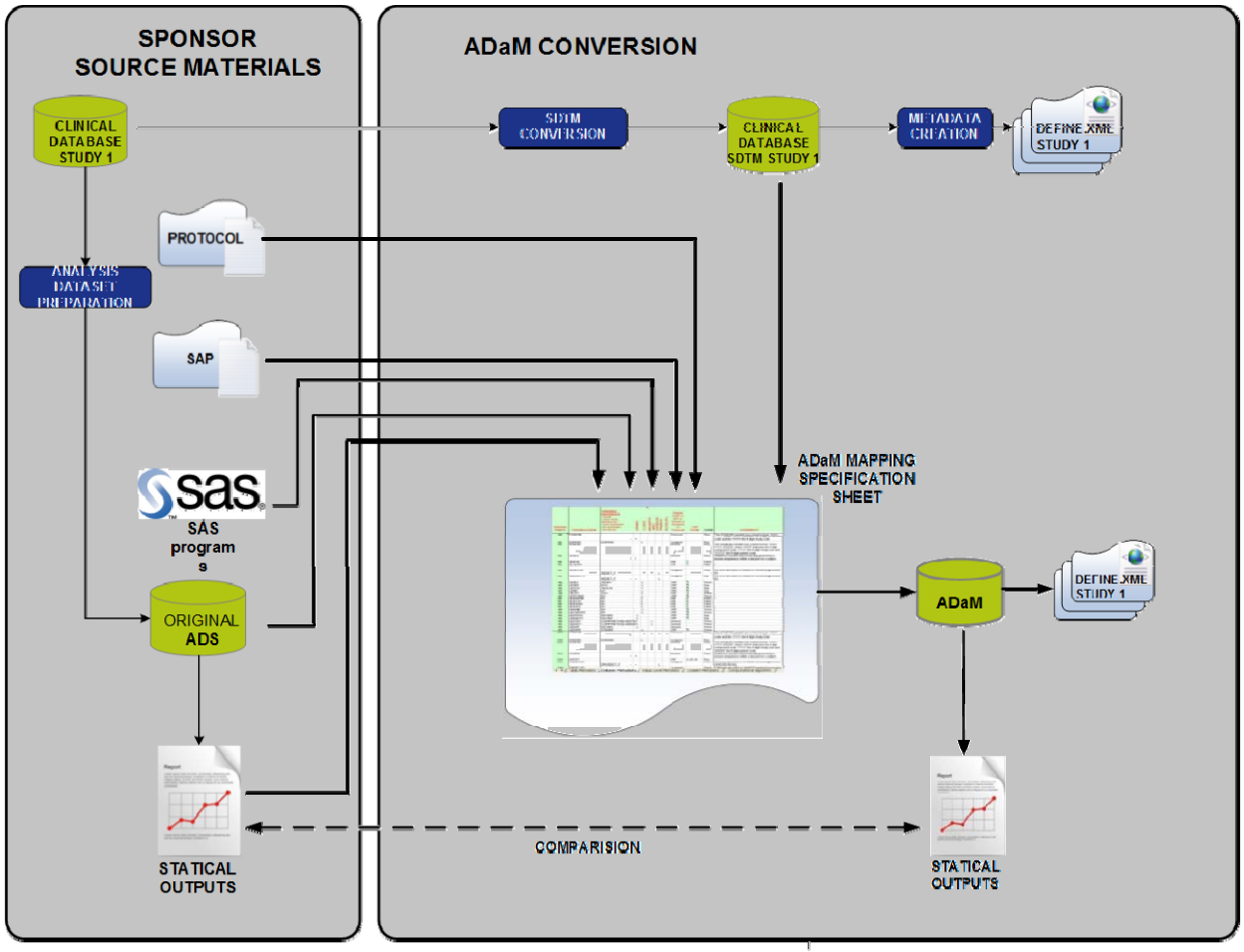
Business & Decision Life Sciences

ADaM Conversions:
The Good, the Bad and the Ugly



Business & Decision Life Sciences

Case Study 1: ADaM Conversion with SAP and TFLs as sole Sources



Business & Decision Life Sciences

Case Study 1: ADaM Conversion with SAP and TFLs as sole Sources

THE GOOD

- **Good Traceability** between the statistical outputs, the ADaM datasets and the SDTM domains
- Since BDLS was responsible for the **SDTM Conversion**, issues with SDTM domains were resolved rapidly
- Having the ADS in ADaM structure facilitated the addition of this legacy study as part of an **Analysis Pooled Database**

Business & Decision Life Sciences

Case Study 1: ADaM Conversion with SAP and TFLs as sole Sources

THE BAD

- **Hidden differences with SDTM** found during validation:
  - Different Coding Dictionary Versions in SDTM database
  - For consistency with newer studies, some categorical variables were updated with new combination of values
- Due to **SDTM standards restrictions** some values used for analysis were not available in the SDTM datasets
  - Key variables in SDTM domains cannot be empty

→ If the initial analysis missing values were summarized in a table, these values were not available anymore in the SDTM datasets

Business & Decision Life Sciences

Case Study 1: ADaM Conversion with SAP and TFLs as sole Sources

T

H

E

- **ADaM specifications** created from SAP and protocol
  - Rules specified in Protocol and SAP were differently interpreted in the same way as in the original analysis

U

→ Multiple solutions were tried before having the proper result

G

- **QC:** manual comparisons were needed
  - **ADS** not available
  - **Listings** in pdf format

L

Y

→ Resolving all these issues was time consuming

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

- Parrallel Approach

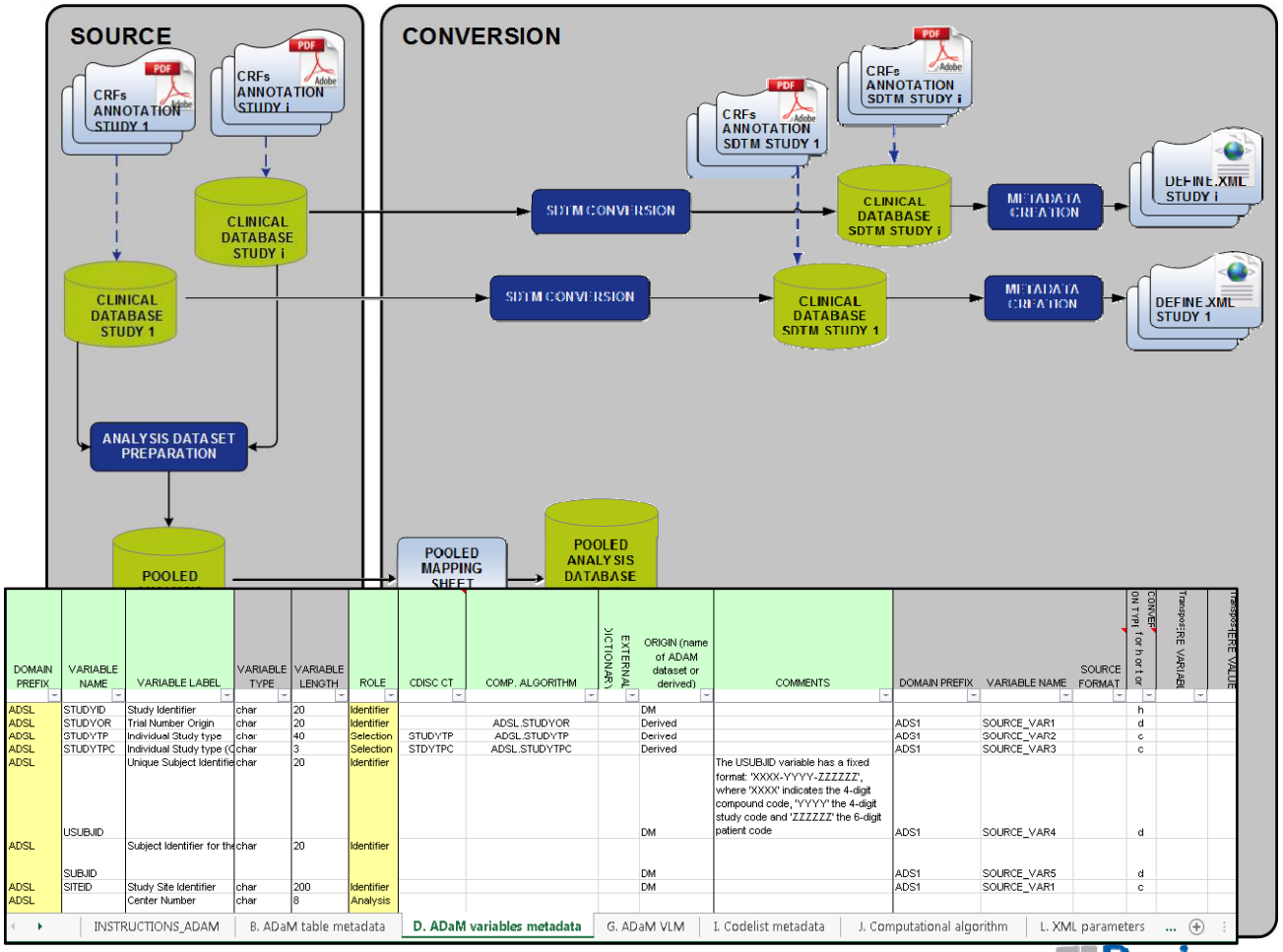
- Creation of a **pooled ADaM mapping**
  - From Sponsor's specifications

- Creation of **pooled ADaM datasets**
  - From Sponsor's Analysis datasets

- Extraction of the **individual ADaM mapping and datasets**

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

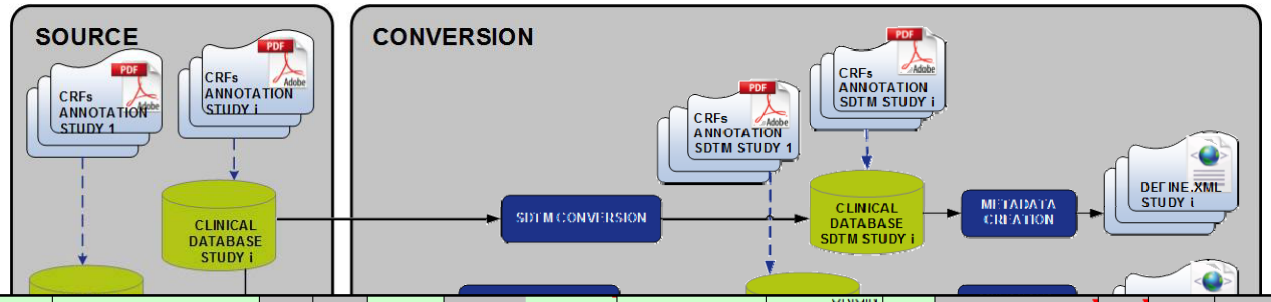


DOMAIN PREFIX	VARIABLE NAME	VARIABLE LABEL	VARIABLE TYPE	VARIABLE LENGTH	ROLE	CDISC CT	COMP. ALGORITHM	EXTERNAL DICTIONARY	ORIGIN (name of ADAM dataset or derived)	COMMENTS	DOMAIN PREFIX	VARIABLE NAME	SOURCE FORMAT	CONVERSION TYPE (c or h or t or	Transpose VARIABLE	Transpose VALUE
ADSL	STUDYID	Study Identifier	char	20	Identifier				DM					h		
ADSL	STUDYOR	Trial Number Origin	char	20	Identifier		ADSL.STUDYOR		Derived		ADS1	SOURCE_VAR1		d		
ADSL	STUDYTP	Individual Study type	char	40	Selection	STUDYTP	ADSL.STUDYTP		Derived		ADS1	SOURCE_VAR2		c		
ADSL	STUDYTPC	Individual Study type (C	char	3	Selection	STDYTPC	ADSL.STUDYTPC		Derived		ADS1	SOURCE_VAR3		c		
ADSL	USUBJID	Unique Subject Identifie	char	20	Identifier				DM	The USUBJID variable has a fixed format: 'XXXX-YYYY-ZZZZZZ', where 'XXXX' indicates the 4-digit compound code, 'YYYY' the 4-digit study code and 'ZZZZZZ' the 6-digit patient code	ADS1	SOURCE_VAR4		d		
ADSL	SUBJID	Subject Identifier for the	char	20	Identifier				DM		ADS1	SOURCE_VAR5		d		
ADSL	SITEID	Study Site Identifier	char	200	Identifier				DM		ADS1	SOURCE_VAR1		c		
ADSL		Center Number	char	8	Analysis											

INSTRUCTIONS_ADAM | B. ADaM table metadata | D. ADaM variables metadata | G. ADaM VLM | I. Codelist metadata | J. Computational algorithm | L. XML parameters ...

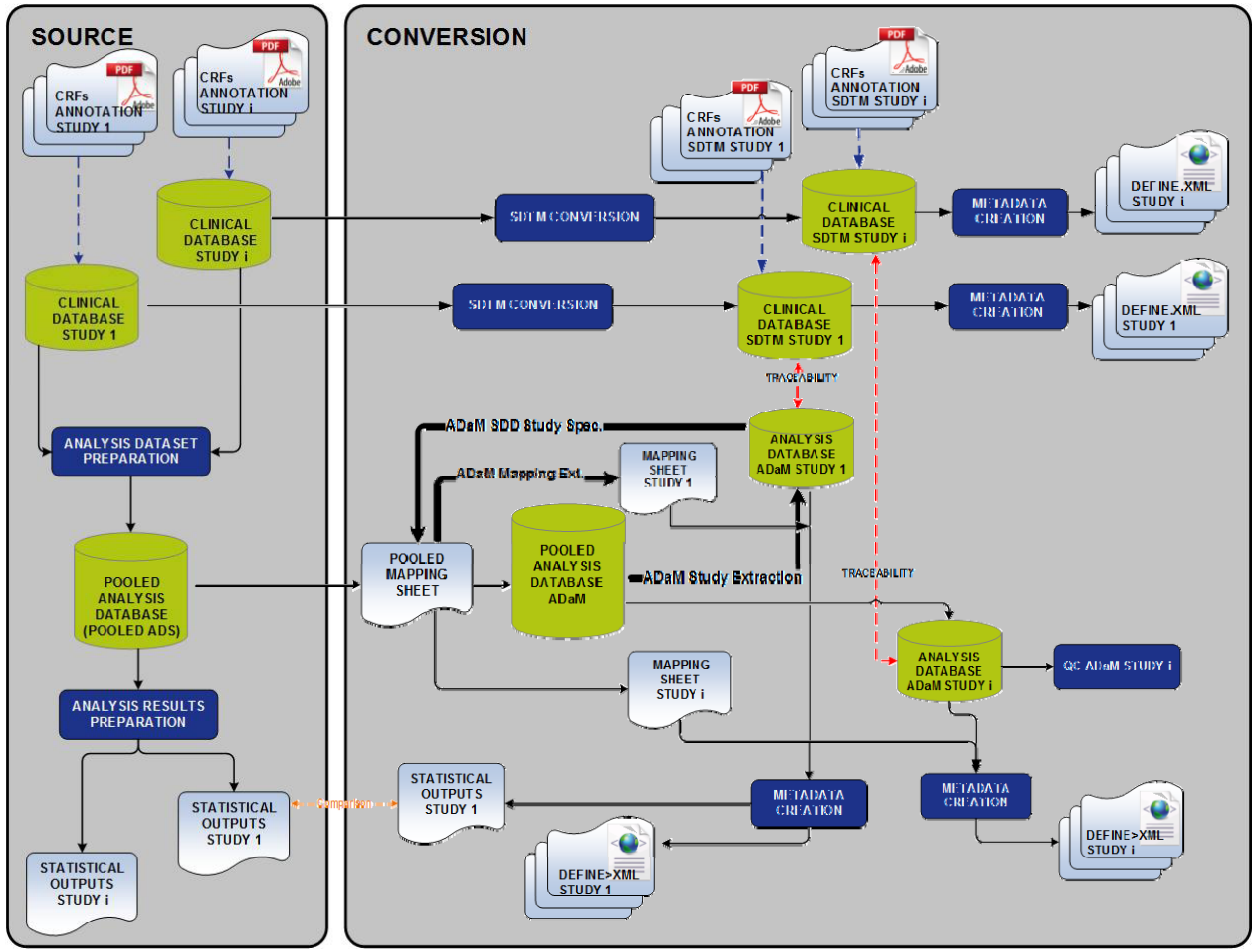
Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset



DOMAIN PREFIX	VARIABLE NAME	VARIABLE LABEL	VARIABLE TYPE	VARIABLE LENGTH	ROLE	USER DEFINED FORMAT	CDISC CT	COMP. ALGORITHM	EXTERNAL DICTIONARY	ORIGIN (name of ADAM dataset	COMMENTS	MAIN PRE	VARIABLE NAME	SOURCE FORMAT	c or f or h or	STUDY 0001	STUDY 0002	STUDY 0003	STUDY 0004	STUDY 0005	STUDY 0006
ADAE	AESDT	Start Date of Adverse Event	num	8	Timing	DATE9.		ADAE.AESDT		Derived		AE	AEOND		c	YES	YES	YES	YES	YES	YES
ADAE	AESTM	StartTime of Adverse Event	num	8	Timing	TIME8.		ADAE.AESTM		Derived		AE	AEONT		c	YES	YES	YES	YES	YES	YES
ADAE	AESTDTC	Start Date/Time of Adverse Event	char	20	Analysis					AE		AE	AEONY		d	YES	YES	YES	YES	YES	YES
ADAE	AESDTF	Start Date Imputation Flag	char	3	Analysis		DATEFL	ADAE.AESDTF		Derived		AE	AEONIMP/AEONYMD		d	YES	YES	YES	YES	YES	YES
ADAE	AEEDT	End Date of Adverse Event	num	8	Timing	DATE9.		ADAE.AEEDT		Derived		AE	AEEND		c	YES	YES	YES	YES	YES	YES
ADAE	AEENDTC	End Date/Time of Adverse Event	char	20	Analysis					AE		AE	AEEND		d	YES	YES	YES	YES	YES	YES
ADAE	AEDTF	Analysis End Date Imputation Fla	char	3	Analysis		DATEFL	ADAE.AEDTF		Derived		AE	AEENDIMP/AEENDYM		d	YES	YES	YES	YES	YES	YES
ADAE	AEENRF	End Relative to Reference Period	char	20	Analysis		AEENRF			AE		AE	AEEND		f	YES	YES	YES	YES	YES	YES
ADAE	BGLU	Blood Glucose Value	char	25	Analysis		BLOODGLU	ADAE.BGLU		Derived		AE	BGLU		f	YES	YES	YES	YES	YES	YES
ADAE	BGLUN	Blood Glucose Value (N)	num	8	Analysis		BLOODGLN	ADAE.BGLUN		Derived		AE	BGLU		c	YES	YES	YES	YES	YES	YES
ADAE	AETERM	Reported Term for the Adverse Ev	char	200	Analysis					AE		AE	AEMN		c	YES	YES	YES	YES	YES	YES
ADAE	AESYM	Adverse Event Symptom	char	40	Analysis		AESYM	ADAE.AESYM		Derived		AE	SYMAE		f	NO	NO	NO	#N/A	NO	NO
ADAE	AESYMN	Adverse Event Symptom (N)	num	8	Analysis		AESYMN	ADAE.AESYMN		Derived		AE	SYMAE		c	NO	NO	NO	#N/A	NO	NO
ADAE	AELLT	Lowest Level Term	char	200	Analysis					AE		AE	AELLT		c	YES	YES	YES	YES	YES	YES
ADAE	AELLTCD	Lowest Level Term Code	char	10	Analysis					AE		AE	AELLTC		c	YES	YES	YES	YES	YES	YES
ADAE	AEACN	Action Taken With Study Treatm	char	40	Analysis		AEACN			AE		AE	AEACTA/AEACTE		d	YES	YES	YES	YES	YES	YES

Case Study 2:
ADaM Conversion of a Pooled Dataset



Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

T

H

E

- BDLS was in charge of the SDTM and ADaM conversion, this **facilitated the communication** between the two teams and accelerated the problem solving

G

- With a **mature Pooled ADaM conversion process**, data extraction for each study is fast, and consistency checks across studies are no longer needed

O

- **Codelist were re-usable** in ADaM due to its consistency across studies

O

- A **comment file** with all questions from and to the sponsor, facilitates the tracking of all answers and decisions

D

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

T

H

E

- Not well documented **Derivation Rules** caused difficulties to define the traceability

B

- **Studies were added into the pooled database** at different time points, while derivations were not properly updated

A

- **SDTM limitations** affected the ADaM conversion
  - some raw variables were not available in SDTM, but were mentioned in ADS specifications
  - Specifications had to be redefined in cooperation with the sponsor

D

- Due to traceability between SDTM/ADaM, **timelines** for ADaM datasets were influenced by SDTM timelines

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

T H E U G L Y

- **Structure of the pooled original ADS datasets** changed with the inclusion of new studies

→ Impact on the pooled mapping, pooled datasets and extractions

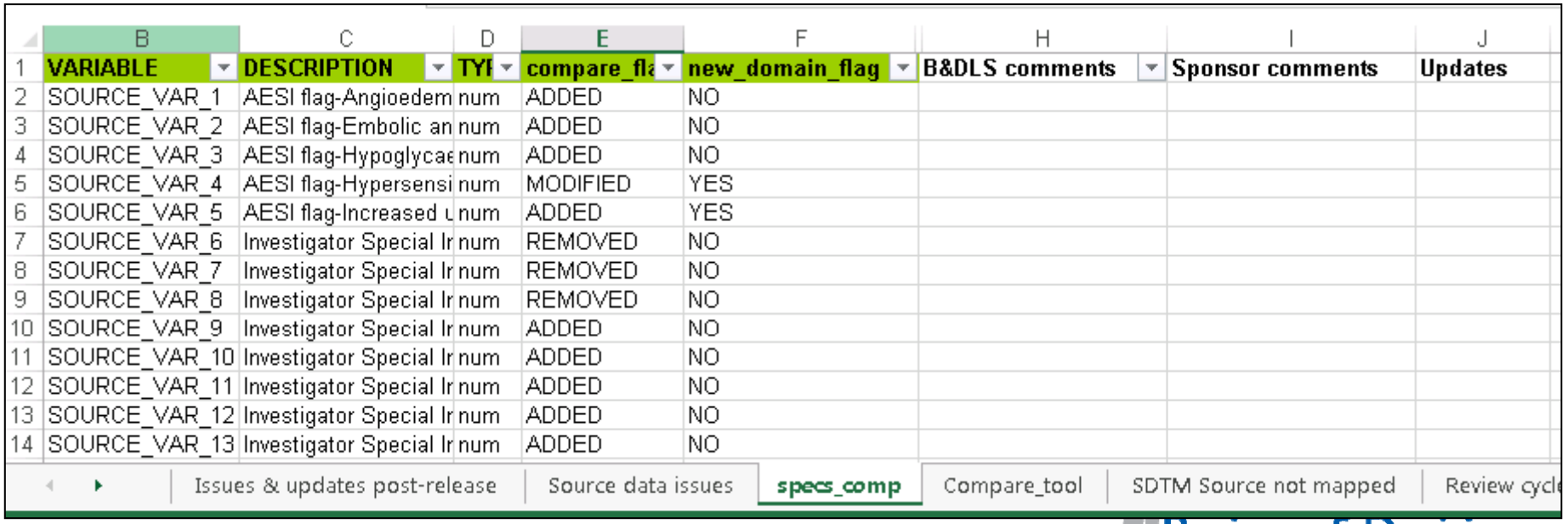
- **Source variables** which had some manipulation during the conversion (e.g. variable derived, with a format or transposed) were **sensitive to changes in its content** at each transfer

→ New compare tools to check changes in dataset structure between the original ADS received

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

- BDLS developed **new compare tools** to check changes in dataset structure between the original ADS received, like number of records, empty, new, and missing variables that were present before
- **Specifications** were **compared**: results of a compare were included in the communication file to the sponsor



	B	C	D	E	F	H	I	J
1	VARIABLE	DESCRIPTION	TYP	compare_fla	new_domain_flag	B&DLS comments	Sponsor comments	Updates
2	SOURCE_VAR_1	AESI flag-Angioedem	num	ADDED	NO			
3	SOURCE_VAR_2	AESI flag-Embolic an	num	ADDED	NO			
4	SOURCE_VAR_3	AESI flag-Hypoglycae	num	ADDED	NO			
5	SOURCE_VAR_4	AESI flag-Hypersensi	num	MODIFIED	YES			
6	SOURCE_VAR_5	AESI flag-Increased u	num	ADDED	YES			
7	SOURCE_VAR_6	Investigator Special Ir	num	REMOVED	NO			
8	SOURCE_VAR_7	Investigator Special Ir	num	REMOVED	NO			
9	SOURCE_VAR_8	Investigator Special Ir	num	REMOVED	NO			
10	SOURCE_VAR_9	Investigator Special Ir	num	ADDED	NO			
11	SOURCE_VAR_10	Investigator Special Ir	num	ADDED	NO			
12	SOURCE_VAR_11	Investigator Special Ir	num	ADDED	NO			
13	SOURCE_VAR_12	Investigator Special Ir	num	ADDED	NO			
14	SOURCE_VAR_13	Investigator Special Ir	num	ADDED	NO			

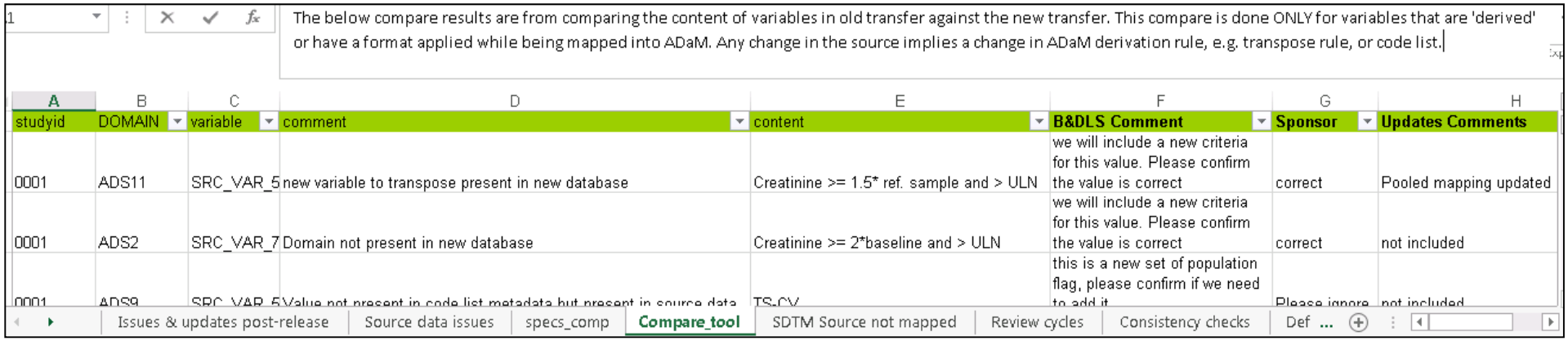
Issues & updates post-release | Source data issues | specs_comp | Compare_tool | SDTM Source not mapped | Review cycl

Business & Decision Life Sciences

Case Study 2:
ADaM Conversion of a Pooled Dataset

- **ADS content** was **compared**:

  Results from content compare for variables being “derived”, “with a CL applied”, “transposed” during ADaM conversion, were included in the communication file to the sponsor



The below compare results are from comparing the content of variables in old transfer against the new transfer. This compare is done ONLY for variables that are 'derived' or have a format applied while being mapped into ADaM. Any change in the source implies a change in ADaM derivation rule, e.g. transpose rule, or code list.

studyid	DOMAIN	variable	comment	content	B&DLS Comment	Sponsor	Updates Comments
0001	ADS11	SRC_VAR_5	new variable to transpose present in new database	Creatinine >= 1.5* ref. sample and > ULN	we will include a new criteria for this value. Please confirm the value is correct	correct	Pooled mapping updated
0001	ADS2	SRC_VAR_7	Domain not present in new database	Creatinine >= 2*baseline and > ULN	we will include a new criteria for this value. Please confirm the value is correct	correct	not included
0001	ADS9	SRC_VAR_5	Value not present in code list metadata but present in source data	TS-CV	this is a new set of population flag, please confirm if we need to add it	Please ignore	not included

Issues & updates post-release | Source data issues | specs_comp | Compare_tool | SDTM Source not mapped | Review cycles | Consistency checks | Def ...

Business & Decision Life Sciences

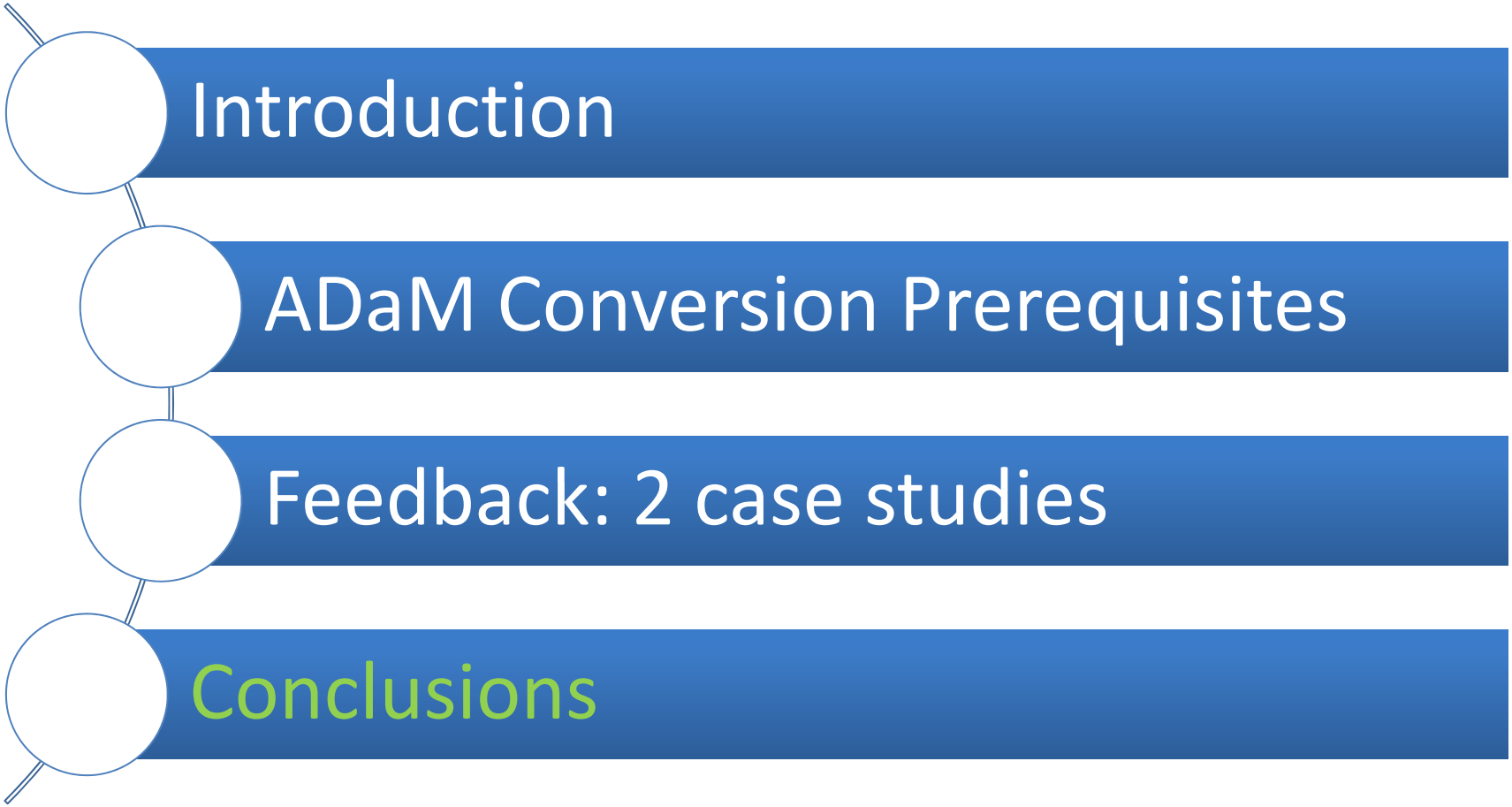
Case Study 2:
ADaM Conversion of a Pooled Dataset

- To ensure traceability with SDTM, a good knowledge of **what was mapped from RAW into SDTM** and in which study, was essential
- **List of RAW variables per study** per domain from each SDTM metadata, specifying if the variable was mapped into SDTM, and if it was empty

STUDY	VIEWNAME	VARNAME	LABEL	VARTYPE	FORMATN	MAPPED	CONFIRMED	EMPTY
RAW_Study1	AE	AEI	AE indicator	num	YN1F.		x	X
RAW_Study1	AE	AEONTM	AE onset time	num	TIME5.	X	x	
RAW_Study1	AE	AEENDC	AE end date continued	char	$CONT1F.	X	x	
RAW_Study1	AE	AEOUT	AE outcome	num	AEOUT1F.	X	x	
RAW_Study1	AE	AEREL	AE drug relationship	num	YN1F.	X	x	
RAW_Study1	AE	AEONDT	AE onset date	num	DATE9.	X	x	
RAW_Study1	AE	AEENDDT	AE end date	num	DATE9.	X	x	
RAW_Study1	AE	AELLT	AE lowest level term	char	$200.		x	
RAW_Study1	AE	AELLTCD	AE lowest level term code	char	$10.		x	
RAW_Study2	AE	AEI	AE indicator	num	YN1F.		x	X
RAW_Study2	AE	AEONTM	AE onset time	num	TIME5.	X	x	
RAW_Study2	AE	AEENDC	AE end date continued	char	$CONT1F.	X	x	
RAW_Study2	AE	AEOUT	AE outcome	num	AEOUT1F.	X	x	
RAW_Study2	AE	AEREL	AE drug relationship	num	YN1F.	X	x	
RAW_Study2	AE	AEONDT	AE onset date	num	DATE9.	X	x	
RAW_Study2	AE	AEENDDT	AE end date	num	DATE9.	X	x	
RAW_Study2	AE	AELLT	AE lowest level term	char	$200.		x	
RAW_Study2	AE	AELLTCD	AE lowest level term code	char	$10.		x	

Business & Decision Life Sciences

ADaM Conversions:
The Good, the Bad and the Ugly



Business & Decision Life Sciences

Conclusions

- **Parallel approach:**
  - ensured reproduction of TFLs
  - consistency between studies within a pooled database
- **Linear approach:**
  - SDTM/ADaM traceability is demonstrated
  - Bigger risk of not reproducing identical results to the initial analysis
- **Source pooled ADS** should be stable and **specifications** should be clear
- SDTM datasets & pooled ADaM datasets should be **finalized** before the individual study extractions can start

Business & Decision Life Sciences

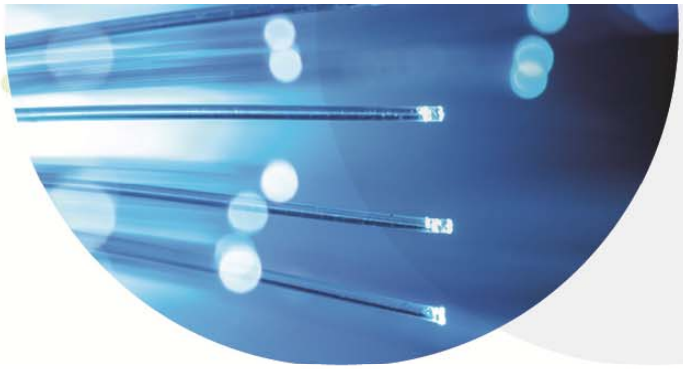
Conclusions

- Dictionary update: SDTM and ADaM uses the **same dictionary version**. Specific ADaM variables are available
- **Good communication** is needed between SDTM, ADaM team members and the sponsor, during the whole process
- **Sponsor's feedback** on draft mapping and draft ADaM datasets is crucial to address all details upfront
- All decisions can be easily tracked if they are properly documented in a **communication file**

Business & Decision Life Sciences

Questions?

Business & Decision Life Sciences



Business & Decision
Life Sciences

Jessica Minkue | Project Manager Statistics | jessica.minkue@businessdecision.com

Business & Decision Life Sciences
Sint-Lambertusstraat 141 rue Saint-Lambert
B-1200 Brussels
T: +32 2 774 11 00
F: +32 2 774 11 99
lifesciences@businessdecision.com
http://www.businessdecision-lifesciences.com/