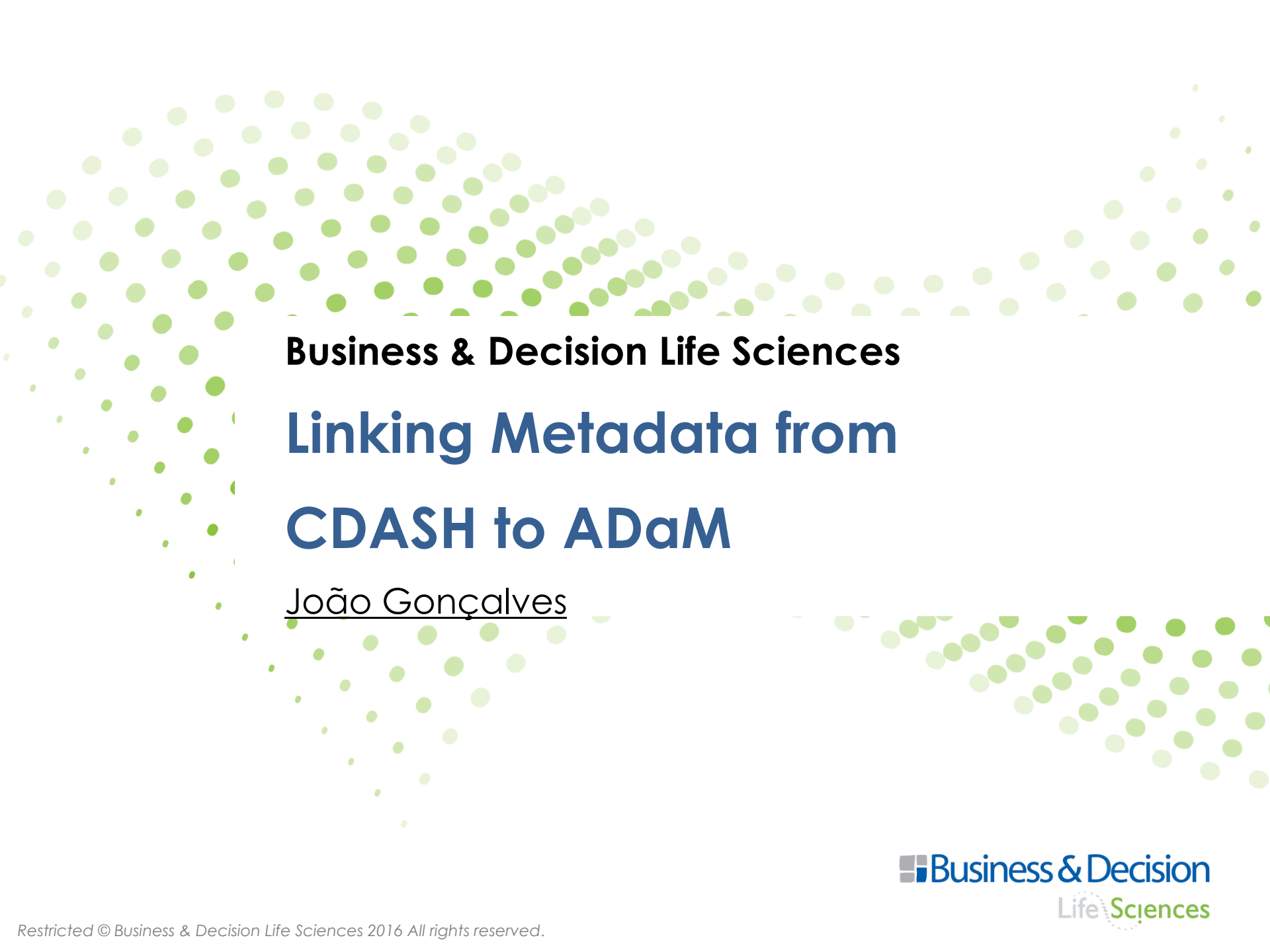


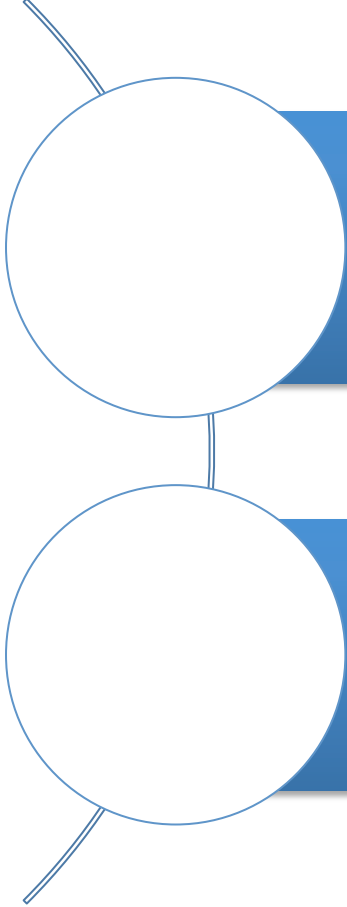
A decorative pattern of green dots of varying sizes is scattered across the slide, primarily concentrated in the upper left and lower right corners, creating a modern, abstract background.

Business & Decision Life Sciences

**Linking Metadata from
CDASH to ADaM**

João Gonçalves

Agenda

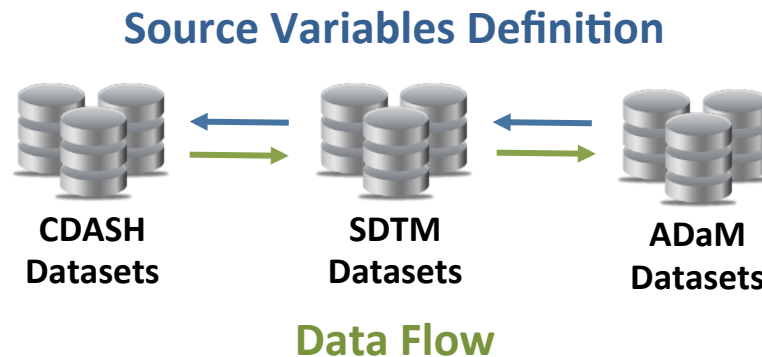


Missing Links

Links Enable Automation

Metadata Links

- **Content of standard variables often involves source variables**
 - SDTM is created based on CDASH, calculations and rules
 - ADaM is created based on copied SDTM and calculations
- **Metadata links are source variable metadata**



Missing Links

ADaM

Character description of
a grouping or pooling of
the subject's age for
analysis purposes.

AGEGRy

DM.AGE AGE

DM.AGEU AGEU

Missing Links

SDTM

ADaM

Character description of
a grouping or pooling of
the subject's age for
analysis purposes. **AGEGRy**

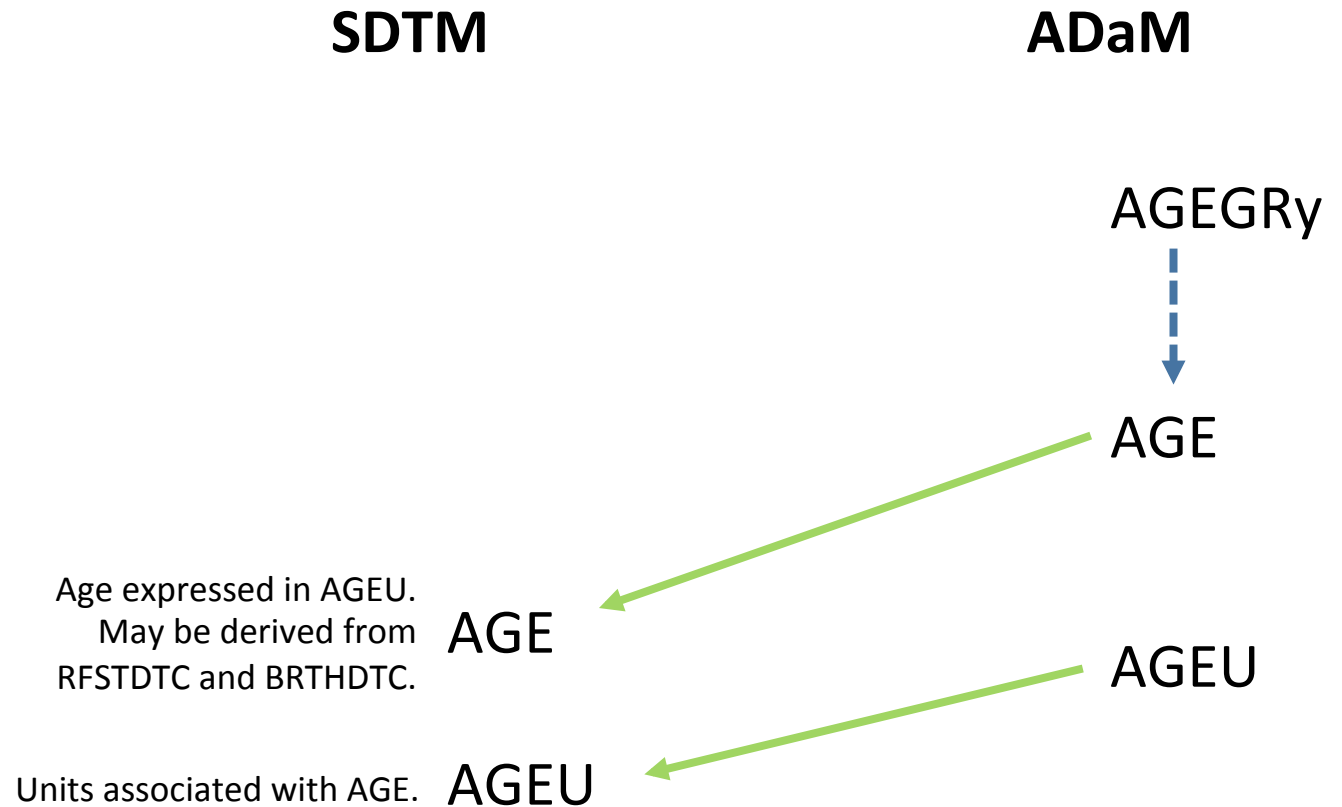
DM.AGE AGE

AGE

DM.AGEU AGEU

AGEU

Missing Links



Missing Links

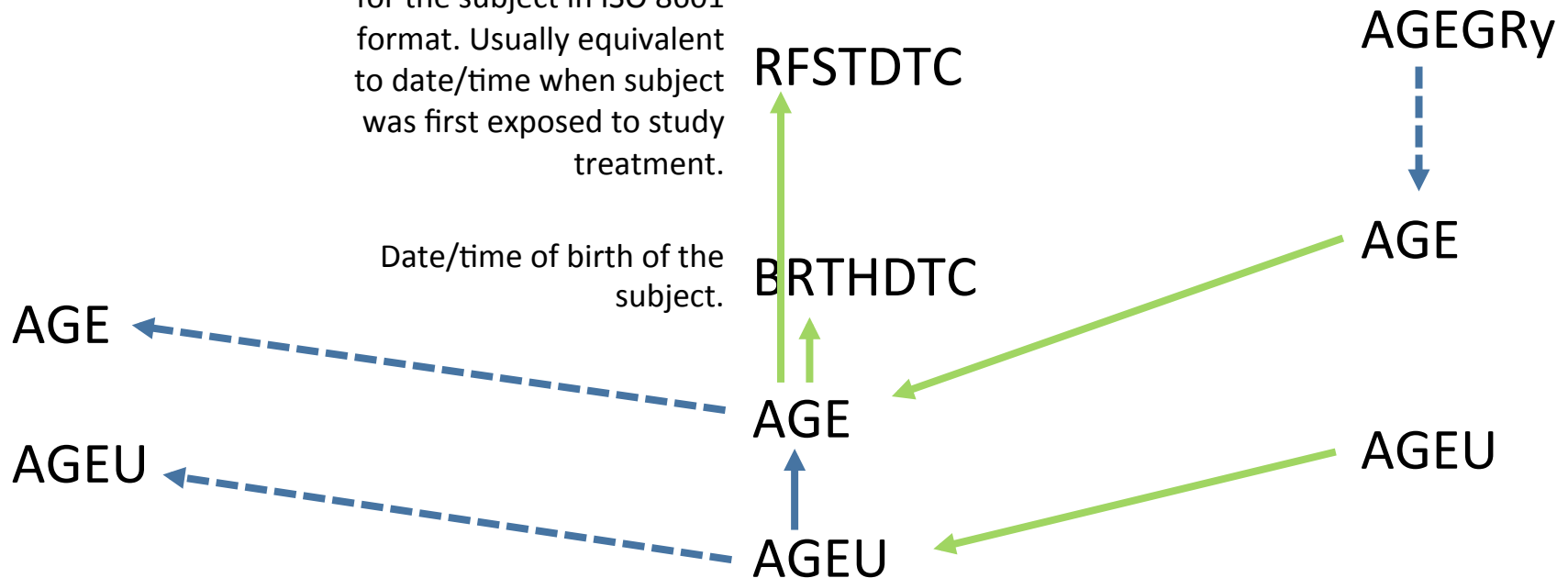
CDASH

SDTM

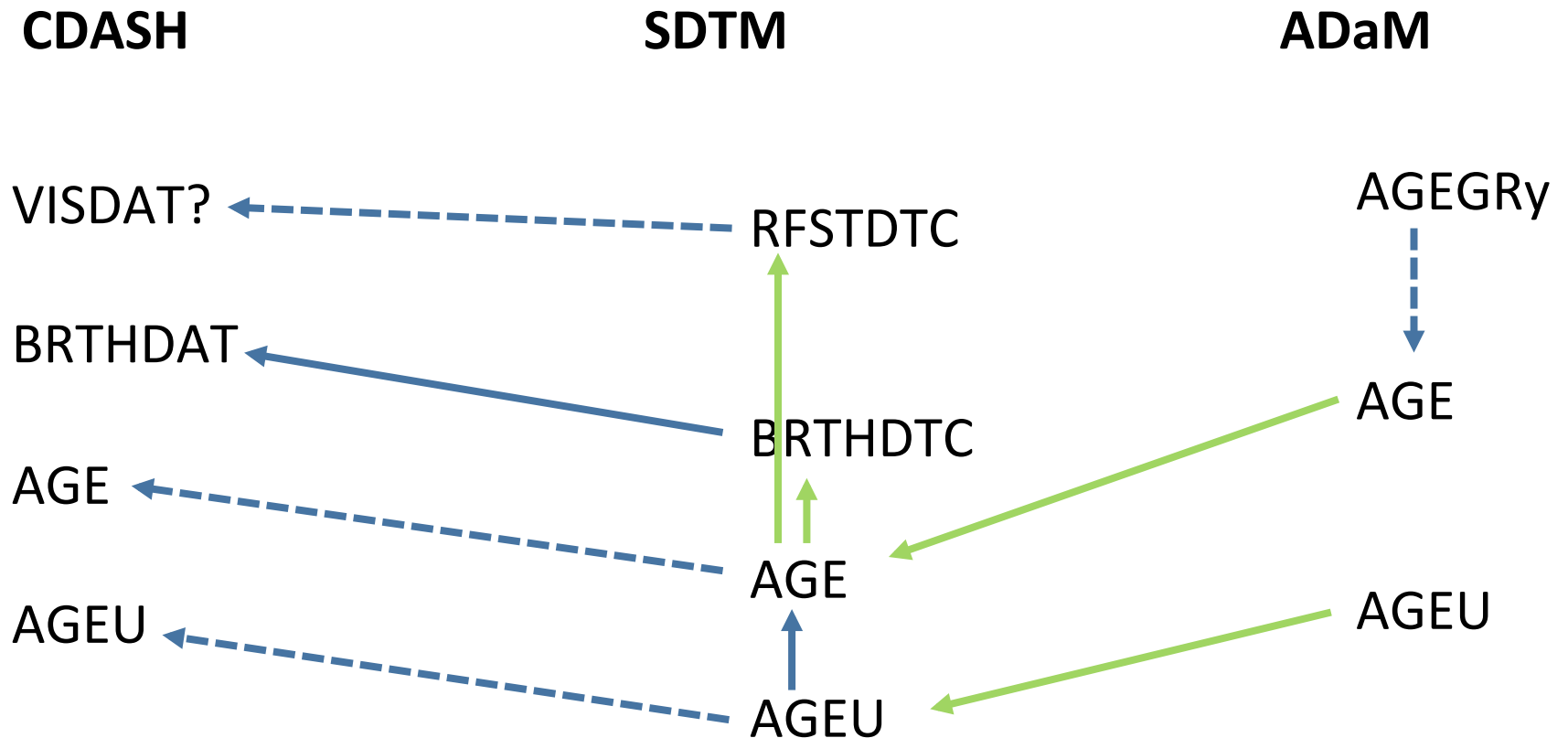
ADaM

Reference Start Date/time
for the subject in ISO 8601
format. Usually equivalent
to date/time when subject
was first exposed to study
treatment.

Date/time of birth of the
subject.



Missing Links



Missing Links

CDASH

VISDAT?

BRTHDAT

AGE

AGEU

SDTM

Reference Start Date/time for the subject in ISO 8601 format. Usually equivalent to date/time when subject was first exposed to study treatment.

Date/time of birth of the subject.

Age expressed in AGEU. May be derived from RFSTDTC and BRTHDTC.

Units associated with AGE.

RFSTDTC

BRTHDTC

AGE

AGEU

ADaM

Character description of a grouping or pooling of the subject's age for analysis purposes.

AGEGRy

DM.AGE AGE

DM.AGEU AGEU

Metadata Links In CDISC Implementation Guides

■ Source variables are specified in the CDISC Notes

- **Textual description** of the calculations or source variable to be copied, or other business rules

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
AGE	Age	Num		Req	DM.AGE. If analysis needs require a derived age that does not match DM.AGE, then AAGE must be added
AGEU	Age Units	Char	(AGEU)	Req	DM.AGEU
AGEGRy	Pooled Age Group y	Char		Perm	Character description of a grouping or pooling of the subject's age for analysis purposes. For example, AGEGR1 might have values of "<18", "18-65", and ">65"; AGEGR2 might have values of "Less than 35 y old" and "At least 35 y old".

Metadata Links In Define-XML

■ Define-XML format

- **Machine-readable** metadata description
- Used for submissions but also to **describe standards**
 - eSHARE Downloads page
 - SHARE API

■ Define-XML predecessor links

- Defined as a **textual description** of the variable

■ Define-XML computational method links

- Defined as a **method** with a **textual description**, **linked** to the variable reference

■ Links are not machine-readable in Define-XML!

Define-XML Predecessor Link

```
<ItemDef OID="IT.ADSL.AGEU" Name="AGEU" SASFieldName="AGEU"
DataType="text" Length="5">
  <Description>
    <TranslatedText xml:lang="en">Age Units</TranslatedText>
  </Description>
  <CodeListRef CodeListOID="CL.AGEU"/>
  <def:Origin Type="Predecessor">
    <Description>
      <TranslatedText xml:lang="en">DM.AGEU</TranslatedText>
    </Description>
  </def:Origin>
</ItemDef>
```

Name of the predecessor variable as text

Define-XML Computation Method Link

```
<ItemRef ItemOID="IT.ADSL.AGEGR1" OrderNumber="17" Mandatory="No"  
MethodOID="MT.ADSL.AGEGR1" />
```

OID of the computational method defined

Definition of the computational method (with an OID)

```
<MethodDef OID="MT.ADSL.AGEGR1" Name="MT.ADSL.AGEGR1"  
Type="Computation">
```

```
<Description>
```

```
<TranslatedText xml:lang="en">AGEGR1 = "&lt;18" if AGE &lt;18.  
AGEGR1 = "&gt;=18-&lt;25 " if AGE 18-24. AGEGR1 = "&gt;=25-&lt;65 " if  
AGE 25-64. AGEGR1 = "&gt;65" if AGE &gt;65.</TranslatedText>
```

```
</Description>
```

```
</MethodDef>
```

Pseudo code / description of the method as text

Mapping Specifications

■ Industry practice to represent source variable metadata

- Semi-structured format, usually in Excel, organization-dependent
- Allows basic machine-readability
 - Metadata traceability checks via highlighting formulae

DOMAIN	VARIABLE	LABEL	TYPE	LN	ROLE	CT	METHOD	ORIGIN	SOURCE DOMAIN	SOURCE VAR
ADSL	AGE	Age	num	8	Analysis			DM	DM	AGE
ADSL	AAGE	Analysis Age (Years)	num	8	Analysis		ADSL.AAGE	Derived	ADSL	AAGE
ADSL	AGEU	Age Units	char	5	Analysis	AGEU		DM	DM	AGEU
ADSL	AGEGR1	Age Group 1	char	20	Analysis	AGEGR1L	ADSL.AGEGR1	Derived	ADSL	AAGE
ADSL	AGEGR2	Age Group 2	char	20	Analysis	AGEGR2L	ADSL.AGEGR2	Derived	ADSL	AAGE

Missing Links in Define-XML

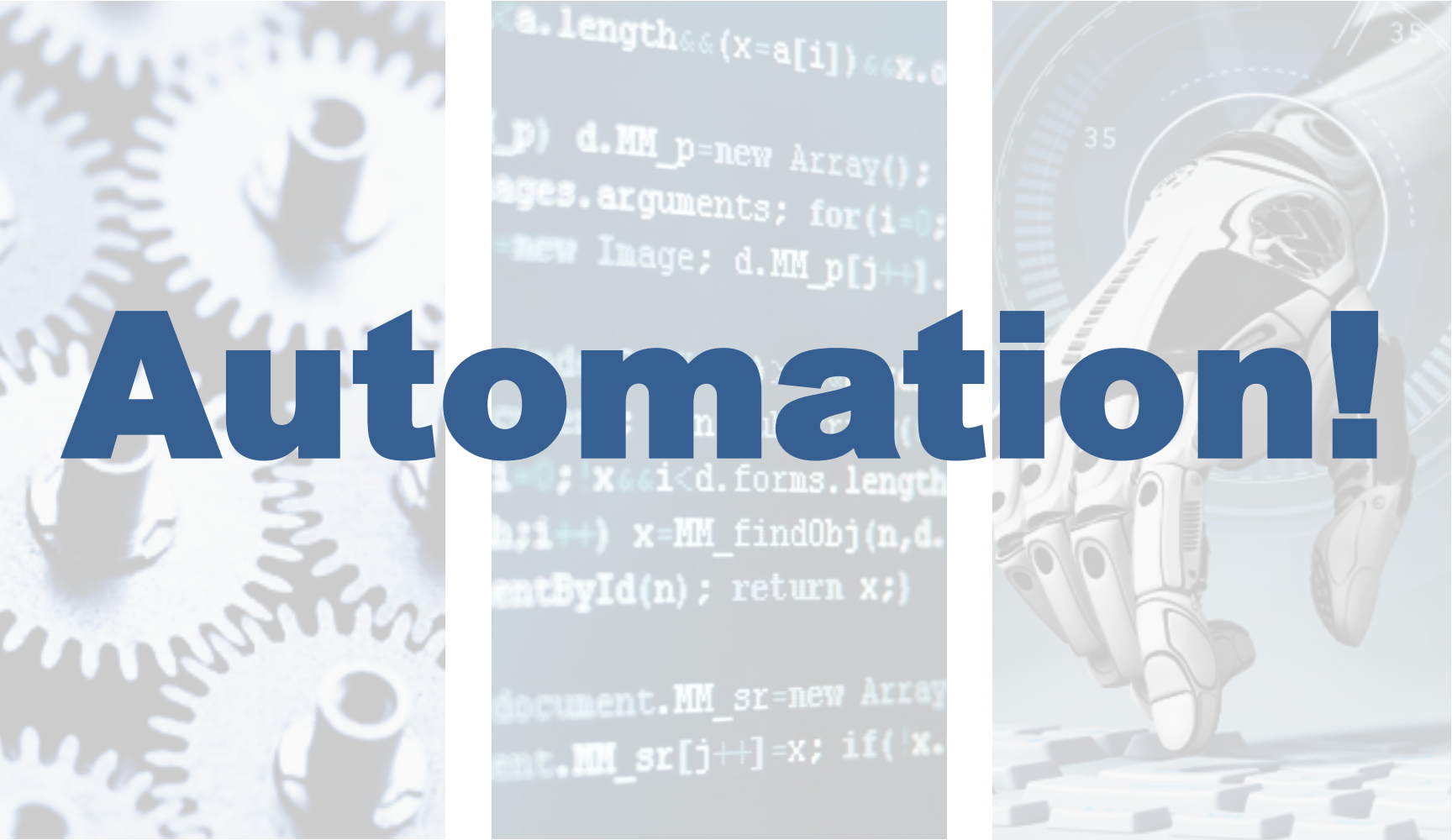
■ Links supported in Analysis Results definitions...

- Each **AnalysisResult** references a **ParameterOID**, relies on a number of **AnalysisDatasets**, which contain **AnalysisVariables** and **WhereClauseRefs** necessary to calculate it

■ Could be similarly supported for ADaM to SDTM and SDTM to CDASH...

```
<ItemDef OID="IT.ADSL.AGEGR1" Name="AGEGR1" DataType="text" Length="20">
<CodeListRef CodeListOID="CL.AGEGR1"/>
  <def:Origin Type="Derived"/>
  <ext:SourceDatasets>
    <def:DocumentRef leafID="LF.SDTM-Define"/>
    <ext:SourceDataset ItemGroupOID="IG.DM" >
      <ext:AnalysisVariable ItemOID="IT.DM.AGE"/>
    </ext:SourceDataset>
  </ext:SourceDatasets>
</ItemDef>
```

Why machine-readable metadata links?

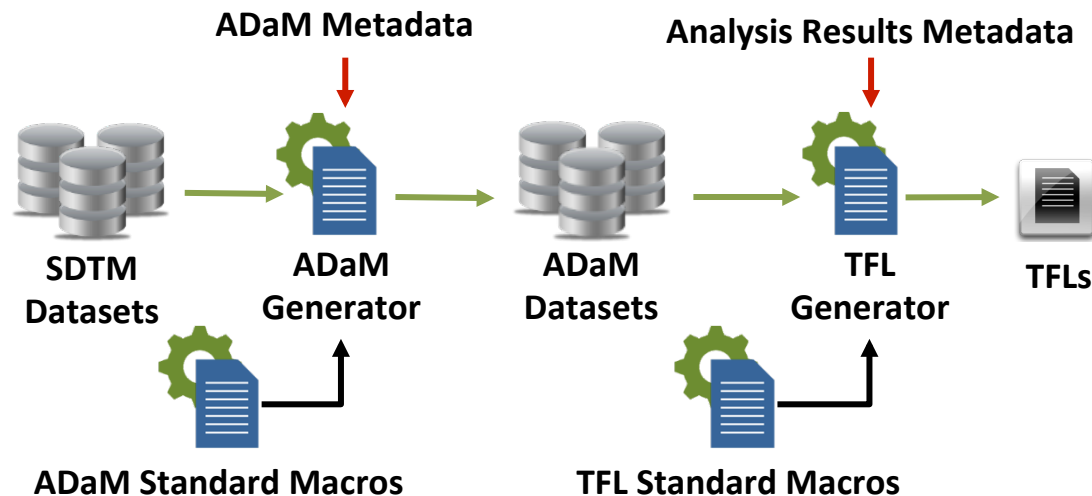
The background of the slide is a collage of three images. On the left is a close-up of interlocking metal gears. In the center is a snippet of JavaScript code. On the right is a robotic hand with a digital overlay showing the number '35'.

Automation!

Auto-generation Proof of Concept – Step #1

■ Generate ADaM datasets and TFLs based on study metadata and standard Macros

- Selected 3 common tables, which require 3 ADaM datasets (ADSL, ADVS, ADAE)
- Goal: decouple study-specific metadata and code



ADaM Standard Macros

■ Re-usable across studies and within the same study

- Study-specific configuration moved to macro parameters

```
%MSDTM_AGEGRy(  
    AGE_GROUP_NUMBER=1,  
    AGE_GROUP='<18 years;>=18-<25 years;>=25-<65 years;',  
    AGE_GROUP_LIST='<18;<25;<65;>65',  
    AGE_GROUP_OTHER='>=65 years'  
);
```

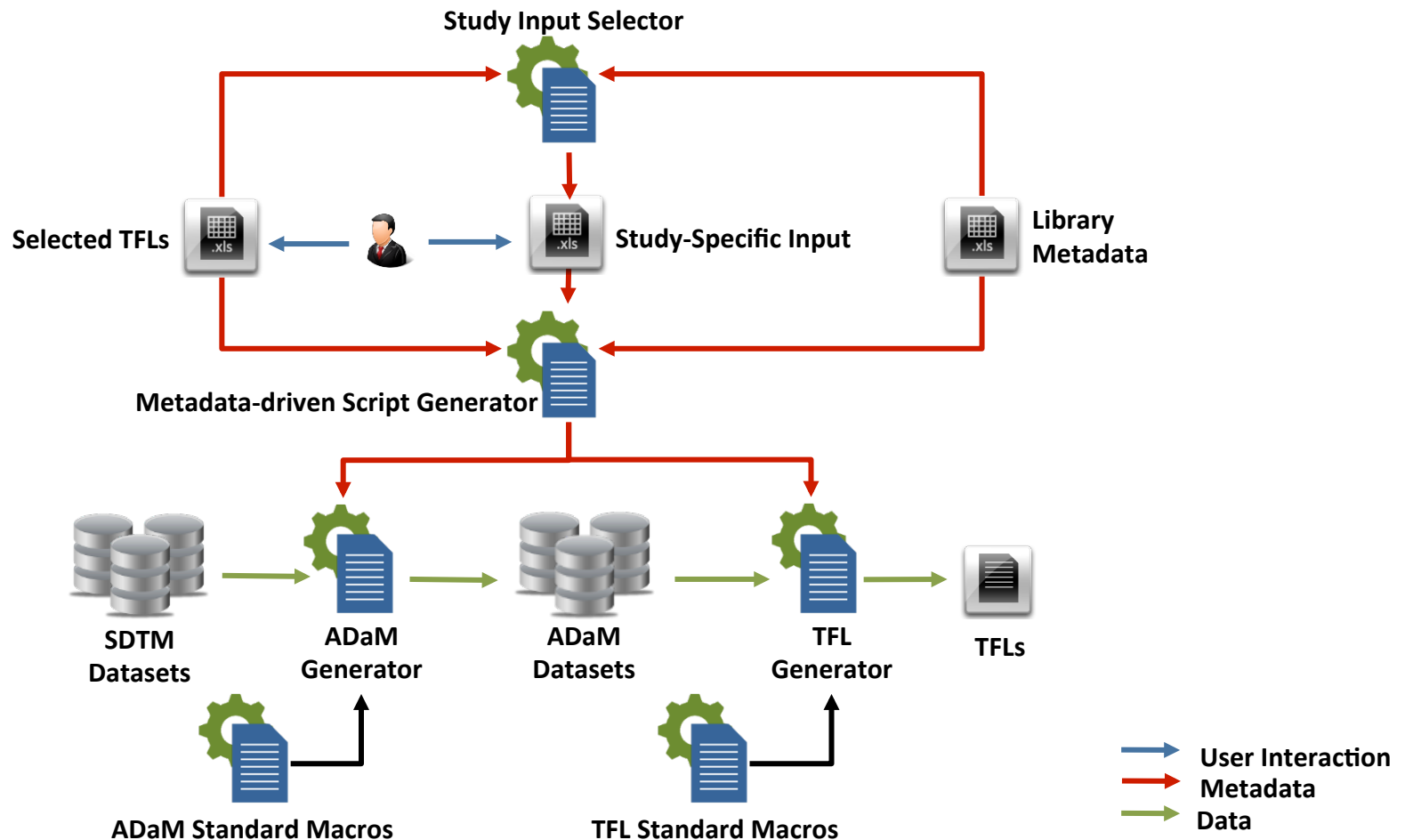
```
%MSDTM_AGEGRy(  
    AGE_GROUP_NUMBER=2,  
    AGE_GROUP='<41;>= 41 and < 65;',  
    AGE_GROUP_LIST='<41;<65;',  
    AGE_GROUP_OTHER='>= 65'  
);
```

TFL Standard Macros

■ More complex but same principle for TFLs

```
%TABLE_DESC_STATS(  
    STUDY='XX YY mg (STUDY_ID_001)',  
    NAME='Table 14.1.5.1',  
    DESCRIPTION='Demographics - Intention-To-Treat population',  
  
    VARIABLES='ADSL.AGE;ADSL.AGEGR2;ADSL.SEX;ADSL.RACE;ADSL.HEIGHT;ADSL.WEI  
GHT;ADSL.BMI;',  
    GROUPS='ADSL.ACTARMCD=A;ADSL.ACTARMCD=F;',  
    GROUP_LABEL_VARIABLE='ADSL.ACTARM',  
    WHERE='ADSL.ITTFL=Y',  
    VARIABLE_FOOTNOTES='ADSL.WEIGHT;Weight at baseline (Week 1). If  
weight is missing at baseline, weight at screening is  
used;ADSL.BMI;Body Mass Index (BMI) (kg/m2) = Weight (kg) / [Height  
(m)]2;',  
    FOOTNOTES='N = Number of subjects in specific group. n = Number  
of subjects. Calculation of percentages based on N.' ■ Business & Decision  
);
```

Auto-generation Proof of Concept – Step #2



Library-level Computational Methods

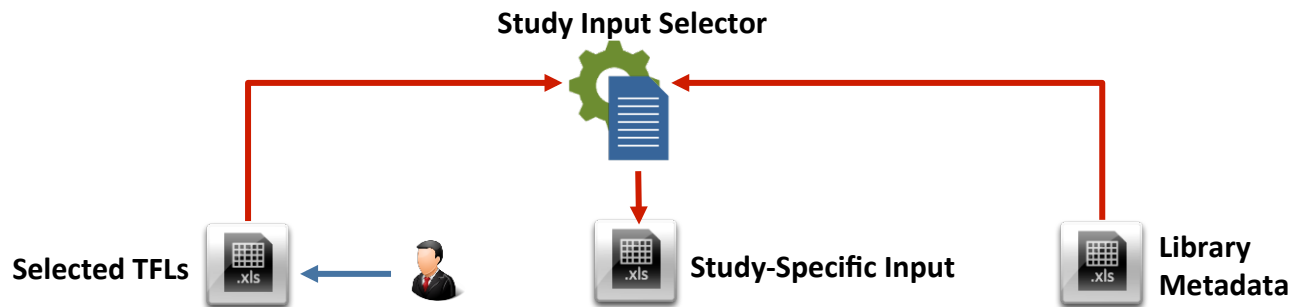
■ Computational Methods that abstract study-specific input

- A computational method is implemented by a Standard Macro
- Macro generalization exercise identified study-specific input

Reference	Description	Pseudo Code	Standard Macro	Study-Specific Input
ADSL.AAGE	Equals to ADSL.AGE	Equals to ADSL.AGE	MSDTM_AAGE	
ADSL.AGEGry	Creation of the age groups. A subsequent number will be used as a suffix to distinguish between the different groupings.	Equals to %AGE_GROUP[n] when ADSL.AAGE < %AGE_GROUP_LIST[n], for the lowest n possible; Equals to %AGE_GROUP_OTHER if ADSL.AAGE > %AGE_GROUP_LIST[n], for the highest n possible	MSDTM_AGEGry	%AGE_GROUP_NUMBER - value for the y suffix %AGE_GROUP - list or array containing the values to be assigned to the AGEGRy variable, except the one for the highest age group %AGE_GROUP_LIST - list or array containing upper age limits in the same order as %AGE_GROUP %AGE_GROUP_OTHER - the value to be assigned to the AGEGRy variable for the highest age group

Study Specification – Select TFLs

- TFLs required for the study are selected by the user



- The Study Input Selector script uses Library Metadata to determine required study-specific input

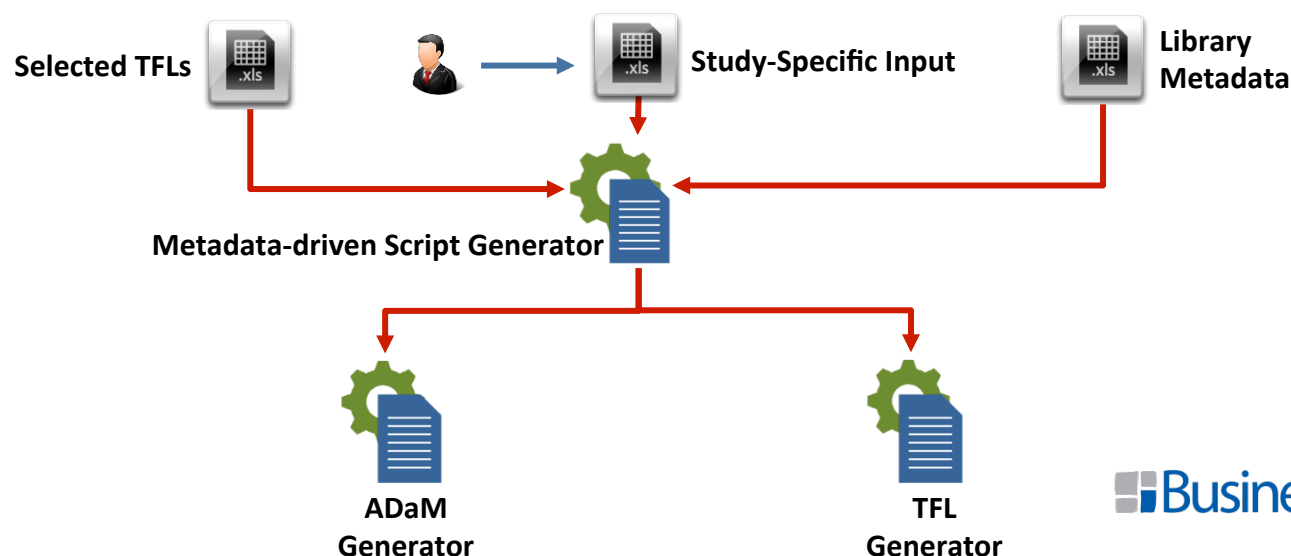
Reference	Param1	Param2	Param3	Param4
ADSL.AGEGRy	%AGE_GROUP_NUMBER - value for the y suffix	%AGE_GROUP - list or array containing the values to be assigned to the AGEGRy variable, except the one for the highest age group	%AGE_GROUP_LIST - list or array containing upper age limits in the same order as %AGE_GROUP	%AGE_GROUP_OTHER - the value to be assigned to the AGEGRy variable for the highest age group

Study Specification – Study-Specific Input

- The user inputs the required (and possibly optional) study-specific input

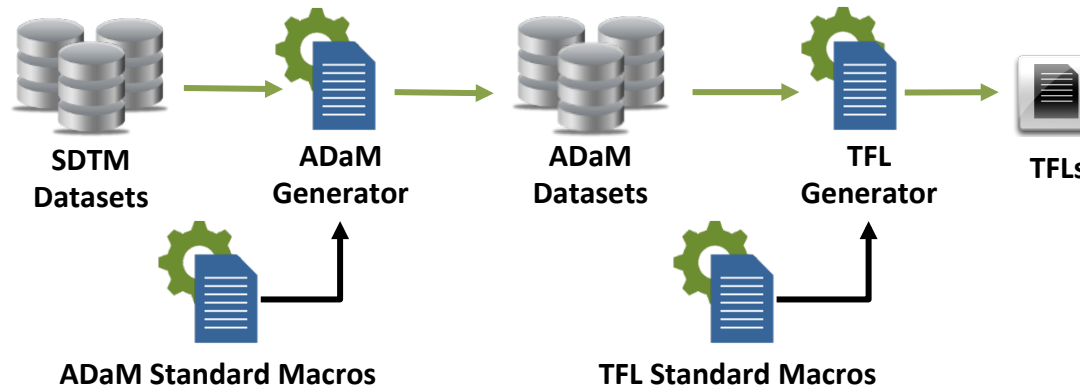
Reference	Param1	Param2	Param3	Param4
ADSL.AGEGRy	1	<18 years;>=18-<25 years;>=25-<65 years;'	'<18;<25;<65;'	'>=65 years'
ADSL.AGEGRy	2	< 41;>= 41 and < 65;'	'<41;<65;'	'>= 65'

- The Metadata-driven Script Generator uses all the metadata to generate the SAS scripts that will process the data



Auto-generation Execution

- The SAS Scripts can then be run against the SDTM data ...



- ... generating ADaM data and TFLs

XX YY mg (STUDY_ID_001)

Table 14.1.5.1: Demographics - Intention-To-Treat population

		STATISTIC	XX	PLACEBO	Total
Age		MAX	77.0	85.0	85.0
		MEAN	57.8	59.5	58.6
		MEDIAN	58.5	61.0	61.0
		MIN	38.0	27.0	27.0
		N	58.0	57.0	0115
		STD	10.6	11.5	11.1
Age Group 1	>=25-<65 years	n (%)	43 (74.1%)	35 (61.4%)	78 (67.8%)
	>=65 years	n (%)	15 (25.9%)	22 (38.6%)	37 (32.2%)
Sex	F	n (%)	32 (55.2%)	35 (61.4%)	67 (58.3%)
	M	n (%)	26 (44.8%)	22 (38.6%)	48 (41.7%)
Height at Baseline (cm)		MAX	0183	0188	0188
		MEAN	0166	0166	0166
		MEDIAN	0166	0164	0165
		MIN	0152	0143	0143
		N	57.0	56.0	0113
		STD	09.0	10.4	09.7

Conclusion

- **End-to-end metadata linking enables the new age of automation for clinical trials**
 - Metadata-driven data processing and reporting
 - Reduced manual work (and errors)
 - Metadata-driven study specification
 - Specifying the required analyses renders the data that needs to be collected

- **Linking be easily implemented through Define-XML**
 - Already existing for Analysis Results Metadata
 - Could be extended to link CDASH to SDTM and SDTM to ADaM



Thank you