

---

# **Standardising The Standards: The Benefits of Consistency**

***Nathan James, Roche Products Ltd.***

# SDTMs – An End User's Perspective

- ✓ Improved submission of clinical data to Regulators
- ✓ Improved transmission of clinical data between CRO and Sponsor

**Is the implementation of the SDTM model *always* the same?**

**CLAIM:**

The interpretation and implementation of the model can vary between companies

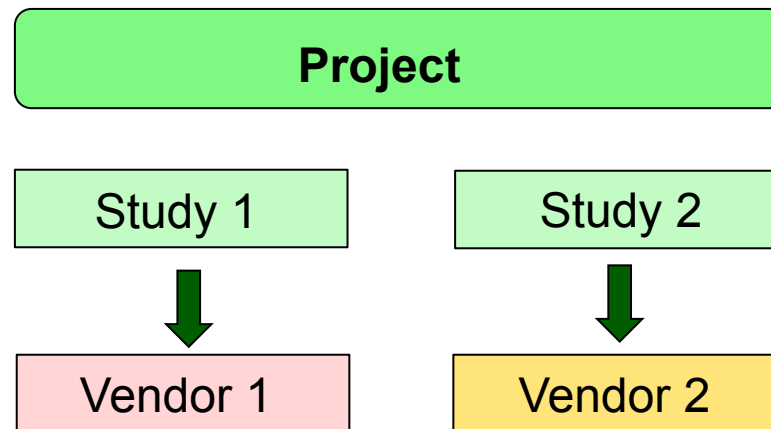
# A Simple Analogy

**“Bring me a standard red apple!”**



# Background

- SDTM build was **outsourced** for the project
- 2 studies had the SDTM construction with **2 different vendors**



- Future **pooling** of studies likely
- **Assumed** standard framework and structure, regardless of who was constructing the SDTMs.

# The Reality

- SDTMs **differed** in a variety of ways
- 2 categories:
  - Expected differences
  - Unexpected differences

# Expected Differences

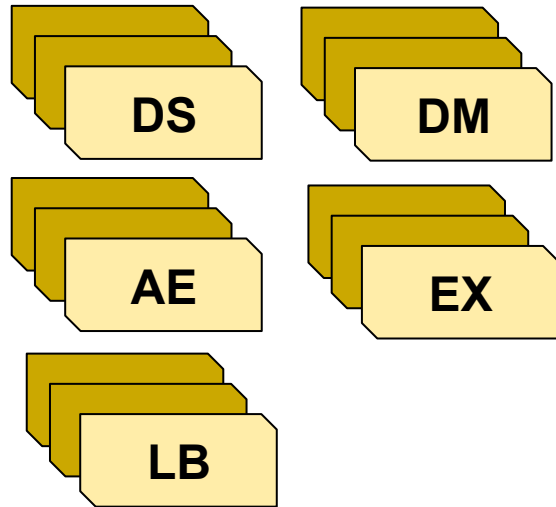
- There will **always** be expected differences
- Based on the similarity of the **collected data**
- **Safety domains** more comparable than efficacy domains

# Expected Differences

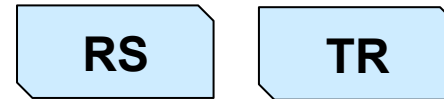
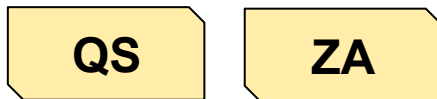
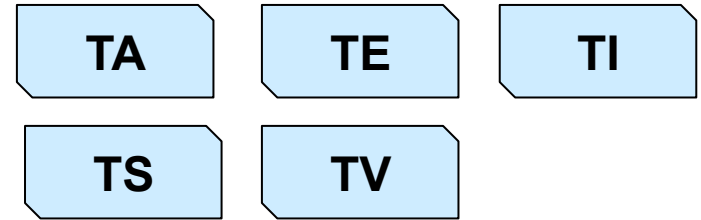
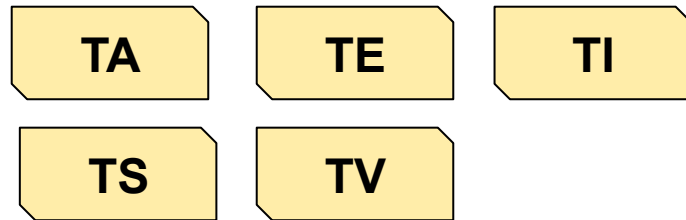
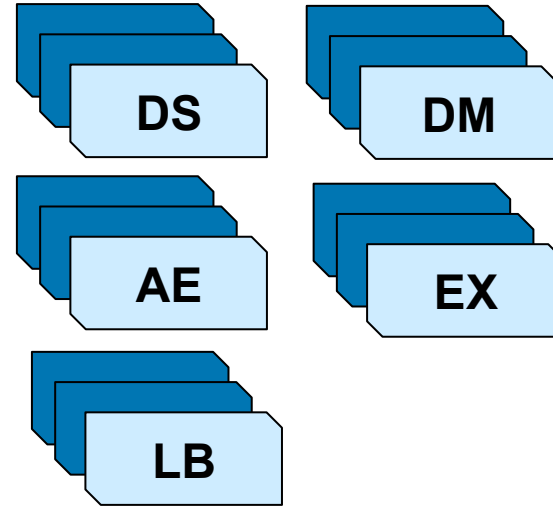
1. Domains delivered
2. Study specific elements
3. Trial Design Domains
4. SDTM Versions

# Expected Differences

## STUDY 1



## STUDY 2





# Expected Differences

1. Domains delivered
2. Study specific elements
3. Trial Design Domains
4. SDTM Versions

# Expected Differences

1. Domains delivered
2. Study specific elements
3. Trial Design Domains
4. SDTM Versions

# Expected Differences

1. Domains delivered
2. Study specific elements
3. Trial Design Domains
4. SDTM Versions

# Unexpected Differences

1. Subjective interpretations of the model
2. Mapping decisions

# Unexpected Differences

- Related to **derivations** performed within the model
  - *Trial Design Domains*
  - *Exposure domain*
- Input from an **analysis perspective**
  - *Reference start date [DM.RFSTDTC]*
  - *Baseline flags [--BLFL]*

# Unexpected Differences

1. Subjective interpretations of the model

2. Mapping decisions

# Unexpected Differences

- Data that has **no obvious mapping** into the model
  - *Deaths collected as an event*





# Unexpected Differences

- Data that has **no obvious mapping** into the model
  - *Deaths collected as an event*
  - *Non-standard elements*
  
- No **right or wrong** way of mapping some data
  - *Different vendors take different approaches*

# Automated Checking Tool

- **Extent** of the differences was unknown
- **Automated tool** to check for discrepancies developed
- **More benefits** to the tool than first perceived

# Automated Checking Tool

## REQUIREMENTS

- ✓ List **all discrepancies** between 2 sets of SDTM data
- ✓ Focus on **consistency** of tabulation and not conformance
- ✓ Dynamic
- ✓ Simple to use

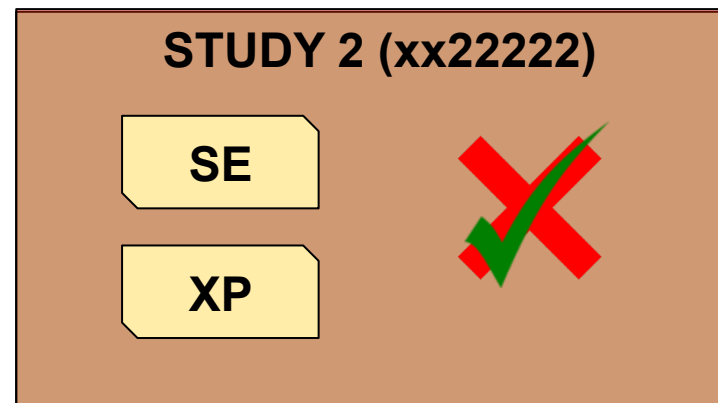
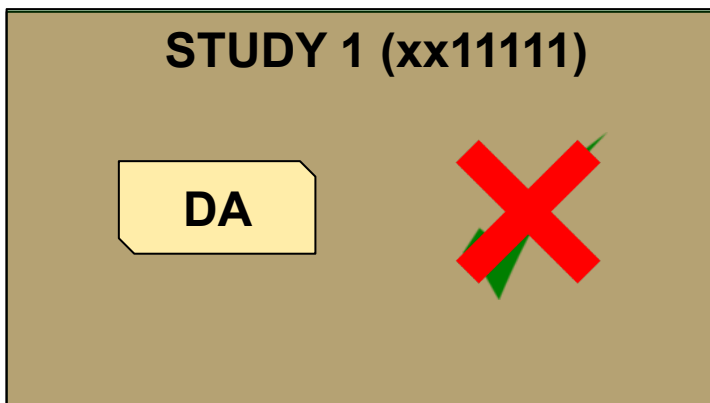
```
%sdtm_consistence_check( study1 = /opt/BIOSTAT/sdtm/cdpt9999/xx11111  
                          , study2 = /opt/BIOSTAT/sdtm/cdpt9999/xx22222  
                          );
```

# Automated Checking Tool

## DOMAIN CHECK

- Checks differences in the **existence of domains**

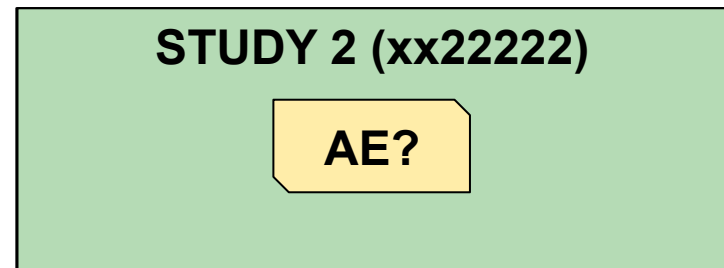
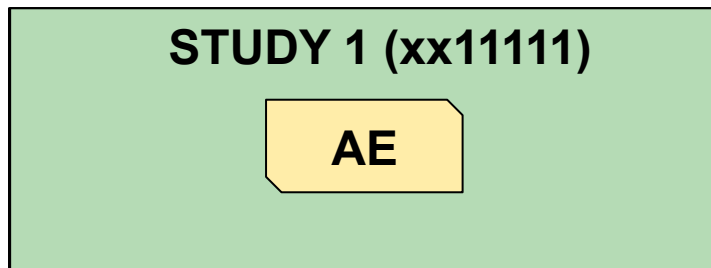
	A	B
1	Member Name	COMMENT
2	DA	Domain DA is in cdpt9999/xx11111 but not in cdpt9999/xx22222
3	SE	Domain SE is in cdpt9999/xx22222 but not in cdpt9999/xx11111
4	XP	Domain XP is in cdpt9999/xx22222 but not in cdpt9999/xx11111



# Automated Checking Tool

## VARIABLE CHECK

- Checks differences in the **variables** – **existence** and **attributes**
- Only **common domains** compared



AE.AEDECOD?

AE.AEDECOD



Label	Format
Length	Type
Informat	

Label	Format
Length	Type
Informat	



# Automated Checking Tool

	A	B	C	D	E
1	Member Name	Column Name	COMMENT	Column Type	Column Length
2	AE	AEACN1	Domain AE, variable AEACN1 is in cdpt9999/xx22222 but not in cdpt9999/xx11111	char	16
3	AE	AEAUTFD	Domain AE, variable AEAUTFD does not have the same length	char	200
4	CM	CMGRPID	Domain CM, variable CMGRPID is in cdpt9999/xx11111 but not in cdpt9999/xx22222		.
5	ZD	VISIT	Domain ZD, variable VISIT does not have the same informat	char	200
6	ZD	ZDSPID	Domain ZD, variable ZDSPID does not have the same label	char	200

	F	G	H	I	J	K	L	M
1	Column Label	Column Format	Column Informat	Column Type	Column Length	Column Label	Column Format	Column Informat
2	Action taken with XXXXX				.			
3	Autopsy Findings			char	100	Autopsy Findings		
4				char	40	Group ID		
5	Visit Name			char	200	Visit Name		\$16.
6	Sponsor ID			char	200	Sponsor-Defined Identifie		

# Automated Checking Tool

## VALUE CHECK

- Checks differences in the **controlled terminology**
- Values checked only on **common variables on common domains**
- List of variables **to be excluded** added to the macro

	A	B	C	D
1	MEMNAME	NAME	VALUE	COMMENT
2	AE	AEOCCRF	EVENT OCCURRED DURING INFUSION	Domain AE, variable AEOCCRF, value EVENT OCCURRED DURING INFUSION is in cdpt9999/xx22222 but not in cdpt9999/xx11111
3	AE	AEOCCRF	DURING INFUSION	Domain AE, variable AEOCCRF, value DURING INFUSION is in cdpt9999/xx11111 but not in cdpt9999/xx22222
4	DS	DSDECOD	RANDOMIZED	Domain DS, variable DSDECOD, value RANDOMIZED is in cdpt9999/xx22222 but not in cdpt9999/xx11111
5	DS	DSDECOD	RANDOMIZATION	Domain DS, variable DSDECOD, value RANDOMIZATION is in cdpt9999/xx11111 but not in cdpt9999/xx22222

# Automated Checking Tool

## VALUE CHECK

- Current tool runs better on **events** and **intervention** domains
  - *all observations considered for each variable*
- **Findings domains** are more problematic
  - *checks not performed on a **parameter level***
- Possible **future enhancement** of the tool.



# Tool Benefits

- ✓ Time/cost saving over manual checks
- ✓ Datasets more consistent → Improved pooling potential
- ✓ Run at start of a study → Pooling considered early
  - No impact on analysis programs
- ✓ Run of an individual study → monitor mapping/data changes
- ✓ Development of standard programs and maintenance → targeted

# Tool Benefits

**LB.LBMETHOD:**

**STUDY 1**  
PHOTOGRAPH

**STUDY 2**  
PHOTOGRAPH

## Standard/Project Program extract:

```
data example;
  set dataset1;
  if LBMETHOD = 'PHOTOGRAPH' then do;
    ...
  end;
run;
```

**LB.LBMETHOD:**

**STUDY 3**  
PHOTOGRAPH

*Known* new term in LBMETHOD → Search for programs containing LBMETHOD

# Tool Benefits

- ✓ Time/cost saving over manual checks
- ✓ Datasets more consistent → Improved pooling potential
- ✓ Run at start of a study → Pooling considered early
  - No impact on analysis programs
- ✓ Run of an individual study → monitor mapping/data changes
- ✓ Development of standard programs and maintenance → targeted
- ✓ Extends to Analysis Datasets or SDTMs constructed on different versions of the model

# Conclusion

- Plan for pooling activities early
- More consistent SDTMs → better pooling & standard programs
- Develop an automated tool
- Further benefits to this tool than just SDTM consistency
- *Why* do these differences exist?
- Never assume SDTMs will be structurally identical

***Doing now what patients need  
next***