
# Discover the Remarkable End-to-End Strategy behind Merck KGaA new Standard Library

Presented by Catherine Laugel, PhD, Senior Data Standards Expert, DMMI, Merck KGaA

# Meet the Speaker

## Catherine Laugel

Title: Senior Data Standards Expert

Organization: Merck KGaA

- At Merck: Reference for data collection Standards, involved in study set-up and co-leading a trans functional Governance group for the review and approval of Merck KGaA Standard Library.
- 13+ years' experience in end-to-end data management of clinical trials, including database design, data collection and quality control, in both CRO and pharma companies.
- Education : PhD in Organic Chemistry in Fundamental Research from Strasbourg (France) and 2 years post-doctoral position in Berlin (Germany).
- Hobbies : If time permits between work and my 2 kids, spending remaining energy in training for triathlon !

# Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

- *The author(s) have no real or apparent conflicts of interest to report.*

# Agenda

1. Overview of Merck Standard Team's new strategy to create CRF
2. Example of impact assessment on a Form
3. Detailed steps in our new process
4. Conclusion : advantages of this new process of creation of CRFs

# Overview of Merck Standard Team’s new strategy to create CRF

# Overview of Merck Standard Team's new strategy to create CRF

## Previously


Data Collection Expert designed CRF without any SDTM annotation.

SDTM Expert was involved at a later stage and many issues were discovered during SDTM mapping step, critical during submissions.

End-to-End strategy

# Overview of Merck Standard Team's new strategy to create CRF

## Previously

Data Collection Expert designed CRF without any SDTM annotation.

SDTM Expert was involved at a later stage and many issues were discovered during SDTM mapping step, critical during submissions.


End-to-End strategy

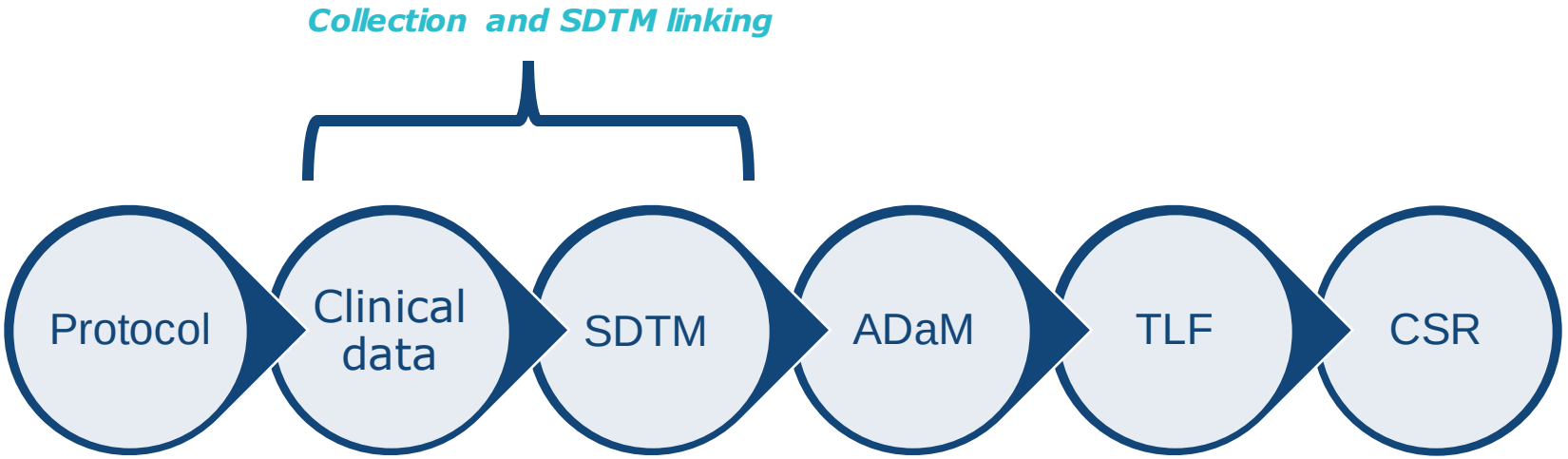
## Now !

Data Collection Expert and SDTM Expert are responsible & accountable for data & SDTM collection, **build in the same time**, leading simultaneously to annotated CRF & defined structures.

# Our Aim

**Previously** → **OUR AIM**

# Our Vision : End to End Process, connecting the elements together



# What we want to achieve with our new strategy of data collection?

Have an end to end strategy from protocol to downstream

Use CDASH for the data collection and SDTM if not available in CDASH

Standardization at Merck but also with other organizations

Reduce time to create data submission

## WHAT ?


Increase quality score (data fitness)

A better enhanced design to help the site in data entry to avoid data error

Speed up the study build (in the future)

Potential reduction of the efforts in programming, data cleaning and reconciliation

**Cost & time saving**

Reduce the time to DB lock eventually

# How we operate to make an improvement ?

Look at the Form in detail to see what can be improved

Impact assessment based on ongoing studies

Look at the industry practice

## HOW ?

Consider the medical requirement

Check the downstream impact for the analysis

Consider the sites side (more guidance, help with data entry, less data collection, less burden on site)


Use CDASH for the data collection or SDTM if not available in CDASH

QC checks on CDASH/SDTM – work on QC automation

Consider the Sponsor side (reconciliation, programming, cleaning)

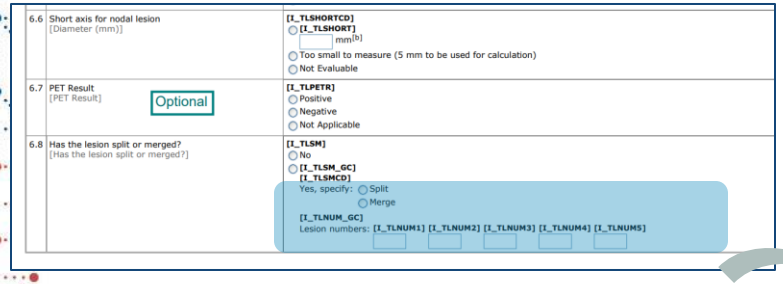
Leveraging in the near future the metadata repository for the forms build

Present the proposal to the study team

Post study implementation upgrade to standards post CCSG approval

# Example of impact assessment on a Form

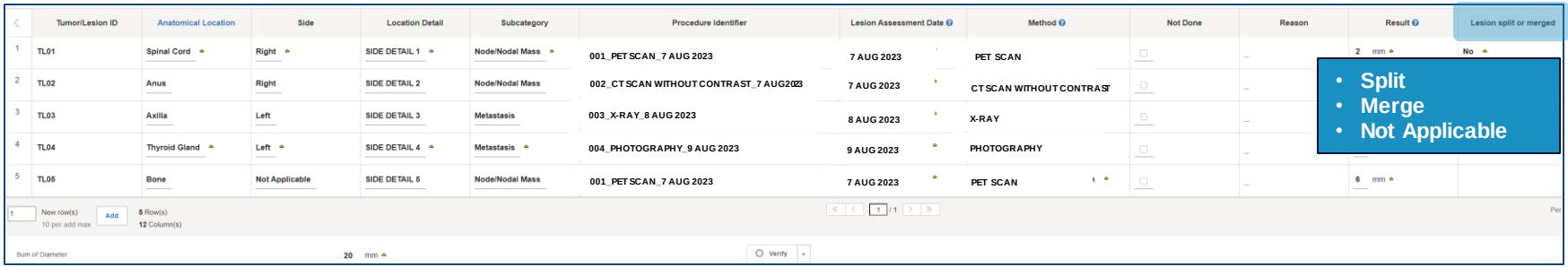
# Forms Updated - Split & Merged lesions details removed




- Split detailed information (Lesion numbers) removed from CRF
- Instructions (eCCGs) to be provided to sites for the reporting of the sum of split lesions in the initial lesion

**How to make an improvement?**

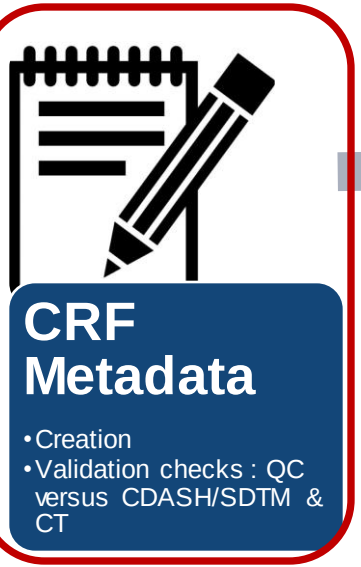
- **Impact analysis** - split reporting represents only **0.029%** of the total number of lesions
- **Look at industry standard** - Not collected in industry standards good practices (e.g. EORTC, other Sponsors…). **EORTC guidance doesn't describe that merged/split lesions need to be followed up** separately with more details, details not present in EORTC CRF.
- **No impact in analysis**
- **Dynamism can be used**
  - ✓ More convenient for sites (resolving major issues that were flagged by KAIZEN initiative)
  - ✓ Faster cleaning
  - ✓ Remove reconciliation effort between Imaging and Tumor Assessments



# Detailed steps in our new process

# Detailed steps in our new process

**For today's presentation**

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT

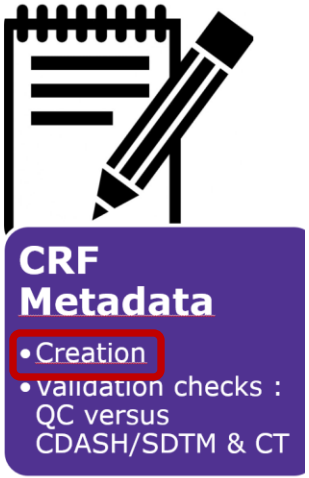

**SDTM Metadata**

- Creation
- Validation checks : QC versus SDTM & CT


**P21 Checks**

- Run P21E or P21C (Individually)
- Combine Metadata
- Run P21E (Combined) at Global Library Level

# Detailed steps in our new process – CRF Metadata




❑ **Creation of CRF Metadata**

✓ **Form Level Naming Convention**

<FormOID> (CRF data) will be: <DOMAIN NAME> <G or M or S> <2-digits serial number>

Form Title = <FormOID>_<Form Name>

DOMAIN NAME = SDTM domain

'**G'** stands for Global. '**M'** stands for Modification to Standards. '**S'** stands for Study-Specific.

- Standard Forms without modifications will have the letter 'G'.
- Standard Forms with modifications will have the letter 'M'. However, if a standard Form gets modified after Go live, the letter 'G' remains the same, as FormOID are not modified post Go Live.
- New study-specific Forms will have the letter 'S'.

The 2-digits serial number indicates the # of unique forms within the same domain.

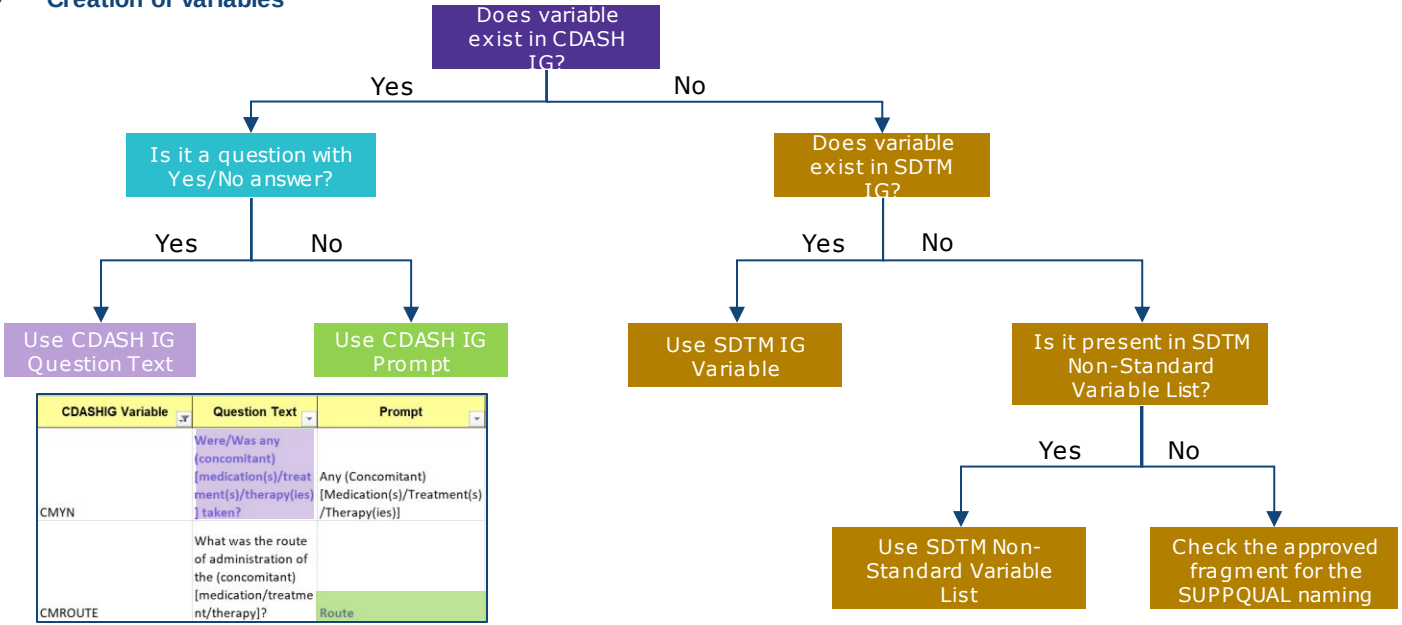
Example: **DMG01_Demographics.**

# Detailed steps in our new process - CRF Metadata

- ❑ **Creation of CRF Metadata**
  - ✓ **Creation of variables**

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT




| CDASHIG Variable | Question Text | Prompt |
|---|---|---|
| CMYN | Were/Was any (concomitant) [medication(s)/treatment(s)/therapy(ies)] taken? | Any (Concomitant) [Medication(s)/Treatment(s)/Therapy(ies)] |
| CMROUTE | What was the route of administration of the (concomitant) [medication/treatment/therapy]? | Route |

**Advantages of following CDASH/SDTM**
Harmonization - Generic Labels that can be used in any trial

# Detailed steps in our new process – CRF Metadata

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT

- **Creation of CRF Metadata**
  - ✓ **Creation of variables – Lab parameters (denormalized structure)**

- **Refer to SDTM Paired View File:**
  https://www.cdisc.org/system/files/members/standard/terminology/SDTM_paired_view_2024_03_29.xlsx
  CDISC LBTESTCD will be used to create FieldOID, and CDISC LBTEST will be used for its corresponding PreText.

| A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| Code | Coded Value Codelist Code | Coded Value Codelist Name | Coded Value for CDISC Submission LBTESTCD | Decoded Value Codelist Code | Decoded Value Codelist Name | Decoded Value for CDISC Submission LBTEST |
| C64431 | C65047 | Laboratory Test Code | ALB | C67154 | Laboratory Test Name | Albumin |
| C147293 | C65047 | Laboratory Test Code | ALBC | C67154 | Laboratory Test Name | Albumin Clearance |
| C74761 | C65047 | Laboratory Test Code | ALBCREAT | C67154 | Laboratory Test Name | Albumin/Creatinine |
| C150814 | C65047 | Laboratory Test Code | ALBEXR | C67154 | Laboratory Test Name | Albumin Excretion Rate |
| C158228 | C65047 | Laboratory Test Code | ALBGALB | C67154 | Laboratory Test Name | Glycated Albumin/Albumin |
| C74894 | C65047 | Laboratory Test Code | ALBGLOB | C67154 | Laboratory Test Name | Albumin/Globulin |

FieldOID (ALS)

PreText (ALS)

| FormOID | FieldOID | Field OID to be updated with LBTESTCD from CDISC | Ordinal | DraftFieldNumber | DraftFieldName | DraftFieldActive | VariableOID | PreText | PreText to be updated with LBTEST from CDISC |
|---|---|---|---|---|---|---|---|---|---|
| LBG01 | ALB_LBORRES | ALB | 6 | | ALB_LBORRES | TRUE | ALB_LBORRES | Albumin | Albumin |
| LBG01 | ALT_LBORRES | ALT | 7 | | ALT_LBORRES | TRUE | ALT_LBORRES | Alanine Aminotransferase (ALT) | Alanine Aminotransferase |

# Detailed steps in our new process – CRF Metadata

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT

## ❑ Creation of CRF Metadata

✓ **Naming conventions for codelists**

- **If codelist is present in CDISC**

The expectation is that the name of the codelists should include the **c-code** of the codelist in order to be able to link it back to CDISC.

**CL_<Code>_<CDISC Submission Value>**

Example: CL_C66767_ACN

| Cod | Codelist Code | Codelist Extensible (Yes/No) | Codelist Name | CDISC Submission Value | CDISC Synonym(s) |
|---|---|---|---|---|---|
| C66767 | | No | Action Taken with Study Treatment | ACN | Action Taken with Study Treatment |

- **CL_<Code>_<CDISC Submission Value>_Subset (if needed)**

Example: CL_C74457_RACE_WHITE

- **If codelist (CL) is not present in CDISC and thus is sponsor defined (SP)**

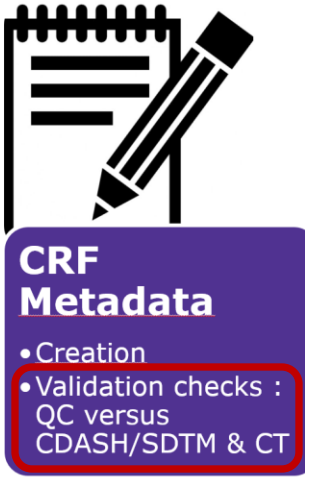
**CL_SP_<Name of the codelist>**

**Advantages of this naming convention**
- Easy to track that CDISC terminology is used
- Automatic check via programming that codelist is per CDISC terminology: gain of time

# Detailed steps in our new process – CRF Metadata

- **QC of CRF Metadata**

Each variable and codelist from ALS are QCed versus CDASH, SDTM and CT thanks to programmed QC checks. Outputs of metadata validation checks are reviewed and metadata are updated accordingly, if applicable.




| Check ID | Dataset | Message | Review Comment |
|---|---|---|---|
| QUES01 | AE | Collection variable "AEREL_PT" in Form "AEG01" does not exist in CDASH or SDTM metadata for the domain. Review and correct as needed or document the deviation. | This is okay, as requested |
| QUES01 | AE | Collection variable "AEREL_IMRT" in Form "AEG01" does not exist in CDASH or SDTM metadata for the domain. Review and correct as needed or document the deviation. | This is okay, as requested |
| QUES01 | AE | Collection variable "AEACN_PT" in Form "AEG01" does not exist in CDASH or SDTM metadata for the domain. Review and correct as needed or document the deviation. | This is okay, as requested |

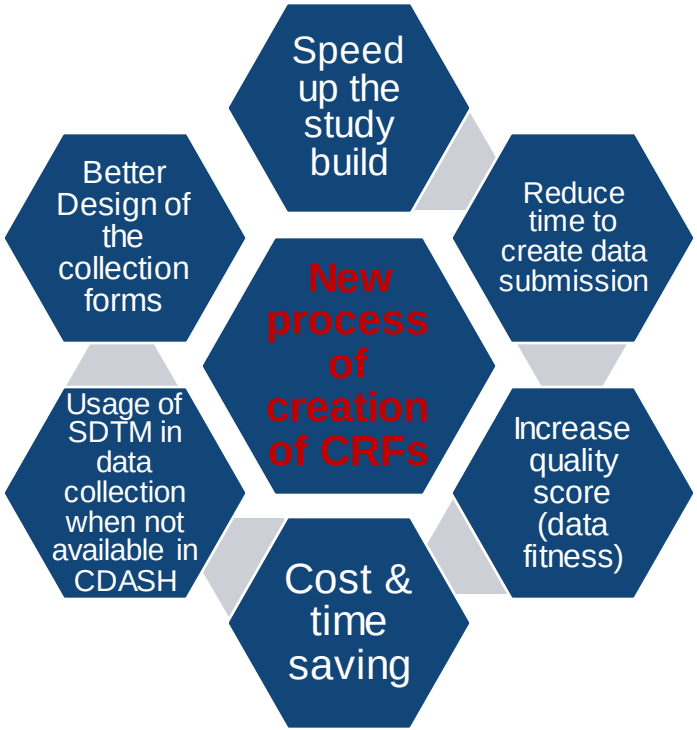
**QC validation checks**

| ALS Form Name | SDTM Domain | ALS Variable Order | ALS Variable | ALS PreText | ALS SASLabel | SDTM Label | ALS Codelist | CDASH Question Text | CDASH Prompt | CDASH Codelist | ALS CDASH Status | Status of CDASH Question Text Vs ALS PreText | Status of CDASH Prompt Vs ALS PreText | Status of CDASH Variable Label Vs ALS SASLabel | Status of SDTM Variable Label Vs ALS SASLabel PreText | Status of CDASH Codelist Vs ALS Codelist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEG01 | AE | 10 | AEACN | Action Taken with Study Treatment | Action Taken with Study Treatment | Action Taken with Study Treatment | C66767 | What action was taken with study treatment? | Action Taken with Study Treatment | C66767 | FIELDOID Present | No Match | Match | Match | Match | Match |

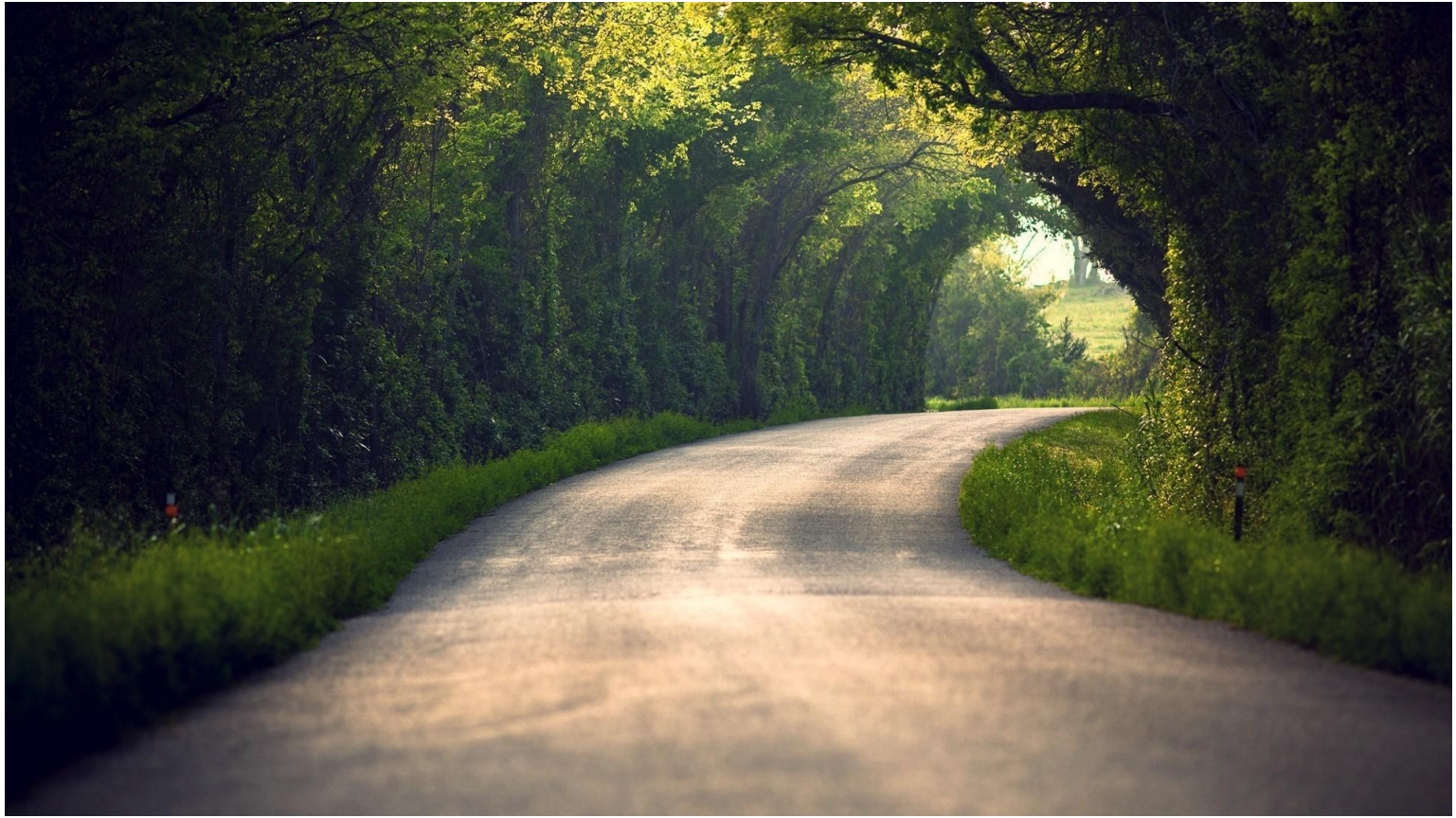
**CDASH compare**

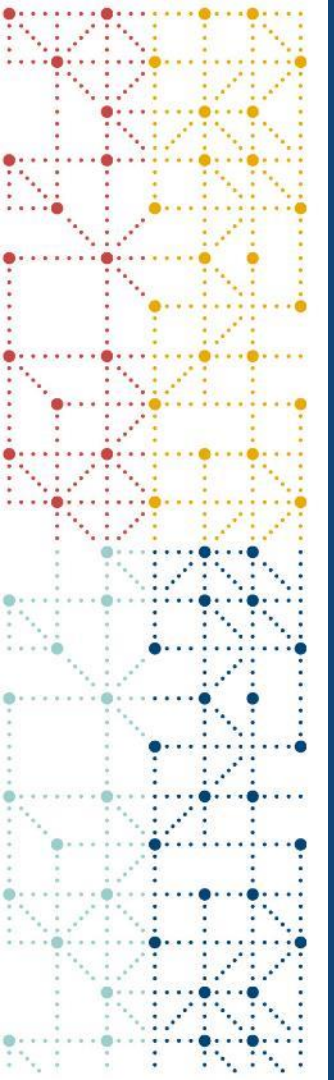
# Conclusion : Advantages of this new process of creation of CRFs

# Conclusion : Advantages of this new process of creation of CRFs



# Take Home Message





# Any question?

# Thank You!

# Back-up slides

# Detailed steps in our new process – CRF Metadata

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT

❑ **Creation of CRF Metadata**

✓ **Creation of variables at study level**

- **Strategy is to share same variables within different Forms to mimic SDTM**
  - If variables with same format (text) or same codelist, and same PreText : same FieldOID will be used.
- **If not possible, the below guidance is applied**
  - 2 variables PRTRT with different format (free text & codelist) or PreText, or same format with different codelists : use different FieldOIDs PRTRT and PRTRT_X

| FormOID | FieldOID | DataFormat | DataDictionaryName | ControlType | PreText |
|---|---|---|---|---|---|
| PRG03 | PRTRT_IMG | $40 | CL_C85492_METHOD_IMAG | RadioButton (Vertical) | Method |
| PRS02 | PRTRT_ENDO | $20 | CL_C85492_METHOD | RadioButton (Vertical) | Procedure Name |
| PRS03 | PRTRT_FEED | $20 | CL_SP_ROUTE_FEED | DropDownList | Procedure Name |
| PRG02 | PRTRT | $200 | | LongText | Procedure Name |
| PRG05 | PRTRT_RT | $40 | CL_SP_PRCAT | DropDownList | Procedure Name |
| PRG07 | PRTRT | $200 | | LongText | Procedure Name |

# Detailed steps in our new process – CRF Metadata

**CRF Metadata**

- Creation
- Validation checks : QC versus CDASH/SDTM & CT

❑ **Creation of CRF Metadata**

✓ **Creation of variables**

- Miscelleanous
  - Use of –DAT (CDASH) and –TIM (CDASH) instead of DTC (SDTM)
  - Ensure Format $200 if free text (ie : AE, CM, PR Forms)

✓ **Normalized structure vs denormalized structure**

Use of normalized structure (Log)

Use SDTM XXTEST and XXORRES in table reporting (i.e.: PEG01 Form)

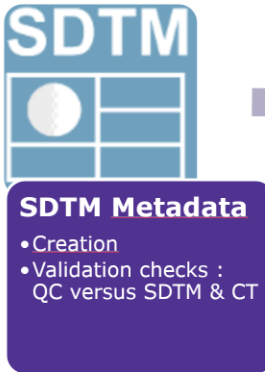
| Log | PETEST | PEORRES |
|---|---|---|
| Log 1 | Gynecological Examination | Normal/Abnormal |
| Log 2 | Dental | Normal/Abnormal |
| Log 3 | Ears, Nose, Throat | Normal/Abnormal |

Use a denormalized structure (non-Log)

Use directly the value of SDTM XXTESTCD as FieldOID (i.e.: VSG01 Form)

| | |
|---|---|
| SYSBP | 120 |
| SYSBPU | mmHg |
| DIABP | 80 |
| DIABPU | mmHg |

# Detailed steps in our new process – SDTM Metadata




## ❑ Creation of SDTM Metadata

CRF is annotated with SDTM and then SDTM Metadata file is created based on CRF Metadata. A file with data examples is also prepared.

| Orde | Dataset | Name | Label | Type | Length | Digits | Format | Core | Codelist | Decoded Variable | VLM | Origin | Method | Comment |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | IE | STUDYID | Study Identifier | text | 200 | | | Required | | | | Protocol | | Identifier |
| 2 | IE | DOMAIN | Domain Abbreviation | text | 2 | | | Required | CL_C66734_DOMAIN | | | Assigned | | Identifier |
| 3 | IE | USUBJID | Unique Subject Identifier | text | 200 | | | Expected | | | | Derived | USUBJID | Identifier |
| 4 | IE | IESEQ | Sequence Number | float | | | | Required | | | | Derived | SEQ | Identifier |
| 5 | IE | IETESTCD | Inclusion/Exclusion Criterion Short Name | text | 8 | | | Required | | IETEST | | CRF | | Topic |
| 6 | IE | IETEST | Inclusion/Exclusion Criterion | text | 200 | | | Required | | | | Assigned | | Synonym Qualifier |
| 7 | IE | IECAT | Inclusion/Exclusion Category | text | 200 | | | Required | CL_C66797_IECAT | | | CRF | | Grouping Qualifier |
| 8 | IE | IEORRES | I/E Criterion Original Result | text | 2 | | | Required | CL_C66742_NY | | | Assigned | | Result Qualifier |
| 9 | IE | IESTRESC | I/E Criterion Result in Std Format | text | 2 | | | Required | CL_C66742_NY | | Y | Assigned | | Result Qualifier |
| 10 | IE | VISITNUM | Visit Number | float | | | | Permissible | | VISIT | | Assigned | | Timing |
| 11 | IE | VISIT | Visit Name | text | 200 | | | Permissible | | | | Assigned | | Timing |
| 12 | IE | VISITDY | Planned Study Day of Visit | integer | | | | Permissible | | | | Protocol | | Timing |
| 13 | IE | EPOCH | Epoch | text | 200 | | | Permissible | CL_C99079_EPOCH | | | Derived | EPOCH | Timing |

## ❑ QC of SDTM Metadata

Each SDTM Metadata file is then QCed versus SDTM and CT thanks to programmed QC checks, similarly as for the previous check on CRF Metadata. What is present in Raw CRF Metadata needs to be present in SDTM Metadata : to facilitate mapping and annotation.

Outputs of metadata validation checks are reviewed and metadata are updated accordingly, if applicable.