# Receiving Outsourced SDTMs

## Automated checks for uploaded CRO SDTM data

15 March 2016 I German CDISC User Network

Wioleta Schulze

# Receiving Outsourced SDTMs
## Automated checks for uploaded CRO SDTM data

### Table of Contents

# Creation of the SDTM data
## Process at Grünenthal (GRT)

- Creation of the SDTM data is **outsourced to CRO**

**Upfront agreement between GRT and the CRO**

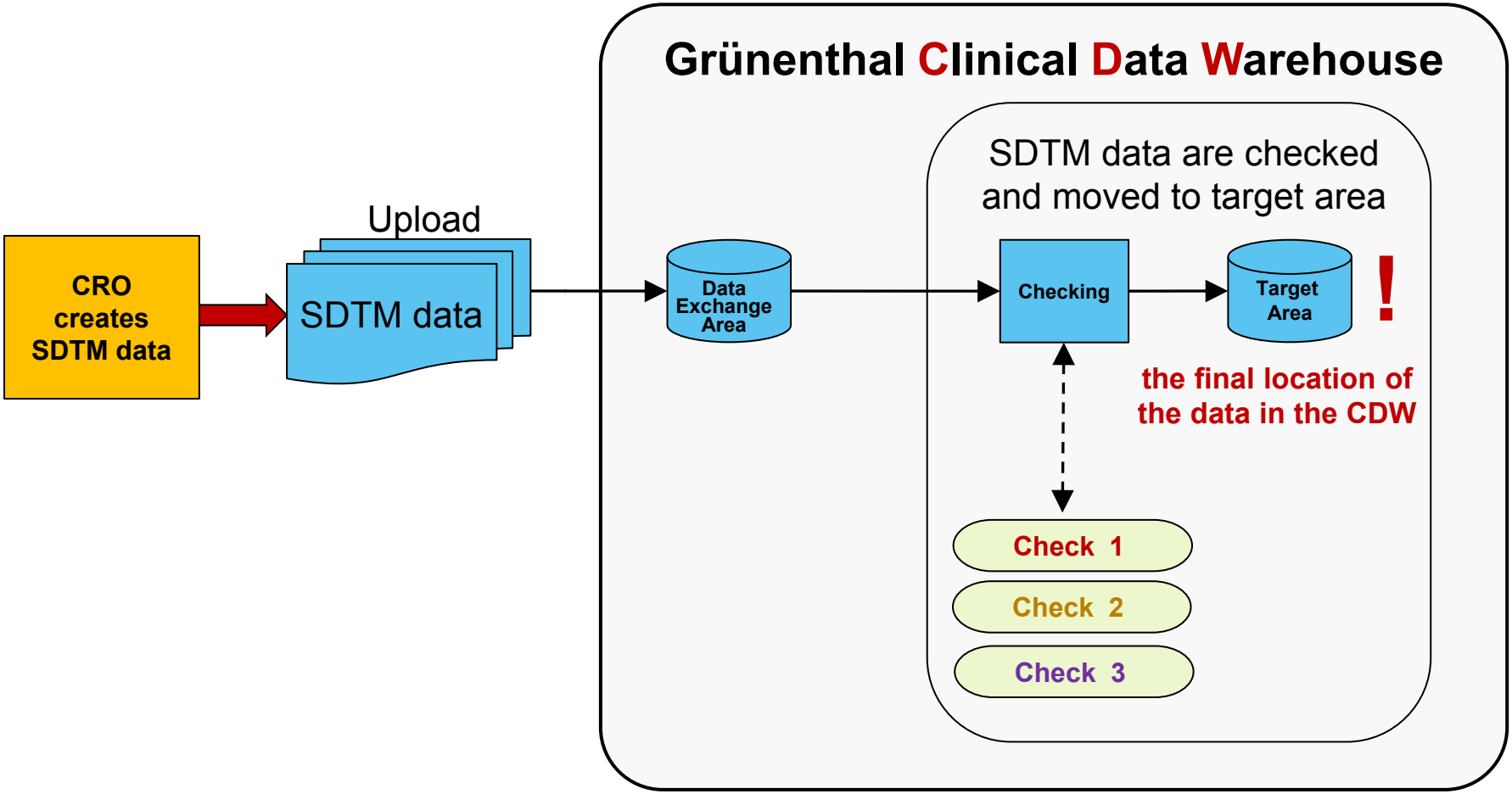
- The CRO creates the SDTM data according to the GRT standards based on the GRT template describing the **GRT SDTM Standard** (excel file) and the corresponding **Instruction for Use** document and **standard annotated CRFs**.
- For every delivery of SDTM data the CRO has to provide an **inventory file** together with the data (based on a GRT template) which includes an overview of the delivered data.

**Process of the receiving SDTM data to GRT**

- The received SDTM data should be checked and stored in the **Clinical Data Warehouse** (CDW) in the **target area**.
- The SDTM data, including the inventory file, will first be uploaded into the **data exchange area** of the Clinical Data Warehouse (CDW).
- Before the SDTM data will be moved to the **target area**, **three automated checks** will be performed.
- The aim of these checks is to find out if:
  - the data are in line with the delivered inventory file
  - all expected files are delivered / no additional files included
  - the delivered SDTM data are in line with the GRT SDTM Standard / data description

# Upload to the Clinical Data Warehouse

## Process at Grünenthal



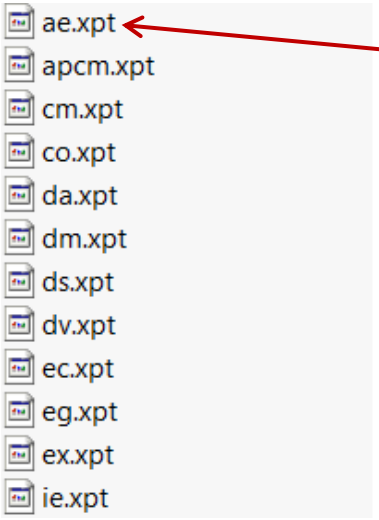
# Automated checks for uploaded CRO SDTM data

## Check 1

Check the SDTM data **against** the inventory file on the dataset level.

- The inventory file is created by the CRO and is a part of the SDTM package. It includes a list of all delivered files, the number of observations per file (only for SAS files) and the data status.

**Example receiving SDTM**




**Example for the inventory file of the SDTM delivery**

| trial id | data cat | filename | transporttype | no rows | data status |
|---|---|---|---|---|---|
| abcd | sdtm | ae.xpt | xport | 20 | intermediate |
| abcd | sdtm | apcm.xpt | xport | 0 | intermediate |
| abcd | sdtm | cm.xpt | xport | 491 | intermediate |
| abcd | sdtm | co.xpt | xport | 292 | intermediate |
| abcd | sdtm | da.xpt | xport | 236 | intermediate |
| abcd | sdtm | dm.xpt | xport | 27 | intermediate |
| abcd | sdtm | ds.xpt | xport | 90 | intermediate |
| abcd | sdtm | dv.xpt | xport | 98 | intermediate |
| abcd | sdtm | ec.xpt | xport | 22 | intermediate |
| abcd | sdtm | eg.xpt | xport | 276 | intermediate |
| abcd | sdtm | ex.xpt | xport | 22 | intermediate |
| abcd | sdtm | ie.xpt | xport | 5 | intermediate |

- The automated check compares the filenames, the specified transport type and the number of rows against the delivered SDTM data package.
- The result of the check will be displayed in a report.

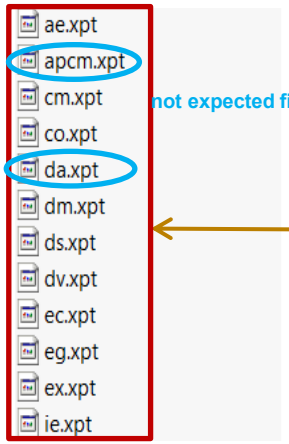
# Automated checks for uploaded CRO SDTM data

## Check 2

Check if all **expected** files have been received.

- At the beginning of a trial a **trial master file** is set up in the **CDW** specifying all expected data categories for the trial. Additionally, a list of expected files and a reference to the GRT SDTM Standard can be included for the data category.

**Example receiving SDTM**




**Example for the trial master file**

| trial_id | data_cat | filename | standard_name | standard_ver |
|---|---|---|---|---|
| abcd | sdtm | ae.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | cm.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | co.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | dm.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ds.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | dv.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ec.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | eg.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ex.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ie.xpt | grt_sdtm | 4_0_0 |

- The automated check compares the delivered SDTM data package against the list of expected files for the respective data category in the trial master file. The check is performed on a filename level.
- The result of the check will be displayed in a report.

# Automated checks for uploaded CRO SDTM data

## Check 3

Check if the received SDTM data are **in line with the GRT SDTM Standard** / data description.

- The trial master file can include a list of expected files for a data category – it can also include a link to the GRT SDTM Standard for every file.

**Example for the trial master file**

| trial_id | data_cat | filename | standard_name | standard_ver |
|---|---|---|---|---|
| abcd | sdtm | ae.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | cm.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | co.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | dm.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ds.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | dv.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ec.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | eg.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ex.xpt | grt_sdtm | 4_0_0 |
| abcd | sdtm | ie.xpt | grt_sdtm | 4_0_0 |

**information about the GRT SDTM Standard (version 4) used for every file**

- The automated check is performed on the dataset level.
- The dataset will be checked against the corresponding GRT SDTM Standard and version.
- The following information is checked: **variable name, variable label, variable type and variable length**.
- All inconsistencies will be displayed in a report.

# Automated checks for uploaded CRO SDTM data
## GRT SDTM Standard

- Example for GRT SDTM Standard, DM domain
  - Check the SDTM data on the dataset level against the defined GRT SDTM Standard **variable name, variable label, type and length**.

**GRT Standard Domains Requirements**

| Domain Prefix | Variable Name | Variable Label | CDISC Notes (for domains) Description (for General Classes) | Required/ Expected [Required/ Expected/ Not use | Controlled terms [empty=No, otherwise source of CT] | Format (length) | Type | Length | Example | Description Grünenthal (How to fill) | Derived [empty = CRF or raw data, otherwise source of derivation] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| DM | DOMAIN | Domain Abbreviation | Two-character abbreviation for the domain. | **Required** | **GRT and SDTM GRT00211 (DOMAIN) C66734 (DOMAIN)** | Char 2 | Char | 2 | DM | Derived variable will be set whilst creating the data sets. | **Controlled Terminology** |
| DM | DTHFL | Subject Death Flag | Indicates the subject died. Should be Y or null. Should be populated even when the death date is unknown. | **Expected** | **SDTM C66742 (NY)** | Char 2 | Char | 2 | Y | Indicates the subject died. Should be Y or null. Should be populated even when the death date is unknown. | |
| DM | AGEU | Age Units | Units associated with AGE. | **Expected** | **SDTM C66781 (AGEU)** | Char 10 | Char | 10 | YEARS | Units associated with AGE. | |
| DM | SEX | Sex | Sex of the subject. | **Required** | **SDTM C66731 (SEX)** | Char 20 | Char | 20 | M | Sex of the subject. | |
| DM | ETHNIC | Ethnicity | The ethnicity of the subject. Sponsors should refer to "Collection of Race and Ethnicity Data in Clinical Trials" (FDA, September 2005) for guidance regarding the collection of ethnicity (http://www.fda.gov/RegulatoryInformation/Guidances/ucm126340.htm). | **Expected** | **SDTM C66790 (ETHNIC)** | Char 40 | Char | 40 | HISPANIC OR LATINO | Ethnicity to be collected according to the FDA guideline.<br><br>If information other than specified in the FDA guideline is collected in ETHNIC, the result is to be mapped to the foreseen categories. | |
| DM | ARMCD | Planned Arm Code | ARMCD is limited to 20 characters and does not have special character restrictions. The maximum length of ARMCD is longer than for other "short" variables to accommodate the kind of values that are likely to | **Required** | **GRT GRT00031 (ARMCD)** | Char 20 | Char | 20 | T-P | The following coding should be used:<br>T or T1…Tx (for the GRT product, numbering indicates the different doses)<br>C or C1…Cx (for the comparator, numbering indicates the different doses or comparators)<br>P or P1…Px (placebo, number indicates different kind of placebo) | **Allocation list** |

Instructions & Abbreviation | History of Changes | **Data Structure** | GRT Controlled Terminology | SDTM Terminology

# Automated checks for uploaded CRO SDTM data

## Summary

1. Check 1 – Completeness check against the inventory file.
2. Check 2 – Expectedness check against the trial master file.
3. Check 3 – Consistency check against the GRT SDTM Standard.

**Many thanks for your attention!**

Wioleta Schulze
Technical Data Associate
Grünenthal GmbH
Zieglerstr. 6
52078 Aachen
wioleta.schulze@grunenthal.com
phone +49 241 569 2078