



Automization of SDTM Controlled Terminology checks

March 15, 2016

Background



Grünenthal (GRT) is currently outsourcing SDTM creation!

We need to check the quality of the SDTM data!

Adherence to Controlled terminology (CT) is one important aspect!

Manual checks?

NO!!!

- Too time-consuming
- Too error-prone
- Too boring



We need some automated solution!

But how?

Use OPEN CDISC checks?

- Good to check OPEN CDISC rules!

CDISC SDTM 3.1.2 Validation Rules

The following is a listing of CDISC SDTM version 3.1.2 validation rules supported by OpenCDISC Validator. The list includes SDTM Validation Specification v.1, CDISC controlled terminology rules, and additional rules defined by the OpenCDISC community.

Domain	Category	Type	WebSDM ID
<input type="text"/>	<Any>	<Any>	<input type="text"/>

Rule ID	Message	Description of Rule
CT0001	Value for --ACN not found in ACN controlled terminology codelist	Variable values should be populated with terms found in 'Action Taken with Study Treatment' (C66767) CDISC controlled terminology codelist
CT0002	Value for AESEV not found in AESEV controlled terminology codelist	Variable values should be populated with terms found in 'Severity/Intensity Scale for Adverse Events' (C66769) CDISC controlled terminology codelist

OpenCDISC

VALIDATOR FOR CDISC

- ✓ Open-source
- ✓ Fully Configurable
- ✓ Validates SDTM and more

- but how do we address GRT terminology or specific study terminology?

We need a GRT solution!

The plan



Technical implementation of checks

- Apply metadata to SDTM or ADaM data
- Derive violations

Reporting

- Show CT violations



1. Metadata generation

Metadata generation from the Excel file containing rules

Variable level (example)

Mapping



Mapping List

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
LB	LBNRIND	Reference Range Indicator	1. Indicates where the value falls with respect to reference range defined by LBORNRL0	Expected	GRT and SDTM GRT00083 (NRIND) C78736 (NRIND)	Char 20	Char	20

Codelists



Code List

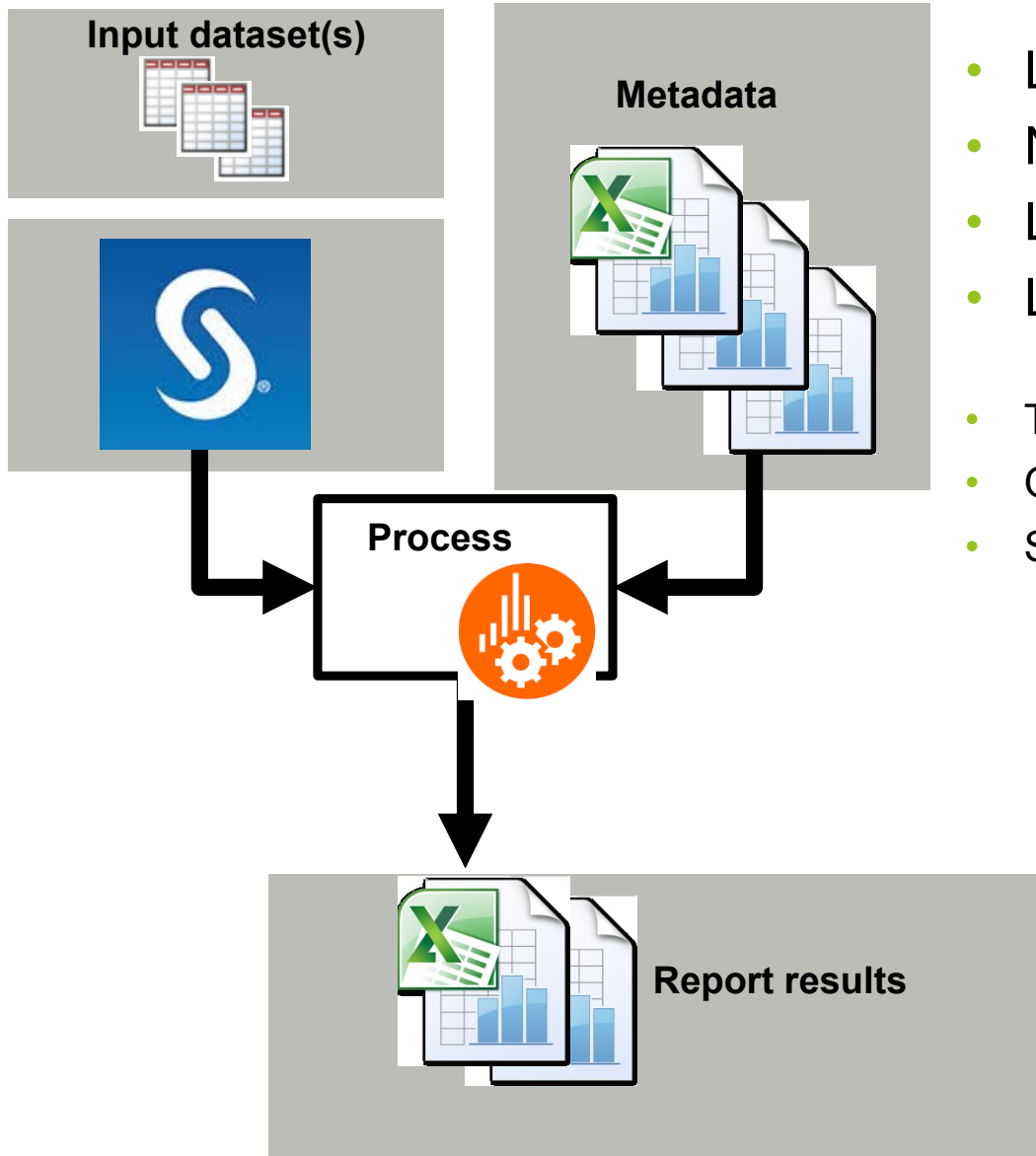
GRT Controlled Terminology					
Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	GRT Submission Value	GRT Synonym(s)
GRT00083	GRT00083	No	NRIND	NRIND	Reference Range Indicator
G000183	GRT00083		NRIND	AL	Alert Low
G000184	GRT00083		NRIND	AH	Alert High
G000185	GRT00083		NRIND	SL	Sponsor Alert Low
G000186	GRT00083		NRIND	SH	Sponsor Alert High
G000187	GRT00083		NRIND	EL	Exclusionary Low
G000188	GRT00083		NRIND	EH	Exclusionary High

SDTM Controlled Terminology				
Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value
C78736		Yes	Reference Range Indicator	NRIND
C78802	C78736		Reference Range Indicator	ABNORMAL
C78800	C78736		Reference Range Indicator	HIGH
C78801	C78736		Reference Range Indicator	LOW
C78727	C78736		Reference Range Indicator	NORMAL

Value level (examples)

Codelist	Condition
CMSCAT	valid only for CMCAT="PROTOCOL SPECIFIED MEDICATION"
CMSCAT	valid only for CMCAT="GENERAL NON-PHARMACOLOGICAL MODALITIES" and "PROTOCOL SPECIFIED NON-PHARMACOLOGICAL MODALITIES"
CMSCAT	Valid only for DSCAT="DISPOSITION EVENT"
DSSCAT	Valid only for DSCAT="OTHER EVENT"
DSDECOD	Valid only for DSCAT="PROTOCOL MILESTONE"

Technical Implementation with SAS: The parameters

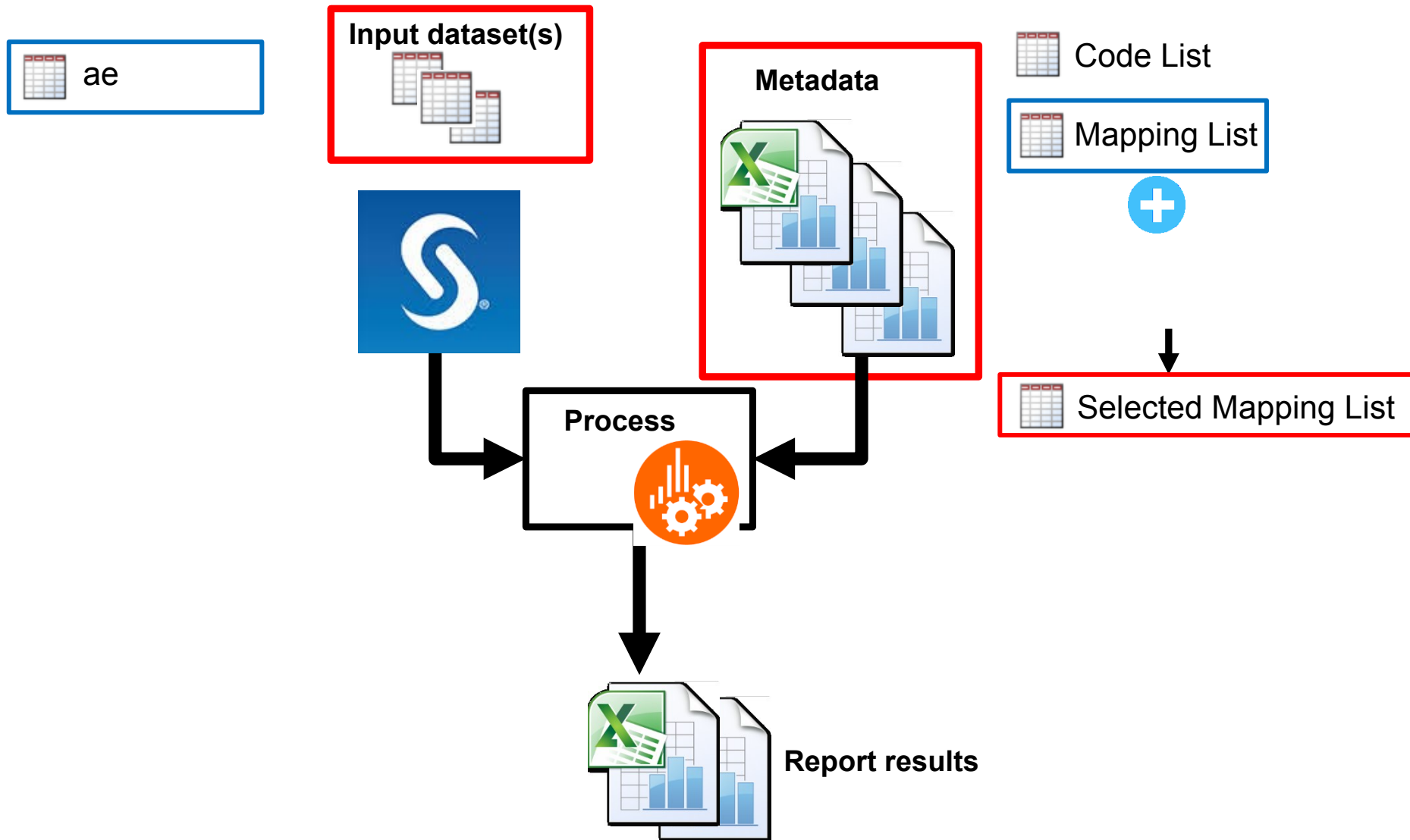


- Location of the data
- Name of the data sets (or `_all_`)
- Location and name of Metadata
- Location and name of report file

- Type of data (SDTM, ADaM)
- Output options
- Selection of data and checks

2. Technical Implementation: Example

Technical Implementation: Data input

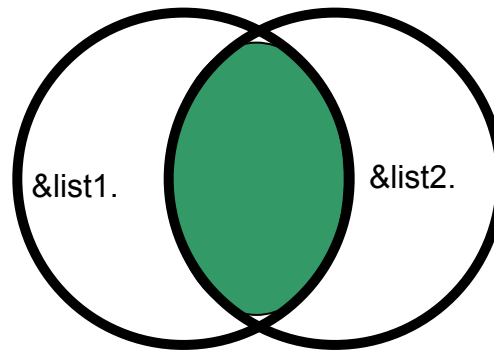


Technical Implementation: Intersection of two lists of values

 ae

 Mapping List

#	Variable	Type	Len
1	STUDYID	Char	200
2	DOMAIN	Char	200
3	USUBJID	Char	200
4	AESEQ	Num	8
5	AESPID	Char	200
6	AETERM	Char	200
7	AEMODIFY	Char	200
8	AEDECOD	Char	200
9	AEBODSYS	Char	200
10	AESEV	Char	200
11	AESER	Char	200
12	AEACN	Char	200
13	AEACNOTH	Char	200
14	AEREL	Char	200
15	AEPATT	Char	200
16	AEOUT	Char	200
17	AESTDTC	Char	200
18	AEENDTC	Char	200

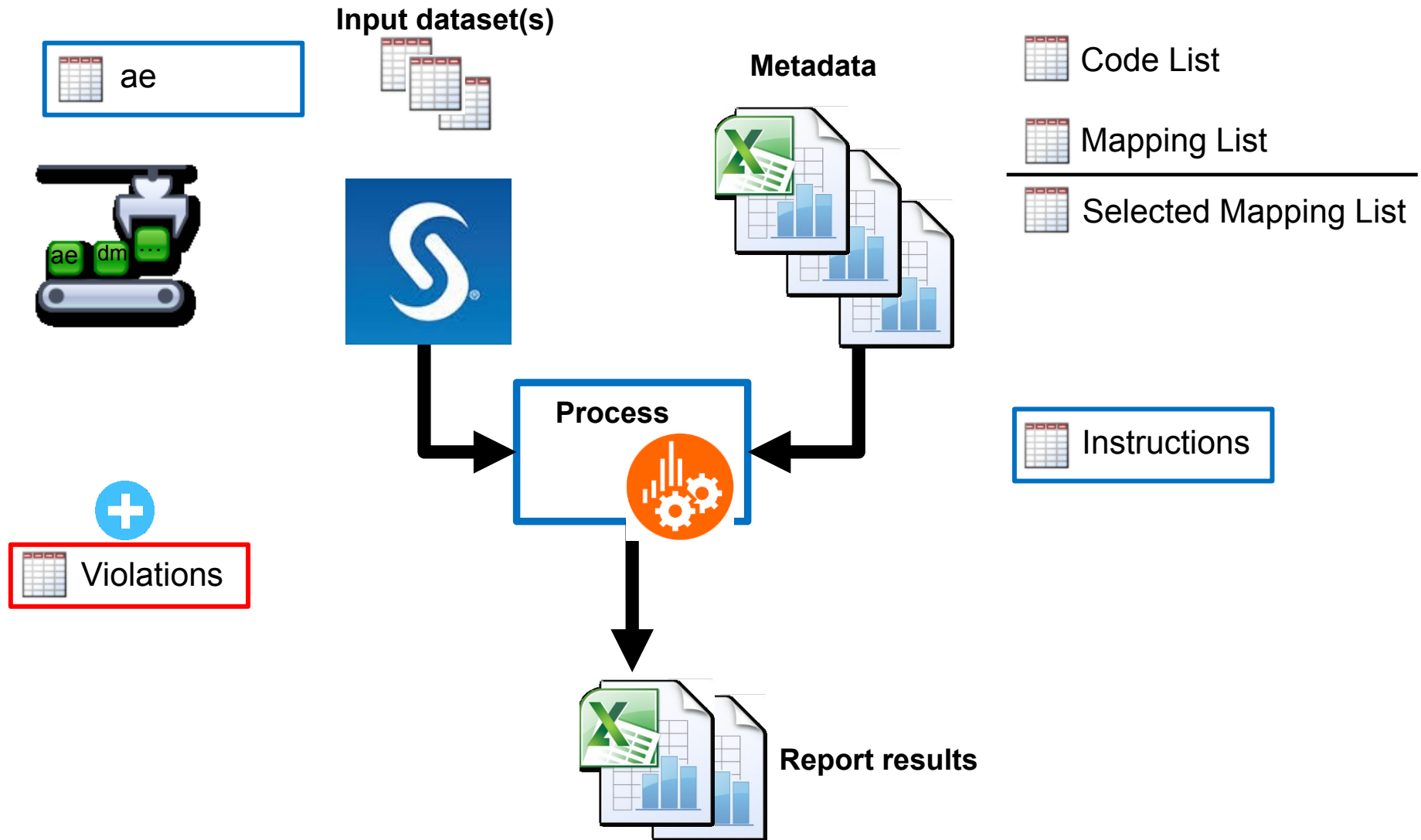


GRT Standard Domains Requirements				
Domain Prefix	Variable Name	Variable Label	Required/Expected	Controlled terms [empty=no, otherwise source of CT]
AE	DOMAIN	Domain Abbreviation	Required	GRT and SDTM GRT00211 (DOMAIN) C66734 (DOMAIN)
AE	AEPRESP	Pre-Specified Adverse Event	Expected	SDTM C66742 (NY)
AE	AESEV	Severity/Intensity	Required	SDTM C66769 (AESEV)
AE	AESER	Serious Event	Expected	SDTM C66742 (NY)
AE	AEACN	Action Taken with Study	Expected	SDTM C66767 (ACN)
AE	AEREL	Causality	Expected	GRT GRT00011 (AEREL)
AE	AEOUT	Outcome of Adverse Event	Expected	SDTM C66768 (OUT)
AE	AECONTRT	Concomitant or Additional	Expected	SDTM C66742 (NY)
AE	AEENRF	End Relative to Reference	Expected	SDTM C66728 (STENRF)
AE	AEENRPT	End Relative to Reference Time Point	Expected	SDTM C66728 (STENRF)
AE	EPOCH	Epoch	Expected	GRT GRT00043 (EPOCH)

 Selected Mapping List

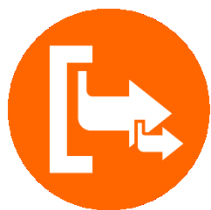
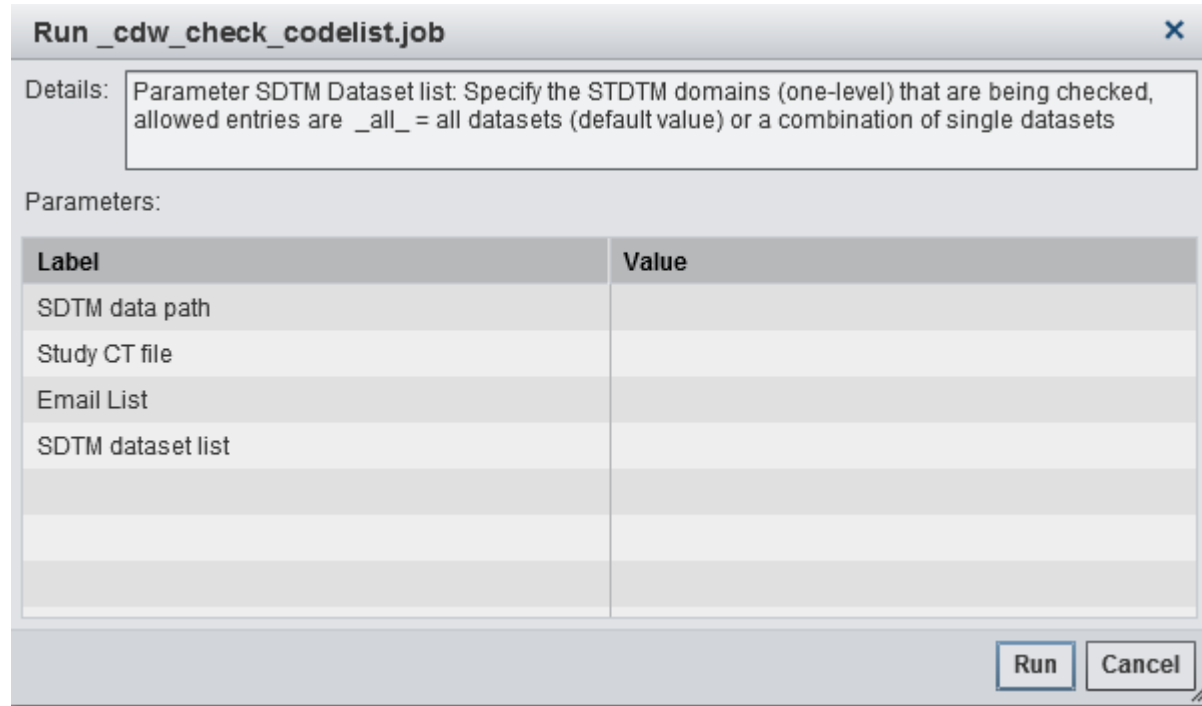
GRT Standard Domains Requirements				
Domain Prefix	Variable Name	Variable Label	Required/Expected	Controlled terms [empty=no, otherwise source of CT]
AE	DOMAIN	Domain Abbreviation	Required	GRT and SDTM GRT00211 (DOMAIN) C66734 (DOMAIN)
AE	AESEV	Severity/Intensity	Required	SDTM C66769 (AESEV)
AE	AESER	Serious Event	Expected	SDTM C66742 (NY)
AE	AEACN	Action Taken with Study	Expected	SDTM C66767 (ACN)
AE	AEREL	Causality	Expected	GRT GRT00011 (AEREL)
AE	AEOUT	Outcome of Adverse Event	Expected	SDTM C66768 (OUT)

Technical Implementation: Generation of results datasets



3. Application, Reporting and Conclusion

Application



SAS Job



SAS
Programs/Macros



Excel Output

Reporting

1	Libname	Domain	SAS Variable	Controlled term	Incorrect value	Extensible	Full comment
2	CDLIST	EG	EGSTRESC	EGSTRESC	101		Number of repetitions: 2.
3	CDLIST	EG	EGSTRESC	EGSTRESC	103		Number of repetitions: 2.
4	CDLIST	EG	EGSTRESC	EGSTRESC	104		Number of repetitions: 2.
5	CDLIST	EG	EGSTRESC	EGSTRESC	107		values: STUDYID= XXXXXX, USUBJID= XXXXX-GERXXXXX, VISITNUM= 3.

Controlled term. Violations

1	CT(non-grouped)	CT decode	Codelist code	xtensil	Long description	Specific code	Possible values	Value synonyms
2	EGSTRESC	EGSTRESC	C71150	Yes	ECG Result	C111088	1ST DEGREE AV BLOCK	Prolongation; Prolonged PR
3	EGSTRESC		C71150			C71044	2:1 AV BLOCK	2:1 AV block
4	EGSTRESC		C71150			C62016	2ND DEGREE AV BLOCK	2nd degree AV block
5	EGSTRESC		C71150			C111091	3RD DEGREE AV BLOCK	complete (third-degree);
6	EGSTRESC		C71150			C114165	50 Hz NOISE	50 Hz Artifact; 50 Hz Noise

Controlled term. Overview

1	number for mapping	Domain	Variable name	Description of violation	Value of --testcd	Value of --test	CT Code
2	C102277	LB	LBTESTCD	Multiple CT entries of --TEST for given entry of --TESTCD	APPTLAS	Lupus Anticoagulant Sensitive APTT	C65047
3	C102277	LB	LBTESTCD	Multiple CT entries of --TEST for given entry of --TESTCD	APPTLAS	Lupus Anticoagulant Sensitive APTT	GRT00084
4	C28407	SC	SCTESTCD	Multiple CT entries of --TEST for given entry of --TESTCD	NATORIG	National Origin	C74559
5	C28407	SC	SCTESTCD	Multiple CT entries of --TEST for given entry of --TESTCD	NATORIGIN	National Origin	GRT00151

CT Test vs Testcd

Conclusion

- Can be used within SAS programs
- ADaM or SDTM
- Multiple studies can be checked simultaneously
- Users can control and modify input metadata (sponsor or study specific codelists)
- Scope is addressed via macro parameters
 - All domains of current study
 - Specific domains across various studies etc.

Questions?



Many thanks for your attention!

Frank Biedermann

Senior Programmer

Contact

phone +49 241 569 2462

mobil +49 151 12263220

email frank.biedermann@grunenthal.com

Grünenthal GmbH

Zieglerstr. 6 • 52078 Aachen

www.grunenthal.com

Guido Wendland

Principal Programmer

Contact

phone +49 241 569 2065

email Guido.Wendland@grunenthal.com

Grünenthal GmbH

Zieglerstr. 6 • 52078 Aachen

www.grunenthal.com



GRÜNENTHAL