



Science For A Better Life

Challenges of different MedDRA versions for integrated data and databases

Peter Bonata, Statistical Analyst

Bayer Pharma, Wuppertal, Germany



Agenda/ Content

- Introduction
- MedDRA
- MedDRA Update
- Workflow to update datasets
- Conclusion



Introduction

- One of the main tasks based on an integrated database is the analysis of Adverse Events.
- Therefore the Medical Dictionary for Regulatory Activities (MedDRA ®⁽¹⁾) is global standard. This coding dictionary is maintained and updated twice a year.

¹ Medical Dictionary for Regulatory Activities (MedDRA).



Introduction

- A process has to be implemented, that ensures during clinical development, for Integrated Summary of Safety (ISS) for submissions and during life cycle management:
 - a unique thesaurus version across different studies in an integrated database
 - allows the traceability of changes,
 - the reproducibility of results using former versions
 - that datasets are still analyses ready.



Introduction

In available documentation (ADaMIG ²) and (draft) guidelines is no clear guidance how to address this topic for the integration of multiple clinical studies:

- The “same single version of MedDRA [should be used] to avoid version specific differences and to correctly identify number of adverse events.
- It is recommended but not required that all levels of terms for the primary path in the MedDRA hierarchy are included (³).
- A sponsor can create additional analysis datasets for AE analysis, even when using a different structure (⁴).

² Analysis Data Model (ADaM) Implementation Guide Version 1.0

³ Analysis Data Model Structure for Occurrence Data Version 1.0 Draft

⁴ Analysis Data Model (ADaM) Data Structure for Adverse Event Analysis Version 1.0



Introduction

- The ADAE structure for the standard AE safety dataset is clearly outlined: Keeping multiple sets of mapping variables is not common, but it might be usable when different versions are used for interim and final analyses or when studies are pooled for an integrated analysis.
 - A proposal for providing traceability is to introduce a “counter” in the variable name.
 - This integer “y” from 1 to 9 refers to a previous version. In the metadata the description is enriched by the dictionary name and version.
 - In addition Standardized MedDRA Query (SMQ) - as groupings of MedDRA terms at the PT level - or Customized MedDRA Queries (CMQ) variables are available in AE datasets. Here is the recommendation to use a number starting from 01 for each query of interest.



Introduction

■ Example:

Table 4.2.10.1 Original or Prior MedDRA Coding Variables

Variable Name	Variable Label	Type	Code List / Controlled Terms	Core	CDISC Notes
DECDORGy	PT in Original Dictionary y	Char	MedDRAy*	Perm	Original preferred term coding of XX.--TERM using MedDRA or other dictionary version X.X.
BDSYORGy	SOC in Original Dictionary y	Char	MedDRAy*	Perm	Original body system coding of XX.--TERM using MedDRA or other dictionary version X.X.
HLGTORGy	HLGT in Original Dictionary y	Char	MedDRAy*	Perm	Original HLGT coding of XX.--TERM using MedDRA or other dictionary version X.X.
HLTORGy	HLT in Original Dictionary y	Char	MedDRAy*	Perm	Original HLT coding of XX.--TERM using MedDRA or other dictionary version X.X.
LLTORGy	LLT in Original Dictionary y	Char	MedDRAy*	Perm	Original LLT coding of XX.--TERM using MedDRA or other dictionary version X.X.
LLTNORGy	LLT Code in Original Dictionary y	Char	MedDRAy*	Perm	Original LLT code of XX.--TERM using MedDRA or other dictionary version X.X.

* For each version of an external dictionary, a different reference name must be used. The individual reference names will point to a dedicated section in the data definition file where all external dictionaries used in the analysis are listed, including dictionary name and version.

Source: CDISC ADaM occurrence Data Structure (Version 1.0 draft), p.22



Introduction

■ Example:

Table 4.2.9.1 Standardized MedDRA Query Variables

Variable Name	Variable Label	Type	Code List / Controlled Terms	Core	CDISC Notes
SMQzzNAM	SMQ zz Name	Char		Cond	The standardized MedDRA queries name. Would be blank for terms that are not in the SMQ. Therefore this variable could be blank for all records if no terms within the study were included in the SMQ. Conditional on whether SMQ analysis is done.
SMQzzCD	SMQ zz Code	Num		Perm	The standardized MedDRA queries number code.
SMQzzSC	SMQ zz Scope	Char	BROAD, NARROW	Cond	The search strategy for SMQs can be narrow or broad. The preferred terms that are narrow in scope have high specificity for identifying events of interest while the broad terms have high sensitivity. By definition, all narrow terms are also considered within the broad score. Therefore, to summarize all broad terms, terms with either narrow OR broad would be considered. Will be null for terms that do not meet the criteria. Conditional on whether SMQ analysis is done.
SMQzzSCN	SMQ zz Scope (N)	Num	1, 2	Perm	Will be null for terms that do not meet the criteria.
CQzzNAM	Customized Query zz Name	Char		Cond	The customized query (CQ) name or name of the AE of special interest category based on a grouping of terms. Would be blank for terms that are not in the CQ. Conditional on whether CQ analysis is done. Examples: "DERMATOLOGICAL EVENTS" "CARDIAC EVENTS", "IARS (INFUSION ASSOCIATED REACTIONS)"

Source: CDISC ADaM occurrence Data Structure (Version 1.0 draft), p.21

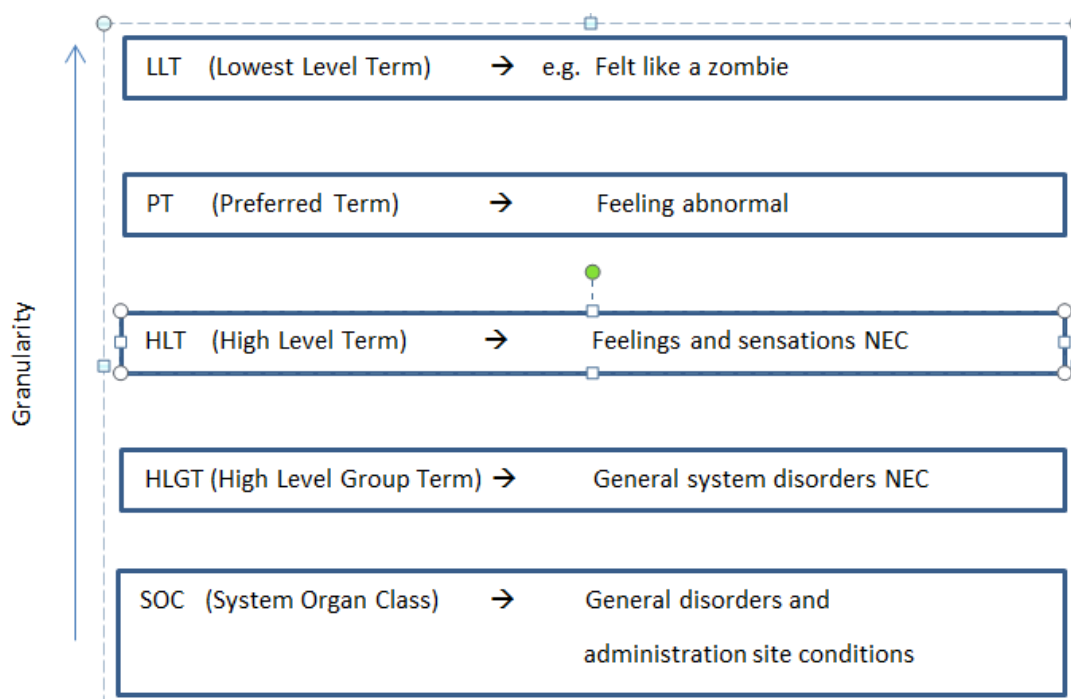


Introduction

- Limitations:
- This counter can only be used when the continuous updating of MedDRA does not exceed 4 ½ years or nine MedDRA versions.
- The detachment of an integer to the numeric thesaurus version is not intuitive and might also lead to confusion and complicates the supportive traceability.
- Not sufficient when not only the primary path is added to the reported terms.
- The implementation of different MedDRA versions for SMQ/ CMQ.

MedDRA

- The structural hierarchy of the MedDRA Terminology is simplified represented as:





MedDRA update

- In cooperation with a centralized global Coding an updating to the controlled vocabulary of the existing thesauri version is initialized.
- To achieve homogenous, consistent and medical accurate coding output a macro is used for this so-called REFRESH process for in house coded studies in an integrated database, that is based on the clinical studies analysis data sets.
- The up-versioning is done by unique terms for the selected thesaurus as the combination of the existing verbatim (AETERM), the modified verbatim (AEMODIFY) and the clarified verbatim (AETERMCV). A consistency is reached as individual terms are coded initially identically in a single clinical study and also across several studies in an integrated data base.



MedDRA update

- An initial SAS® macro gives the necessary information to the coding environment so that their internal macros can update the requested dictionary.



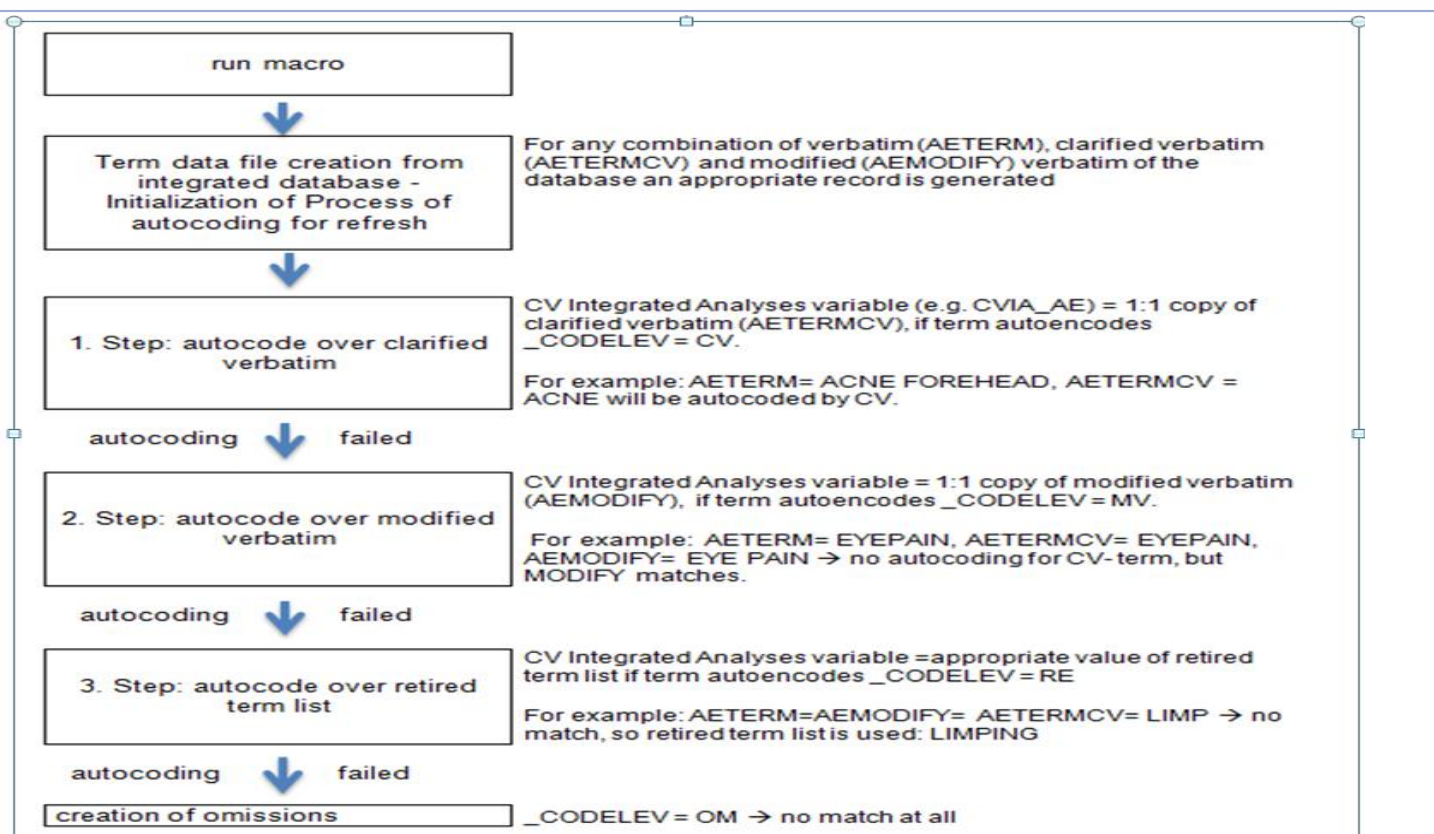
MedDRA update

■ Macro parameters

SOURCE_DIR	= data directory
SOURCE_DATA_SET	= ADAE
TERM_DATA_SET	= ADAExxx
REPORT_DIRECTORY	= RESULTS
IA_CREATE_OMISSIONS	= Y/N
IA_CREATE_TRR	= Y/N → a Term Review Report (datasets/ Microsoft Office® Word and Excel files) for documentation and investigation
ORIGINAL_TERM_VARIABLE	= AETERM
CLARIFIED_TERM_VARIABLE	= AETERMCV
MODIFIED_TERM_VARIABLE	= AEMODIFY
WORK_FLOW_VARIABLE	= AETERMWF
CODE_VARIABLE_1	= AEMLLT
CODE_VARIABLE_2	= needed for Concomitant Medication WHO DD coding
CODE_VARIABLE_3	= needed for Concomitant Medication WHO DD coding
CODING_THESAURUS	= MedDRA
IA_IDENT	= &identifier.
TITLE1_REPORT	=ADVERSE EVENT TERMS....



MedDRA update





MedDRA update

- In a several step approach the existing verbatims will be checked against the current thesaurus. In a first step it is looked for all terms that autocode. This means no further action is required and those terms are updated including the new lowest level term code.
- Terms that don't autocode are named omissions and are delivered to the coding environment, from where they have to be coded manually by medical coders after medical review.
- When this task is successfully completed the macro has to rerun to achieve a complete updated dataset to the most recent version.
- After the coding process is finalized a so called Term data file is available and the encoding information can be used for further processing.



MedDRA update

■ Term level file:

Variables	Description	
AETERM	Reported Term for the Adverse Event	Original verbatim in database
AETERMCV	Clarified Verbatim	Original clarified verbatim in database
AEMODIFY	Modified Reported Term	Original modified verbatim in database
AETERMWF	Workflow Step for Coding	Workflow variable
MODIA_AE	MODIFIED VERB. FOR IA	New modified verbatim for recent version and future updates
AEMLLT	MedDRA Term Code	New AEMLLT in most recent MedDRA version
_CODELEV	FLAG for CODING LEVEL(CV, MV, RE or OM)	Which level is used for update
.....		
QUERY_xxxx		Variables for coding function related to the coding process
.....		
THES_VERSION	THESAURUS VERSION	Most recent version, e.g. 17.0
UPDATE_D	DAY TERM WAS ADDED TO TERM DATA SET	
UPDATE_M	MONTH TERM WAS ADDED TO TERM DATA SET	
UPDATE_Y	YEAR TERM WAS ADDED TO TERM DATA SET	



Workflow to update datasets

- Bayer Integrated Analyses has chosen an approach that creates two datasets:
 - An updated ADAE dataset containing the most recent MedDRA version
 - A so-called AE history dataset ADAEHIST storing the prior dictionary version(s)
- Therefore a macro can be used that uses the information from the coding process and enriches the ADAE dataset with the complete MedDRA hierarchy.



Workflow to update datasets

- Following basic parameters are needed:

DATASET	: dataset to be updated
PREFIX_	: prefix for MedDRA-Variables
TERMDATA	: dataset containing new codes (including library name)
OLDVERSION	: old MedDRA version
REPORTTERM	: variable containing reported term AETERM
CLARIFIEDTERM	: variable containing clarified term AETERMCV
MODIFIEDTERM	: variable containing modified term AEMODIFY
CODEVARIABLE	: variable containing updated code AELLT
APPEND	: if history file already exists =y to append to history file



Workflow to update datasets

- History dataset ADAEHIST
 - Initially creation of a history dataset : → “rename” of current ADAE dataset
 - Update an existing ADAEHIST: → “append” current ADAE to ADAEHIST
 - In both cases an variable with the respective MedDRA- version is added
- Updating ADAE:
 - PROC SQL: merging updated information from data term level file via a left join to the ADAE dataset by using AETERM, AEMODIFY and AETERMCV
 - The AEMODIFY variable is updated.
 - The new LLT– code variable is used to add the hierarchy (i.e. PT, HLT, HLGT, SOC) to ADAE via attaching the corresponding MedDRA formats



Conclusion

- Advantages:
 - ADAE stays well-arranged
 - Does not exceed the allowed size for electronic submissions to the FDA, in ongoing submission the former MedDRA versions can be added to the ADAE dataset.
 - Analyses based on former versions can be reproduced without changing of existing programs. Only a macro parameter for the ADAE(HIST) dataset and the required MedDRA version is needed.
 - Traceability can be achieved, when comparing the datasets respectively the different dictionary versions. This can be supported by a difference file or listing.
- Areas for improvement:
 - The complete MedDRA hierarchy is updated in the term data set.



Science For A Better Life

Thank you!

Peter Bonata
Bayer Pharma AG
D-42113 Wuppertal
Phone: +49 202 365263
Email : peter.bonata@bayer.com