Uppsala Monitoring Centre

WHO Collaborating Centre for International Drug Monitoring

2017

How to use WHODrug for compliance with CM domain in the CDISC SDTM standard

a technical guide for industry

INSPIRE. ENGAGE. TRANSFORM.

Introduction

The Uppsala Monitoring Centre (UMC) has been providing the de facto standard drug dictionary for analysis of concomitant drug data for over 30 years. WHODrug is now the dictionary recommended by leading national authorities globally. At UMC we are constantly evolving our products to meet the needs of users and to ensure regulatory compliance. In recent years the CDISC standards for electronic submission of study data have evolved and regulatory authorities are to require or strongly recommend using the CDISC standards in NDA submissions. Due to the regulatory updates UMC is taking action to facilitate compliance with the standards in order to continue enabling good data analysis by both industry and regulatory agencies

Background

The CDISC SDTM (Study Data Tabulation Model) standard describes how to report data from a clinical trial and includes guidance for reporting of concomitant medications. The CDISC implementation guide does not state how to retrieve the concomitant medication data from a drug dictionary; this guide aims to give data managers and SAS programmers detailed information for how to retrieve the correct data from WHODrug for inclusion in the Concomitant medication (CM) domain and as a result be fully compliant with the CDSIC SDTM standard.

This guide is based on the CDISC SDTM standard version 1.4[1] and implementation guide version 3.2. There are five variables within the CM domain that are most relevant with relation to WHODrug: CMTRT, CMMODIFY, CMDECOD, CMCLAS and CMCLASCD. The following table displays the fields and explanation from the implementation guide:

CMTRT	Verbatim medication name that is either pre-printed or collected on a CRF.
CMMODIFY	If CMTRT is modified to facilitate coding, then CMMODIFY will contain the modified text.
CMDECOD	Standardized or dictionary-derived text description of CMTRT or CMMODIFY. Equivalent to the generic medication name in WHO Drug. The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external code list attributes. If an intervention term does not have a decode value in the dictionary then CMDECOD will be left blank.
CMCLAS	Drug class. May be obtained from coding. When coding to a single class, populate with class value. If using a dictionary and coding to multiple classes, then follow assumption 4.1.2.8.3 or omit CMCLAS.
CMCLASCD	Class code corresponding to CMCLAS. Drug class. May be obtained from coding. When coding to a single class, populate with class code. If using a dictionary and coding to multiple classes, then follow assumption 4.1.2.8.3 or omit CMCLAS.

[1]Study Data Tabulation Model v1.4, www.cdisc.org

The most important concomitant medication information that can be retrieved from WHODrug is CMDECOD, CMCLAS and CMCLASCD. WHODrug comes in different formats and depending on which format is used, there are different approaches for how to retrieve the data.

CMDECOD

According to the implementation guide the CMDECOD field should contain a standardized or dictionary-derived text of the given medication, corresponding to the generic name in WHODrug. There are generic records for all drugs and drug combinations in WHODrug but they are displayed differently depending on the format. For single ingredient products the generic names are the same as the Preferred Names, independent of format: the Preferred Names have Drug Codes ending with 001. For multi-ingredient drugs in the B2/C formats the Preferred Name is sometimes a trade name and the generic name needs to be retrieved based on other properties.

Furthermore, the generic record may show truncated generic names due to the character limitation of 45 in the B2-format and 80 in the C-format. In the B3- and C3-formats both these issues are resolved. The Preferred Name is always generic and the full generic name can be displayed since the character limitation is set to 1500 instead of 45 or 80. Thus, the B2- and C-formats need workarounds but B3 and C3 do not.

CMDECOD in B2-format

To obtain the generic names without truncation there are two approaches which are valid for both single and multi-ingredient drugs and therefore either of these methods are recommended over using the preferred or generic flagged names:

1 Use Ingredients_longtext, available in the additional features folder for all releases since 2012. The Ingredients_longtext file makes it very easy to exchange the Preferred Name to the active ingredients, both for single and multi-ingredient drugs. The Drug Code of the Preferred Name is used to join the files. The Ingredients_longtext file returns the generic name, and substances in alphabetical order are separated by semicolon. An example from Ingredients_longtext is shown in figure 1.

```
00132201001Caffeine;Mepyramine maleate;Noscapine;Paracetamol;Pheniramine maleate;Phenylpropanolamine hydrochloride;Terpin hydrate
00132301001Aminoacridine;Naphazolinehydrochloride
00132501001Thialbarbital
00132502001Thialbarbital sodium
00132601001Hydrochlorothiazide;Reserpine
00132701001Benzylpenicillin;Streptomycin
00132801001Ajmaline;Raupine;Rescinnamine;Reserpine hydrochloride;Yohimbic acid
00132901001Betaine hydrochloride;Pepsin
00133001001Chlorhexidine
00133002001Chlorhexidine gluconate
00133003001Chlorhexidine hydrochloride
00133004001Chlorhexidine diacetate
00133101001Amiodarone
00133102001Amiodarone hydrochloride
00133201001Caffeine;Calcium pantothenate;Liver extract;Nicotinamide;Pyridoxine hydrochloride;Quinine;Riboflavin;Thiamine hydrochloride
```

Figure 1.Sample from the Ingredients_longtext file. The preferred Drug Code can be used to retrieve list of active ingredients separated by semicolon.

2. The second option is to pull the active ingredients from the original, standard files. This option must be used for WHODrug releases prior to 2012. The active ingredients are found in the files ING.txt and BNA.txt by linking the Drug Code of the Preferred Name to ING.txt, to get a list of CAS numbers and then linking each CAS number to BNA.txt to get the ingredient name corresponding to each CAS number. The relationships between the files are illustrated in figure 2. If there is more than one ingredient, the ingredients should be listed in alphabetical order with a semicolon as separator in order to achieve the same result as in Ingredients_longtext.

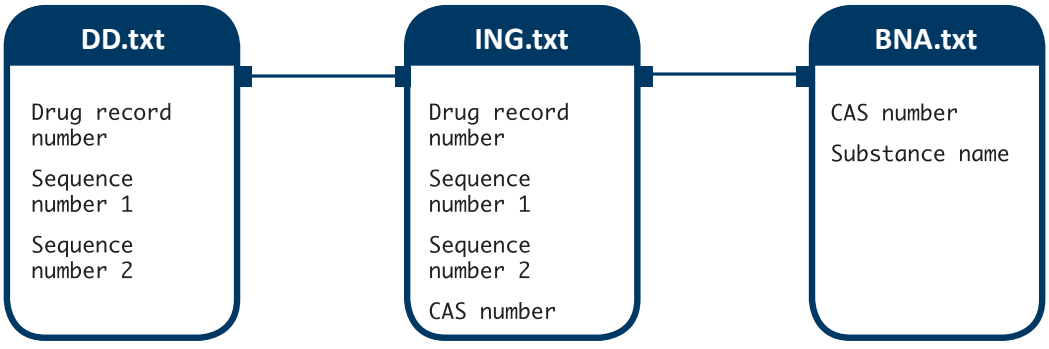

Figure 2. The linkage between drug name and generic name in WHODrug B2-format.

Drug Record Number, Sequence Number 1 and Sequence Number 2 together constitute the Drug Code.

CMDECOD in C-format

As for the B2-format, there are two options for obtaining the full generic name. The C-format allows for more characters (80) but since it may not be enough it is still recommended to use any of the 2 options to ensure the full generic name is extracted.

1 This option is exactly the same as for alternative 1 for the B2-format: use Ingredients_longtext, available in the additional features folder for all releases since 2012. The Ingredients_longtext file makes it very easy to exchange the Preferred Name to the active ingredients, both for single and multi-ingredient drugs. The Drug Code of the Preferred Name is used to join the files. The Ingredients_longtext file returns the generic name and substances are in alphabetical order separated by semicolon. Example from Ingredients_longtext is shown in figure 3.

```
00132201001Caffeine;Mepyramine maleate;Noscapine;Paracetamol;Pheniramine maleate;Phenylpropanolamine hydrochloride;Terpin hydrate
00132301001Aminoacridine;Naphazolinehydrochloride
00132501001Thialbarbital
00132502001Thialbarbital sodium
00132601001Hydrochlorothiazide;Reserpine
00132701001Benzylpenicillin;Streptomycin
00132801001Ajmaline;Raupine;Rescinnamine;Reserpine hydrochloride;Yohimbic acid
00132901001Betaine hydrochloride;Pepsin
00133001001Chlorhexidine
00133002001Chlorhexidine gluconate
00133003001Chlorhexidine hydrochloride
00133004001Chlorhexidine diacetate
00133101001Amiodarone
00133102001Amiodarone hydrochloride
00133201001Caffeine;Calcium pantothenate;Liver extract;Nicotinamide;Pyridoxine hydrochloride;Quinine;Riboflavin;Thiamine hydrochloride
```

Figure 3 . Sample from the Ingredients_longtext file. The preferred Drug Code can be used to retrieve list of active ingredients separated by semicolon.

2. The generic name can also be obtained from the original standard files. This option must be used for WHODrug releases prior to 2012. Use the Medicinalprod_Id in the MP file and link to the Substance_Id in the ING file. Use the Substance_Id to link to the substance name from the SUN file. The relationships between the files are illustrated in figure 2. If there is more than one ingredient, the ingredients should be listed in alphabetical order with a semicolon as separator in order to achieve the same result as in Ingredients_longtext.

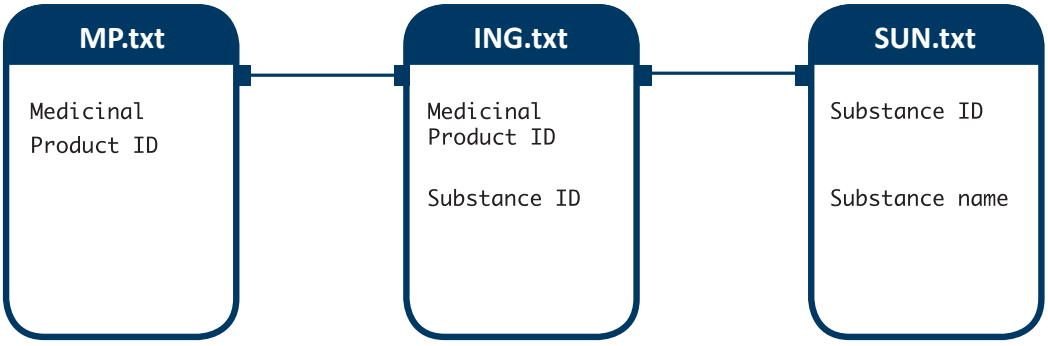

Figure 4. The linkage between drug name and generic name in WHODrug C-format.

Drug Record Number, Sequence Number 1 and Sequence Number 2 together constitute the Drug Code.

CMDECOD in B3- and C3-format

The B3- and C3-formats were introduced in March 2017. These formats are designed to remove the workaround for retrieving generic names in B2- and C-formats. To obtain the generic name in the B3- and C3-formats, simply use the Preferred drug name from drug name field in MP.txt or DD.txt, depending on format used. Based on company conventions the Preferred Name is the Drug Code ending with 01 001 (Preferred Base Name) or ending with 001 (Preferred Salt Name).

CMDECOD is longer than 200 characters

For drugs with many ingredients, the generic name is longer than 200 characters. The SAS export format has a limitation to 200 characters per field, if this format is used for submission, the supplemental dataset needs to be utilised. Note that the guidelines state that the text should be truncated between words, in the case for long generic names the text should be truncated after the semicolon closest to 200 characters. Illustrations of the ordinary and supplemental datasets are shown in table 1 and 2.

Table 1. Illustration of SDTM dataset where CMDECOD is longer than 200 characters.

USUBJID	CMSEQ	CMTRT	CMMODIFY	CMDECOD	CMCLAS	CMCLASCD
AB-21-01	1	….	…	Ascorbic acid;Biotin;Calcium;Carbohydrates nos; Chloride;Choline;Chromium;Colecalciferol; Copper;Cyanocobalamin;Docosahexaenoic acid; Fats nos;Folic acid;Fructooligosaccharides; Iodine;Iron;Magnesium;	….	….

Table 2. Illustration of supplemental dataset for CM domain where CMDECOD is longer than 200 characters.

USUBJID	RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
AB-21-01	CM	CMSEQ	1	CMDECOD1	Standardized Medication Name 1	Manganese;Nicotinic acid;Pantothenic acid;Phosphorus;Phytomenadione; Potassium;Proteins nos;Pyridoxine; Retinol;Riboflavin;Selenium;Sodium; Thiamine;Vitamin e nos;Zinc

CMCLAS and CMCLASCD

CMCLAS and CMCLASCD refer to the classification from the drug dictionary. For WHODrug the classification is WHO ATC classification.

In the implementation guide there are three ways of submitting ATC information described:

1. One single class selected
2. Multiple classes selected
3. No classification

From our current understanding based on available indications the authorities will want to get a classification of the concomitant drugs, but it has not yet been specified if they will require one or all ATC codes. Therefore it is recommended to use option 1 or 2 of the above.

CMDECOD in CRT Japan

CRT Japan is already designed to make it easy to find the generic name: the field 'generic name' in the WHODD Genericnames file can be used directly for CMDECOD.

Single class ATC code

When one of the ATC codes available in WHODrug is submitted it is important to understand that the one ATC code must be manually selected based on information available on the CRF from the investigator. It is not recommended to randomly select one ATC code, for example to choose the first or last of ATC codes in the list. An example of a manually selected ATC code is displayed in table 3.

Table 3. Illustration of SDTM dataset where single class ATC code is used.

USUBJID	CMTRT	CMMODIFY	CMDECOD	CMCLAS	CMCLASCD
AB-21-01	Asprina 03	Aspirina 03	Acetylsalicylic acid; Aluminium glycinate; Magnesium hydroxide	Platelet aggregation inhibitors excl. heparin	B01AC

Multiple classes ATC code

When using WHODrug, the alternative for multiple classes is equivalent to submitting all ATC codes available for a specific drug name. Some drugs have one ATC code only and in these cases the single ATC code structure should be used. Examples of multiple classes ATC code are displayed in Table 4 and 5.

Table 4. Illustration of ordinary dataset where multiple classes ATC code is used. Aspirina 03 has several ATC codes in WHODrug and Tylenol has one ATC code.

USUBJID	CMSEQ	CMTRT	CMMODIFY	CMDECOD	CMCLAS	CMCLASCD
AB-21-01	1	Asprina 03		Acetylsalicylic acid; Aluminium glycinate; Magnesium hydroxide	MULTIPLE	MULTIPLE
AB-21-01	2	Tylenol (actaminophen)	Tylenol	Paracetamol	Anilides	N02BE

Table 5. Illustration of supplemental dataset for CM domain where the decode for Aspirina 03 has several ATC codes.

USUBJID	RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
AB-21-01	CM	CMSEQ	1	CMCLAS1	Medication Class 1	Platelet aggregation inhibitors excl. heparin
AB-21-01	CM	CMSEQ	1	CMCLSCD1	Medication Class Code 1	B01AC
AB-21-01	CM	CMSEQ	1	CMCLAS2	Medication Class 2	Salicylic acid and derivatives
AB-21-01	CM	CMSEQ	1	CMCLSCD2	Medication Class Code 2	N02BA
AB-21-01	CM	CMSEQ	1	CMCLAS3	Medication Class 3	Other agents for local oral treatment
AB-21-01	CM	CMSEQ	1	CMCLSCD3	Medication Class Code 3	A01AD

UMC assistance

On the WHODrug User Area there is more extensive documentation, information and guidance on the specific WHODrug terminology such as Drug Code, ATC codes, text files and best practices. The WHODrug User Area can be found at www.who-umc.org.

For any questions and concerns please contact WHODrug@who-umc.org.