PharmaSUG 2020 - Paper DS-109

Impact of WHODrug B3/C3 Format on Coding of Concomitant Medications

Lyma Faroz, Seattle Genetics, Inc., Bothell WA
Jinit Mistry, Seattle Genetics, Inc., Bothell WA

ABSTRACT

The WHODrug dictionary is the industry standard for coding concomitant medications. As CDISC becomes more prevalent and strongly recommended by regulatory authorities for submission, the dictionary has evolved over time to ensure full compliance. It is maintained by the Uppsala Monitoring Centre (UMC) with updates provided to industry users twice every year, that is, 1st March and 1st September. The previous WHODrug B2/C formats are now up versioned to B3/C3, which make WHODrug coded data fully compliant with the expectations of regulatory authorities and bring heightened efficiency and other benefits to the industry. The older vs newer format length updates have impact on mapping of coding concomitant medication data according to SDTM CM guidelines. To add to that, per a notice in the Federal Register published by FDA in October 2017, the use of the B3 format is required in submissions of studies starting after 15th March 2019. Hence, it is critical for statistical programmers to learn and be aware of these updates and apply them in new studies.

In this paper, we will describe how the WHODrug B3 and C3 formats relate to the U.S. FDA Data Standards Catalog, shed light on aspects relevant to statistical programmers receiving concomitant medication data in these formats, and illustrate efficient ways of handling them in the SDTM CM domain in full compliance with CDISC standards and regulatory submission expectations.

INTRODUCTION

The clinical research industry has been using the WHODrug dictionary for many years to standardize drug names for the analysis of concomitant medications. As many regulatory authorities refer to these guidelines, it became more important for UMC to keep its products updated so that WHODrug users will always be compliant and can use the dictionary without any workarounds.

While the present-day WHODrug dictionary is represented in what is called the B3 and C3 formats (further explained later in this paper), pre-2017 versions of the dictionary provided terms in what was labelled the B2 and C formats. Please refer to the technical guide "How to use WHODrug for compliance with CM domain in the CDISC SDTM standard" to understand how to retrieve the concomitant medication data from the WHODrug dictionary. This paper will not discuss the retrieval process of concomitant medications from the drug dictionary, instead it will list steps to be taken once the concomitant data is retrieved from the dictionary and how changes from B2/C to B3/C3 can be incorporated in SDTM CM domain per CDISC standards.

BACKGROUND

The guidance provided by the Uppsala Monitoring Centre (UMC) is based on CDISC SDTM v1.4 and SDTM-IG v3.2. Based on the structure of the CM domain per SDTM-IG v3.2, variables of interest are CMDECOD (Standardized Medication Name), CMCLAS (Medication Class) and CMCLASCD (Medication Class Code) and would need updating every time the dictionary is up versioned. The old B2 and C formats led to length constraints due to a limit of 45 characters for CMDECOD and 50 for CMCLAS variables, as a result the full text was not displayed and needed a workaround for both these variables. Hence, two of the most awaited changes that were incorporated in the B3 and C3 formats to solve these workarounds are,

1. the preferred base level is connected to salted and unsalted (generic) versions of drugs in combination and the character length limit was increased to 1,500 characters as the '/'s' were replaced with active ingredients in brackets for all non-unique trade names as shown in Table 1 below, and
2. the active ingredients are displayed next to the trade name, instead of drug code as was in the B2 format.

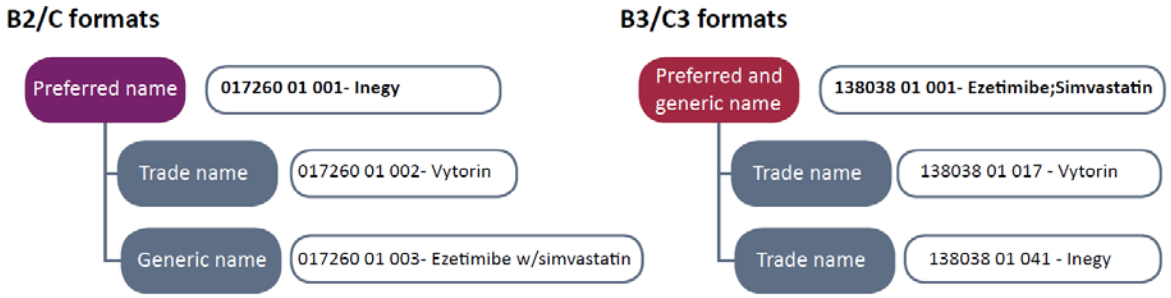
OLD VERSUS NEW

Let us understand the B and C formats. The B-format has information about trade names, ingredients, and ATC classification(s) where the alphanumeric drug code is the unique key. The C-format consists of all B-format information (including the Drug Code), with additional information regarding the countries in which the product is marketed, Marketing Authorization Holders, pharmaceutical forms and strengths; here, the unique key is the alphanumeric Medicinal Product ID. (WHODrug B3 and C3 format, 2018). Throughout this paper we will focus on how the change from the old B2/C to the new B3/C3 formats plays out in terms of how information is captured in two key SDTM CM domain variables, CMDECOD and CMCLAS.

DISPLAY OF SINGLE AND MULTI-INGREDIENTS FOR PREFERRED NAMES AND ATC ASSIGNMENTS

Assignment of preferred names for single and multi-ingredient medications

Medications are made up of either a single ingredient or multiple ingredients. These ingredients were displayed differently in B2/C format as compared to B3/C3 format. For single-ingredient records in both the B2/C and the B3/C3 formats, the preferred name is the same as the generic name, which is the active substance in the medication. However, the formats display differently for multi-ingredient records: the B2/C format displays the first compound with that unique formulation submitted to the UMC to add to the dictionary (often a trade name), whereas the B3/C3 format always displays the generic name. A side-by-side example is provided below:

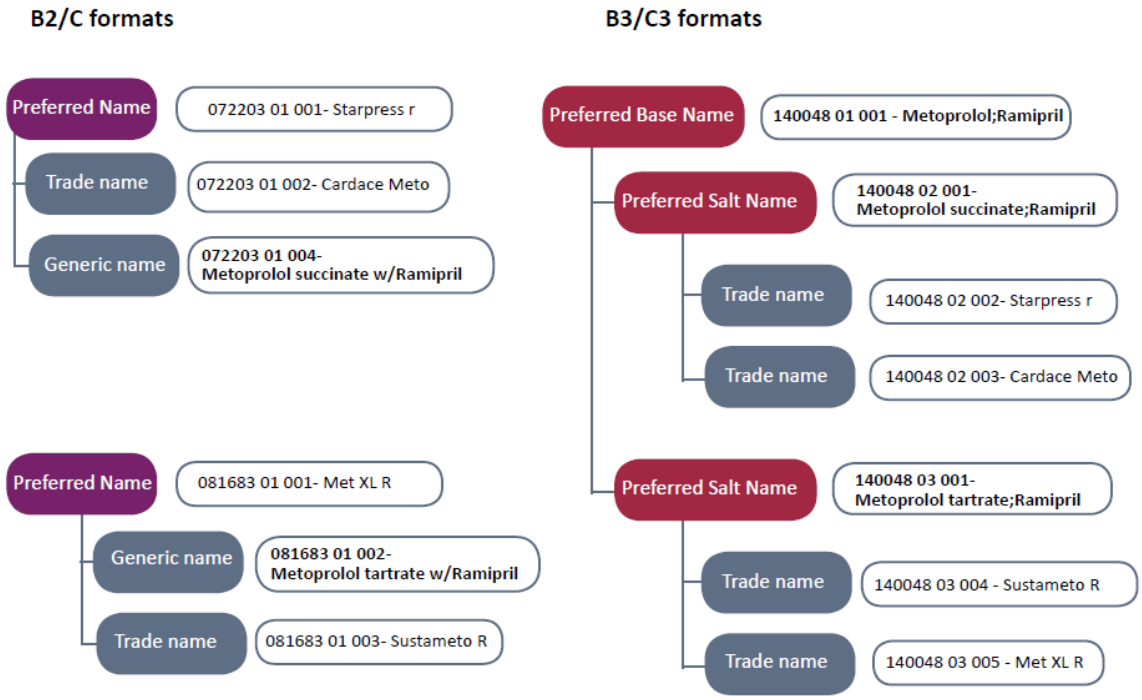


Display 1. (WHODrug B3- and C3- format, 2018), the preferred name in B2/C formats for a multi-ingredient record is a trade name whereas for B3/C3 formats it is a generic name

Assignment of preferred names for single and multi-ingredient salt/base relationships

In B2/C format, all salt variations of an ingredient are connected to the Preferred Base Name for single-ingredient products, whereas for multi-ingredient products, different forms of salts of an ingredient have a separate preferred name and they are not connected as their drug code is different as well.

This is different in the B3/C3 format. All variations of salts for single- as well as multi-ingredient products are connected to the same Preferred Base Name with the same drug code.



Display 2. (WHODrug B3- and C3- format, 2018), the salt variations of metoprolol along with Ramipril have separate preferred names and are not connected to one Preferred Base Name in the B2/C format, whereas in the B3/C3 format, all salt variations of metoprolol are connected to the same Preferred Base Name along with unsalted substance combinations

Display of multi-ingredient medications after ATC assignments

In the B2/C format, as different salt variations of a substance are connected to separate preferred names and drug codes for multi-ingredient products, they do not have the same ATC codes assigned. In contrast, in the B3/C3 format, all salt variations of a substance are connected to the same Preferred Base Name for multi-ingredient products, thus the same ATC codes are assigned for these records.

Display of non-unique names

Remember those odd numbers that were displayed after the preferred medication name captured in the CMDECOD variable in SDTM CM datasets? Those are nothing but /DRECNO + SEQ1/ (drug record number + sequence) and they were added in the B2 format to make the record unique. To aid in the efficient review of the active ingredient for each trade name, the latest update in B3 format replaces that number with [ACTIVE INGREDIENT(S)]. If there is more than one active ingredient, they will be displayed in alphabetical order and separated by semicolons.

B2 format	B3 format
ZYFLOX /00668101/	ZYFLOX [Norfloxacin]

Table 1. Drug displayed in B2 format as trade name along with code, and in B3 format as trade name along with generic name

Display of multi-ingredient generic names

In the B2/C format, the generic names were displayed by appending 'w/' followed by the substance name, whereas in the B3/C3 format, the generic name is separated by a semicolon (;). If multiple substance names

are present, then they are displayed in alphabetical order.

SDTM IMPLEMENTATION

Once the concomitant medication data is retrieved from the WHODrug dictionary in B3/C3 format, the CMDECOD variable length has been increased to 1,500 characters and to 110 characters for the CMCLAS variable. The reason for this increased length is to incorporate the full generic name/ATC text which can help the review process. Currently only 3% of records in WHODrug exceed the 200-character limit imposed by the SAS® V5 transport file format that is required for electronic data set submission to health authorities, with some records extending out to 1,200 characters. In order to still include the full coded term in such cases, we can use the SDTM supplemental CM dataset (SUPPCM). Here, the text string of the generic name must be extracted after the semicolon closest to 200 characters, and the rest of the coded term text can be kept in SUPPCM. Special care should be taken to avoid splitting between words.

USUBJID	CMSEQ	CMTRT	CMDECOD	CMCLAS	CMCLASCD
XYZ-1001	1	FUSION PLUS	ASCORBIC ACID;BIOTIN;CYANOCOBALAMIN;FERROUS FUMARATE;FOLIC ACID;LACTOBACILLUS CASEI;NICOTINIC ACID;PANTOTHENIC ACID;POLYSACCHARIDE-IRON COMPLEX;PYRIDOXINE HYDROCHLORIDE;RIBOFLAVIN;THIAMINE	ALIMENTARY TRACT AND METABOLISM	...

Table 2. SDTM CM dataset where CMDECOD contains generic names separated by semicolons, displayed up to 200 characters with the remainder of the text placed in QVAL of the SUPPCM dataset shown below

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
XYZ	CM	XYZ-1001	CMSEQ	1	CMDECOD1	Standardized Medication Name 1	HYDROCHLORIDE	...	...

Table 3. Supplemental dataset for CM where CMDECOD1 is the extended variable of CMDECOD which stores the portion of the string that exceeds 200 characters. The length of QVAL is also 200 characters, hence multiple CMDECODx can be generated where x is any number.

Display in concomitant medication tables (the final look)

B3/C3 formats help with the review of concomitant medications, especially for multi-ingredient products as we can see the full generic names for drug combinations. The following images will show how these changes in formats are reflected at the table level.

Table xx.xx.xx Concomitant Medications
(Population) Analysis Set

WHO Drug class Preferred Term	Cohort 1 (N=xx) n (%)	Cohort 2 (N=xx) n (%)	Total (N=xx) n (%)
Alimentary tract and metabolism			
Fusion Plus	xx(xx)	xx(xx)	xx (xx)

Page 1 of 1

Dictionary:WHODrug Global B3 201903

Display 3. Table display using a B2/C format

Table xx.xx.xx Concomitant Medications
(Population) Analysis Set

WHO Drug class Preferred Term	Cohort 1 (N=xx) n (%)	Cohort 2 (N=xx) n (%)	Total (N=xx) n (%)
Alimentary tract and metabolism			
Ascorbic acid;biotin;cyanocobalamin;ferrous fumarate;folic acid;lactobacillus casei;nicotinic acid;pantothenic acid;polysaccharide-iron complex;pyridoxine hydrochloride;riboflavin;thiamine hydrochloride	xx(xx)	xx(xx)	xx (xx)

Page 1 of 1

Dictionary:WHODrug Global B3 201903

Display 4. The same table after WHODrug is updated to B3/C3 format

CONCLUSION

The U.S. Food and Drug Administration (FDA) recognized this recent WHODrug update to B3/C3 format by publishing a notice in the Federal Register in October 2017 showing their support and requiring studies that started after March 15, 2019, to use this updated version. It is also mentioned in the U.S. FDA Data Standards Catalog. We encourage everyone to confirm with their study teams if the WHODrug Dictionary is updated and make necessary changes as suggested in this paper to the SDTM CM domain so that it is compliant with regulatory requirements.

REFERENCES

CDISC Study Data Tabulation Model (SDTM) v1.4, Study Data Tabulation Model Implementation Guide (SDTM-IG) v3.2: https://www.cdisc.org/standards/foundational/sdtm

WHODrug B3- and C3- Formats – Implementation Guide, How to use WHODrug for compliance with CM domain in the CDISC SDTM standard: https://www.who-umc.org/whodrug/b3c3-formats-upversioning/b3c3-documents/

U.S. FDA Data Standards Catalog: https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources

Notice in the Federal Register: https://www.govinfo.gov/content/pkg/FR-2017-10-24/pdf/2017-23029.pdf

ACKNOWLEDGMENTS

We would like to acknowledge our manager Shefalica Chand and John Shaik for inspiring us to contribute our learnings towards PharmaSUG proceedings and for their constant support. We would like to thank our clinical data coding specialist Hank Dennis for providing valuable comments on this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Contact the authors at:

Lyma Faroz
Seattle Genetics, Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
lfaroz@seagen.com

Jinit Mistry
Seattle Genetics, Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
jmistry@seagen.com