

---

# Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry

**U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research (CDER)  
Center for Biologics Evaluation and Research (CBER)**

**December 2023**

**Real-World Data/Real-World Evidence (RWD/RWE)**

---

# Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry

*Additional copies are available from:  
Office of Communications, Division of Drug Information  
Center for Drug Evaluation and Research  
Food and Drug Administration  
10001 New Hampshire Ave., Hillandale Bldg., 4<sup>th</sup> Floor  
Silver Spring, MD 20993-0002  
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353  
Email: [druginfo@fda.hhs.gov](mailto:druginfo@fda.hhs.gov)*

<https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs>

*and/or*

*Office of Communication, Outreach and Development  
Center for Biologics Evaluation and Research  
Food and Drug Administration  
10903 New Hampshire Ave., Bldg. 71, Room 3128  
Silver Spring, MD 20993-0002  
Phone: 800-835-4709 or 240-402-8010  
Email: [ocod@fda.hhs.gov](mailto:ocod@fda.hhs.gov)*

<https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances>

**U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research (CDER)  
Center for Biologics Evaluation and Research (CBER)**

**December 2023**

**Real-World Data/Real-World Evidence (RWD/RWE)**

**TABLE OF CONTENTS**

<b>I.</b>	<b>INTRODUCTION AND SCOPE.....</b>	<b>1</b>
<b>II.</b>	<b>REGULATORY BACKGROUND.....</b>	<b>2</b>
<b>III.</b>	<b>APPLYING CURRENTLY SUPPORTED DATA STANDARDS TO STUDY DATA DERIVED FROM REAL-WORLD DATA SOURCES.....</b>	<b>3</b>
	<b>A. Challenges in Real-World Data Standardization.....</b>	<b>3</b>
	<b>B. Documentation of Processes for Managing Real-World Data .....</b>	<b>4</b>
	<b>C. Considerations for Conforming Real-World Data to Currently Supported FDA Study Data Standards .....</b>	<b>4</b>
	<b>D. Considerations for Mapping Real-World Data To Study Data Submission Standards .....</b>	<b>5</b>
	<b>E. Considerations for Data Transformations.....</b>	<b>6</b>
	<b>APPENDIX: EXAMPLES OF MAPPING HEALTH CARE DATA TO CDISC SDTM ....</b>	<b>8</b>

# **Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry<sup>1</sup>**

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

## **I. INTRODUCTION AND SCOPE**

### **21st Century Cures Act and Real-World Data**

The 21<sup>st</sup> Century Cures Act,<sup>2</sup> signed into law on December 13, 2016, is intended to accelerate medical product development and bring innovations faster and more efficiently to the patients who need them. Among other provisions, the 21<sup>st</sup> Century Cures Act added section 505F to the Federal Food, Drug, and Cosmetic Act (FD&C Act) (21 U.S.C. 355g). Pursuant to this action, calling for FDA to issue guidance on the use of real-world evidence (RWE) in regulatory decision-making, FDA has created a framework for a program to evaluate the potential use of real-world data (RWD) to generate RWE to help support the approval of new indication(s) for drugs<sup>3</sup> already approved under section 505(c) of the FD&C Act (21 U.S.C. 355(c)) or to help support or satisfy post-approval study requirements (RWE Program).<sup>4</sup>

This guidance provides recommendations to sponsors for complying with section 745A(a) of the FD&C Act (21 U.S.C. 379k-1(a)) when submitting RWD as study data in applicable drug submissions. FDA is issuing this guidance as part of its RWE Program and to satisfy, in part, the mandate under section 505F of the FD&C Act (21 U.S.C. 355g) to issue guidance on the use of RWE in regulatory decision-making.<sup>5</sup>

---

<sup>1</sup> This guidance has been prepared by the Center for Drug Evaluation and Research (CDER) in cooperation with the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

<sup>2</sup> Public Law 114-255.

<sup>3</sup> For the purposes of this guidance, all references to *drugs* include both human drugs and biological products.

<sup>4</sup> See the document *Framework for FDA's Real-World Evidence Program* (December 2018), available at <https://www.fda.gov/media/120060/download>. The framework and RWE Program also cover biological products licensed under the Public Health Service Act.

<sup>5</sup> See section 505F(e) of the FD&C Act.

## *Contains Nonbinding Recommendations*

This guidance addresses considerations for the use of data standards<sup>6</sup> currently supported by FDA in applicable drug submissions containing study data<sup>7</sup> derived from RWD sources. For the purposes of this guidance, FDA defines *RWD* as data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources. Examples of RWD include data from electronic health records (EHRs); medical claims data; data from product and disease registries;<sup>8</sup> patient-generated data (including data from in-home-use settings); and data gathered from other sources that can inform on health status, such as mobile devices.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in FDA guidance means that something is suggested or recommended, but not required.

## **II. REGULATORY BACKGROUND**

Under section 745A(a) of the FD&C Act, at least 24 months after the issuance of a final guidance document in which FDA has specified the electronic format for submitting certain submission types to the Agency, such content must be submitted electronically and in the format specified by FDA.<sup>9</sup> The guidance for industry *Providing Regulatory Submissions in Electronic Format—Standardized Study Data* (Study Data Guidance)<sup>10</sup> and the technical specifications referenced therein describe electronic submission requirements under section 745A(a) of the FD&C Act for clinical and nonclinical study data contained in new drug applications (NDAs), abbreviated new drug applications (ANDAs), certain biologics license applications (BLAs), and certain investigational new drug applications (INDs) submitted to CDER or CBER.<sup>11</sup> Given that these electronic submission requirements apply to study data submitted in the covered application types, they apply to RWD that is submitted as study data in such applications. That is, RWD submitted as study data to NDAs, ANDAs, certain BLAs, and certain INDs, as further described in section II.A of the Study Data Guidance, must be in an electronic format that the

---

<sup>6</sup> For the purposes of this document, the term *data standards* means defined rules, conventions, guidelines, characteristics, methods, formats, and terminologies that provide structure and consistency for exchange and utilization of data (see <https://www.cdisc.org/standards/glossary>).

<sup>7</sup> See FDA's guidance for industry *Providing Regulatory Submissions in Electronic Format—Standardized Study Data* (June 2021). We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at <https://www.fda.gov/regulatory-information/search-fda-guidance-documents>. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-regulatory-submissions-electronic-format-standardized-study-data>

<sup>8</sup> For the purposes of this guidance, the term *registries* means organized systems that collect clinical or other data in standardized formats for populations defined by a particular disease, condition, or exposure.

<sup>9</sup> For additional information on how FDA interprets and intends to implement the electronic submission requirements of section 745A(a) of the FD&C Act, see the guidance for industry *Providing Regulatory Submissions in Electronic Format—Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act* (December 2014).

<sup>10</sup> For additional information, see the guidance for industry *Providing Regulatory Submissions in Electronic Format—Standardized Study Data* (June 2021).

<sup>11</sup> See section II of the Study Data Guidance for more information on the types of submissions subject to electronic submission requirements for standardized study data and what submissions are exempt from such requirements.

## *Contains Nonbinding Recommendations*

Agency can process, review, and archive, unless such submission is exempt from the electronic submission requirements or if FDA has granted a waiver.<sup>12</sup> Currently, as stated in the Study Data Guidance, the Agency can process, review, and archive electronic submissions of clinical and nonclinical study data (including those derived from RWD sources) that use the standards specified in the Data Standards Catalog (Catalog).<sup>13</sup> As that guidance explains, the Catalog provides a listing of currently supported<sup>14</sup> and/or required standards, their uses, the date FDA will begin (or has begun) to support a particular standard, the date such support ends (or will end), the date the requirement to use a particular standard will begin (or has begun), the date such requirement ends (or will end), and other pertinent information.

FDA is issuing this guidance to provide recommendations to sponsors for complying with section 745A(a) of the FD&C Act using standards specified in the Catalog when submitting study data derived from RWD sources in applicable drug submissions.

### **III. APPLYING CURRENTLY SUPPORTED DATA STANDARDS TO STUDY DATA DERIVED FROM REAL-WORLD DATA SOURCES**

#### **A. Challenges in Real-World Data Standardization**

FDA recognizes the challenges involved in standardizing study data derived from RWD sources for inclusion in applicable drug submissions. These challenges include but are not limited to:

- The variety of RWD sources and their inconsistent formats (e.g., EHR, registry)
- The potential use of more than one type of RWD source in a study (such as combining EHR and claims data)
- The differences in source data captured regionally and globally using different standards, terminologies,<sup>15</sup> and exchange formats<sup>16</sup> for the representation of the same or similar data elements<sup>17</sup>
- Certain information only existing in non-structured documentation (such as text in physician notes)

---

<sup>12</sup> Sponsors or applicants may apply for a waiver from the requirement to use specific versions of FDA-supported standards for the submission of study data using the waiver request process described in section II.D of the Study Data Guidance.

<sup>13</sup> The Catalog is available at <http://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm>.

<sup>14</sup> For the purposes of this document, the term *supported* means the receiving Center has established processes and technology to support receiving, processing, reviewing, and archiving files in the specified standard.

<sup>15</sup> For the purposes of this document, the term *terminologies* means the body of terms used for particular technical application to standardize a medical term for the submission of nonclinical and clinical study data.

<sup>16</sup> For the purposes of this document, the term *exchange format* means a data format for converting from one file or database structure to another. For example, XML is commonly used as a data exchange format.

<sup>17</sup> See data elements at [https://csrc.nist.gov/glossary/term/data\\_element](https://csrc.nist.gov/glossary/term/data_element).

## *Contains Nonbinding Recommendations*

- A wide range of methods and algorithms which could be used to create datasets intended to aggregate data
- The many aspects of health care data that can affect the overall quality of the data, including business processes and database structure, inconsistent vocabularies and coding systems, and de-identification methodologies used to protect patient data when shared
- The various levels of access a sponsor has to any data sources used in the study

Approaches to these challenges will depend on the specific data being used and the research being conducted. For this reason, sponsors should discuss such issues with the appropriate FDA review division(s) as early as possible,<sup>18</sup> preferably before the sponsor conducts the study.

### **B. Documentation of Processes for Managing Real-World Data**

During data curation<sup>19</sup> and data transformation,<sup>20</sup> adequate processes should be in place to ensure confidence in the resultant data. Documentation of these processes may include but is not limited to electronic documentation (e.g., audit trails, quality control procedures, etc.) of data additions, deletions, or alterations from the source data system to the final study analytic data set(s). Sponsors should also document in their applicable drug submission changes to data to conform to the current FDA-supported data standards, and the potential impacts of these changes. For more information, see the *Study Data Technical Conformance Guide*<sup>21</sup> (TCG), which supplements the Study Data Guidance and standards listed in the Catalog, to provide specifications, recommendations, and general considerations on how to submit standardized study data and documentation using FDA-supported data standards located in the Catalog.

### **C. Considerations for Conforming Real-World Data to Currently Supported FDA Study Data Standards**

The FDA Data Standards Catalog is updated periodically to reflect changes to the supported and required data standards, exchange formats, and terminologies that FDA can process, review, and archive. FDA plans to update the Catalog, and/or issue other guidance documents, to reflect standards for study data that are derived from RWD sources as they are developed and evaluated. Currently, and absent a waiver, sponsors submitting clinical and nonclinical study data (including those derived from RWD sources) in submissions subject to section 745A(a) of the FD&C Act are required to use the formats described in the Study Data Guidance and the supported study data standards listed in the Catalog. Sponsors should refer to the specifications, recommendations, and general considerations provided in the TCG when submitting study data

---

<sup>18</sup> See the draft guidance for industry *Formal Meetings Between the FDA and Sponsors or Applicants of PDUFA Products* (December 2017) available at the FDA guidance web page. When final, this guidance will represent FDA's current thinking on this topic.

<sup>19</sup> For the purposes of this document, the term *data curation* means processing of source data through the application of standards for exchange, integration, sharing, and retrieval.

<sup>20</sup> For the purposes of this document, the term *data transformation* means the process of converting data from one format or structure into another format or structure.

<sup>21</sup> The *Study Data Technical Conformance Guide* is available at <https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources>.

## *Contains Nonbinding Recommendations*

in an applicable submission to FDA. When seeking to conform RWD to data standards supported by FDA, sponsors should consider the relevant data transformations, conversions, or mappings<sup>22</sup> that may be needed to produce study datasets in required formats in an applicable drug submission.

Sponsors should discuss, as early as possible, with the appropriate FDA review division, any planned submission of study data derived from RWD sources in an applicable submission and their approaches to mapping and transformation of data. Such discussions should include approaches for taking data from its original RWD format to intermediary format(s), including those used for analysis, and ultimately to the study data set prepared for submission to FDA (using the applicable FDA-supported data standards as listed in the Catalog). Sponsors should describe these approaches, including in the protocol, data management plan, and/or final study reports.

FDA recognizes that a range of approaches may be used to apply the supported study data standards (e.g., Clinical Data Interchange Standards Consortium's (CDISC's) Study Data Tabulation Model (SDTM) or Analysis Data Model (ADaM)) to RWD sources such as EHR or claims data. FDA encourages sponsors to discuss such approaches with FDA.<sup>23</sup> With adequate documentation of the conformance methods used and their rationale, study data derived from RWD can be transformed into SDTM and ADaM datasets and submitted to FDA in an applicable submission.

### **D. Considerations for Mapping Real-World Data To Study Data Submission Standards**

FDA is aware that, for nearly every data domain,<sup>24</sup> there is wide divergence in the terminologies used and their precise meaning between RWD sources and FDA-supported/required data standards. Examples range from the meaning and specific terms used for race/ethnicity, terminology systems for medications, and interpretation of the measurements of vital signs from health care records. Even for seemingly similar variables, there can be differences in the way these variables are defined between RWD sources and FDA-supported data standards. For example, sex as a variable may be codified in CDISC's terminology as a concept based on physical characteristics, whereas EHRs may, in some cases, use gender identity. In such cases, sponsors should document the potential impact of mapping the sex variable or other variables to CDISC's terminology on the study findings.

The sponsor should select an appropriate mapping approach that best fits the characteristics of the data and the nature of the study. Sponsors should explain the rationale for the chosen method

---

<sup>22</sup> For the purposes of this document, the term *mapping* means the process of defining data element relationships between two distinct data models.

<sup>23</sup> Questions regarding data standards for electronic study data can be submitted to [CDER-edata@fda.hhs.gov](mailto:CDER-edata@fda.hhs.gov) or [CBER-edata@fda.hhs.gov](mailto:CBER-edata@fda.hhs.gov) for submissions to CDER or CBER, respectively.

<sup>24</sup> For the purposes of this document, the term *data domain* means "a collection of logically related observations (with a common, specific topic) that are normally collected for all subjects in a clinical investigation. NOTE: The logic of the relationship may pertain to the scientific subject matter of the data or to its role in the trial/study. Example domains include laboratory test results, adverse events, concomitant medications" (<https://www.cdisc.org/kb/articles/domain-vs-dataset-whats-difference>).



## ***Contains Nonbinding Recommendations***

in the Study Data Reviewer's Guide.<sup>25</sup> Additionally, documentation of the sponsor's rationale for choosing particular data elements from currently supported study data standards to represent study data derived from RWD sources is critical. The sponsor should provide a description of the general approach and anticipated impact of data mapping as a part of to the Study Data Reviewer's Guide to highlight the domains involved. Furthermore, the sponsor should include a data dictionary that documents the definition of every data element derived from RWD sources that were mapped to the study data standard and all relevant information about the element, such as its relationships to other data, origin, usage, and format. The Study Data Reviewer's Guide can also be used to guide reviewers to detailed mappings which can be placed in the Define-XML file<sup>26</sup> and data domain files (see the Appendix for illustrative examples of possible ways to use Define-XML for this purpose). At any time, sponsors can consult the most recent version of the TCG to see FDA's current thinking on the use of Define-XML and the Study Data Reviewer's Guide.

### **E. Considerations for Data Transformations**

Sponsors may encounter challenges when transforming RWD into data that are consistent with FDA-supported data standards. Examples of these challenges include (but are not limited to) management of semantic concepts (terms) that are present at multiple locations in a health record (such as medication information), inconsistent coding or miscoding of concepts (e.g., drugs or diagnoses), changes in data collection or coding practices (e.g., International Classification of Diseases-9 (ICD-9) and ICD-10 codes) that occurred during the study, or missing information (either because information is not typically recorded in health care settings or due to inconsistent data entry).

Sponsors should document data challenges encountered during transformation to an FDA-supported data standard and a justification of their approach to enable the application of an FDA-supported data standard. Detailed representation of the mapping of standards and terminologies used when transforming data to current study data standards can be handled using the Define-XML (see the Appendix) and domain data files. In addition, for more complex discussions, narrative descriptions should also be presented in the Study Data Reviewer's Guide, either in the body or as an appendix, with appropriate directions for reviewers to the Define-XML and dataset/domains for more detail, if needed.

---

<sup>25</sup> For the purposes of this document a *Study Data Reviewer's Guide* means a document provided by sponsors to FDA to describe any special considerations or directions or conformance issues that may facilitate an FDA reviewer's use of the submitted data and may help the reviewer understand the relationships between the study report and the data. For more information, see the Study Data Technical Conformance Guide, available at <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/study-data-technical-conformance-guide-technical-specifications-document>.

<sup>26</sup> See description of Define-XML file at <https://www.cdisc.org/standards/data-exchange/define-xml>.

## *Contains Nonbinding Recommendations*

### **APPENDIX: EXAMPLES OF MAPPING HEALTH CARE DATA TO CDISC SDTM**

Differences in the coding systems used between real-world data (RWD) and traditional clinical trial data can usually be addressed using the Define-XML file, which is included in all standard Study Data Tabulation Model (SDTM) submissions. The Define-XML file, along with the appropriate use of *Alias* data element, provides one mechanism for communicating the transformation of external coding systems to the appropriate SDTM controlled terminology.<sup>27</sup>

An example of this approach involves race/ethnicity data, where the Food and Drug administration (FDA) anticipates both heterogeneity among electronic health records (EHRs) as well as between EHR and Clinical Data Interchange Standards Consortium (CDISC) terminologies. In the guidance for industry *Collection of Race and Ethnicity Data in Clinical Trials* (October 2016), FDA recommends that at least five specific categories be used to define race:

- (1) American Indian or Alaska Native
- (2) Asian
- (3) Black or African American
- (4) Native Hawaiian or Other Pacific Islander
- (5) White

RWD sources, however, may not follow the same system of coding. Given that FDA recommends using the race and ethnicity categorization outlined in the *Collection of Race and Ethnicity Data in Clinical Trials* guidance mentioned above, a sponsor should map the RWD terminology system to the relevant SDTM terminology. To achieve this objective, the *Alias* element in Define-XML file can be used to document the conversions to a single nomenclature while ensuring traceability.<sup>28</sup>

Table 1 illustrates how race can be transformed from non-standardized to standardized data using FDA-supported data standards. In Table 1, the *Alias* column shows the original codes present in an EHR system and the *Code* column shows the relevant mapped term in the current FDA-supported CDISC controlled terminology:

---

<sup>27</sup> For the purposes of this document, a *controlled terminology* means a “finite set of values that represent the only allowed values for a data item. These values may be codes, text, or numeric.” (CDISC Glossary at <https://www.cdisc.org/standards/glossary>.)

<sup>28</sup> For the purposes of this document, traceability means documentation of curation and transformation procedures used throughout the data life cycle.

## Contains Nonbinding Recommendations

**Table 1: Approach To Using Define-XML To Indicate Decision Involved in Transforming Non-Standardized Data (Race Data) To Standardized Data (i.e., SDTM and ADaM)**

*Illustrative example of an approach representing cross-mapping of coding systems, in this case for Race data, to CDISC coding in the Define-XML file. Note that this table does not recommend **how** to map coding systems to CDISC terminology. Additionally, this table demonstrates only one possible approach out of many for representing the mapping choices made.*

Permitted Value (Code)*	Display Value (Alias)**
Race [RACE, C74457]	
AMERICAN INDIAN OR ALASKA NATIVE [C41259]	American Indian or Alaska Native, Native American, Native of Alaska; Context: EHR Mapping
ASIAN [C41260]	Asian, Chinese; Context: EHR Mapping
BLACK OR AFRICAN AMERICAN [C16352]	Black or African American, Black; Context: EHR Mapping
NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER [C41219]	Native Hawaiian or Other Pacific Islander, Samoan; Context: EHR Mapping
OTHER [*]	Other; Context: EHR Mapping
WHITE [C41261]	White, Mexican; Context: EHR Mapping

\*Permitted Value (Code): vocabulary that is provided in the study data tabulations and conformant with CDISC controlled terminologies.

\*\*Display Value (Alias): vocabulary that was used in the original data set (i.e., EHR value). Code/Alias: Respective CDISC elements.

Differences in controlled terminology between RWD systems and FDA submission data standards may make mapping terminology challenging. Furthermore, FDA is aware that in non-interventional studies,<sup>29</sup> a sponsor may use data aggregated from multiple RWD systems. Such situations can complicate the use of terminologies further, since different RWD sources, such as EHRs, might use different coding systems for the same concept or might use the same coding system but use different default codes for the same item.

Various approaches can be applied to permit the use of RWD in applicable drug submissions. Examples of potential approaches are: (1) translating the codes to their mapped structured definitions with subsequent mapping to appropriate CDISC controlled terminologies, which provides the most detail but is labor-intensive; or (2) alternatively mapping all original codes to the least granular analogous codes, and then mapping those to CDISC controlled terminologies, which is less labor-intensive yet necessitates that detail of a more specific categorization will not be represented in the submitted, standardized dataset. It is up to the sponsors to determine the best approach to mediating data transformation, as well as to document and justify their approach accordingly. However, if details that are essential to the consideration of the safety and effectiveness of a drug are absent, the latter approach may not be appropriate. Whatever approach is used, the application of *Alias* to achieve CDISC standard controlled terminology is one mechanism to document the normalization of nomenclature into a format developed by CDISC.

An example where concepts and terminology do not map precisely and directly is the SDTM intervention domain capturing *drugs prescribed*. Domains containing *drugs prescribed* data may be mapped to a template SDTM intervention domain. (FDA anticipates that an EHR system may

<sup>29</sup> For the purposes of this document, a *non-interventional* study is a type of study in which patients are not assigned to a study arm according to a protocol, but instead receive the drug of interest during routine clinical care.

## *Contains Nonbinding Recommendations*

use prescription coding from RxNorm,<sup>30</sup> FDA’s National Drug Code (NDC),<sup>31</sup> or other such systems.) Additionally, and unless a sponsor uses EHR data where the prescription dispensing information is retained or opts to link EHR information with medical claims data, uncertainty will persist regarding the actual prescription (e.g., whether a generic pharmaceutical agent was substituted). In such cases, sponsors may apply the *Alias* element within Define-XML to normalize the coding systems (as illustrated in Table 2), where the final dataset (see Table 3) will have prescription dosing information uniformly reflecting the less specific prescription coding system.

The standard modeling would be accompanied by the Define-XML code list as follows:

**Table 2: Approach To Using Define-XML To Indicate Decision Involved in Transforming Non-Standardized Data (Drugs Prescribed) to Standardized Data (i.e., SDTM and ADaM)**

*Illustrative example of an approach to representing cross-mapping of coding systems, in this case for prescription data, to CDISC coding in the Define-XML file. Note that this table does not recommend **how** to map coding systems to CDISC terminology. Additionally, this table demonstrates only one possible approach out of many for representing the mapping choices made.*

Permitted Value (Code)*	Display Value (Alias)**
FLUoxetine	FLUoxetine 40 milligram (mg) Oral Capsule, RxCUI = 383919, 0093-7198-56, FLUoxetine 40 mg Oral Capsule Generic Permitted; Context: EHR Mapping
PROzac	PROzac 40 mg Oral Capsule, RxCUI = 313989, 0777-3107-30; Context: EHR Mapping

\*Permitted Value (Code): vocabulary that is provided in the study data tabulation.

\*\*Display Value (Alias): vocabulary that was used in the original data set (i.e., EHR value). Code/Alias: Respective CDISC elements.

The standard modeling of prescription data in the domain file would appear as shown below (Table 3):

**Table 3: Example of the Drugs Prescribed Data in the Respective Domain Upon Mapping**

*Illustrative example of an approach to representing the values from data, in this case for prescription data, to CDISC format in the relevant data domain file. Note that this table does not recommend **how** to map data to CDISC standards. Additionally, this table demonstrates only one possible approach out of many for representing the mapping choices made. The column headers represent CDISC data elements. For more information, see <https://www.cdisc.org/standards/foundational/sdtm>.*

Original Value (from original dataset)	--TRT	--MODIFY	--Alias	--DOSE	--DOSU	--ROUTE	--DOSFRM
FLUoxetine 40 mg Oral Capsule	FLUoxetine	To be populated by the sponsor to assist in coding to standard terminology	Generic Drug Name in WHO Drug <sup>32</sup> (either from original data system or assigned by medical coding)	40	mg	ORAL	CAPSULE

<sup>30</sup> See <https://www.nlm.nih.gov/research/umls/rxnorm/index.html>.

<sup>31</sup> See <https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory>.

<sup>32</sup> World Health Organization's Drug Dictionary, see <https://who-umc.org/whodrug/>.

***Contains Nonbinding Recommendations***

			vendor for sponsor)				
FLUoxetine 40 mg Oral Capsule Generic Permitted	FLUoxetine	(same as above)	(same as above)	40	mg	ORAL	CAPSULE
RxCUI = 383919	FLUoxetine	(same as above)	(same as above)				
0093-7198- 56	FLUoxetine	(same as above)	(same as above)				
PROzac 40 mg Oral Capsule	PROzac	(same as above)	(same as above)	40	mg	ORAL	CAPSULE
RxCUI = 313989	PROzac	(same as above)	(same as above)				
0777-3107- 30	PROzac	(same as above)	(same as above)				

Although only a few examples are presented here, sponsors should use elements of the Define-XML file and relevant domain data files to communicate how the health care terminology of all data domains were normalized to CDISC standard terminology. As in the examples shown above, the technical details of all transformations are best placed in the data files.