

## Tips and tricks when developing Trial Design Model Specifications that provide Reductions in Creation time

Ruth Marisol Rivera Barragan, Chiltern International, Inc. (Manor, TX)

### ABSTRACT

This presentation will give you some key points when developing trial design model specification under SDTM versions (up to 3.2). After the presentation, you will have a clearer idea of what inputs are needed to gather, as well as understand in advance of starting work on the trial design domains, in order to move efficiently through the definition process.

It is important to have: the latest protocol; a summary of the Inclusion/Exclusion criteria changes; the SDTM Specification Template; the latest version of Electronic Case Report Form (eCRF) as well as the specifications for DM, IE, SE and SV domains before beginning to create the Trial Design domains. The presentation will also recognize the main differences within SDTM IG version 3.1.3 and older versions for the Trial Design Model (TDM) domains.

A discussion will present the critical cross checking between TDMs and other SDTM domains that is necessary to consider for the most efficient specification development, for example:

1. TA (Trial Arms) and DM (Demographics) domain.
2. TA (Trial Arms) and TE (Trial Element).
3. TE (Trial Element) and SE (Subject Elements) domain.
4. TI (Trial Inclusion/Exclusion) and IE (Inclusion/Exclusion) domain.
5. TV (Trial Visits) and SV (Subject Visits) domain.

Finally, the presentation will give the audience useful and specific examples on how to find the required information on external references shown in the SDTM IG that may not be as straightforward as it initially appears when working with Trial Summary domain in version 3.1.3 (like REGID, UNII, SNOMED CT, NDF-RT, etc.)

### INTRODUCTION

A Trial Design Model (TDM) domain is a special purpose data set, which represent information about the study design but do not contain subject data. There are six TDM domains that are well defined on the SDTM Implementation Guide version 3.2, they are:

1. Trial Arms (TA), describes each planned arm in the trial.
2. Trial Elements (TE) contains the definitions of the elements that appear in the TA dataset.
3. Trial Visits (TV), describes the planned Visits in a trial.
4. Trial Summary (TS), allows the sponsor to submit a summary of the trial in a structured format.
5. Trial Inclusion Exclusion (TI) which contains all the inclusion and exclusion criteria for the trial.
6. Trial Disease Assessments (TD) domain provides information on the protocol-specified disease assessment schedule (this is present in SDTM IG version 3.2 but not in the previous, TD will not be referenced further in this paper).

The specifications for these datasets are different from the rest of the SDTM domains. If you have been assigned to develop TDM specifications for the first time and not sure where to start, this document will help you with that, giving you some useful tips and extra resources.

This presentation will give you some key points when developing trial design model specification under SDTM versions (up to 3.2). After the presentation, you will have a clearer idea of what inputs are needed to gather, as well as understand in advance of starting work on the trial design domains, in order to move efficiently through the definition process.

## INPUTS NEEDED TO DEVELOP TDM SPECS

1. Latest Protocol version.
2. A summary of the Inclusion/Exclusion criteria changes (if any) including all the wording updates and protocol version number where they were changed.
3. SDTM Implementation Guide (IG). It's very important to confirm first the version in which the study should be based.
4. SDTM Specification Template (based in the version mentioned above).
5. Latest Electronic Case Report Form (eCRF)
6. Updated specs for DM, IE, SE and SV domains (if available).

Note: TDM domains can be created either prior to or after the other SDTM domains. If prior, TDM domains such as TA, TE, TV, and TI will be the source of SDTM domain values. If TDM domains are created after the other SDTM domains, it will be the opposite way. Further information will be provided about these cross checks in this paper.

## WHERE TO START?

Once you have all the inputs needed, you can follow these steps:

1. Read the protocol synopsis.
2. Read the SDTM IG Section 7, please put special attention to the version you have to use, since there are substantial differences between versions. (Ask if unsure). Look carefully at the examples provided there, they are very useful to have a starting point.
3. Use the specifications template (usually available with the SDTM IG package) to start creating the TDM domains:
  - For creating Trial Arms (TA) data set specs, you will need to read the protocol and make a list or diagram (whatever works better for you) including all the 'Arms' within in the study and fill each row on the template specs with the information required for the corresponding 'Arm'.
  - Once you have the 'Arms' defined, you have to work on the Trial Elements (TE) data set, which will contain the definitions of the elements that appear in the Trial Arms (TA) dataset. Make sure to understand the difference between an 'Arm' and an 'Element', you can find these and some other helpful definitions at Section 7.1.2 on SDTM IG (Definitions of Trial Design Concepts).
  - To create the Trial Visits (TV) data set specifications, you will only to copy the visit values from SV (if available), otherwise, get them from the protocol in the 'Schedule' section.
  - The Trial Summary (TS) will require that each PARAMCD described on the SDTM IG appendix is filled with the study information present in the protocol. The paper will show some specific examples in the following section.
  - Since the Trial Inclusion Exclusion (TI) has to contain the inclusion and exclusion criteria for the trial, try to get them directly from IE (Inclusion/Exclusion) data set specifications, if not available, get them directly from the protocol.
4. Make sure that all the variables that are shown as Core = "Required" or "Expected" are present in the domain. You can find this information in the 'Appendix C –Controlled Terminology – Appendix C1: Trial Summary codes on the SDTM IG v 3.2.

## SPECIFIC EXAMPLES FOR GETTING TSVAl VALUES LISTED ON THE SDTM IG V3.1.2 AND UP.

In this section, a review of some specific examples on how to find the TSVAl values on external links (referenced in the SDTM IG) will be presented. It may not be as straightforward as it initially appears when working with Trial Summary domain in version 3.1.2 and up, such as the following:

### HOW TO GET TSVAl VALUES FOR TSPARMCD="REGID" (REGISTRY IDENTIFIER).

- Go directly to the CT.GOV website at [www.clinicaltrials.gov](http://www.clinicaltrials.gov)
- Then, on the left side put for the protocol number of your study in the 'Search for Studies' box and hit "Search" as shown in the example below.

The screenshot displays the ClinicalTrials.gov homepage. At the top left is the logo "ClinicalTrials.gov" with the tagline "A service of the U.S. National Institutes of Health". To the right is a descriptive paragraph: "ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world. Learn more about clinical studies and about this site, including relevant history, policies, and laws." Below this is a navigation bar with links: "Find Studies", "About Clinical Studies", "Submit Studies", "Resources", and "About This Site". A status line indicates "ClinicalTrials.gov currently lists 211,437 studies with locations in all 50 States and in 193 countries." The main content area features a "Search for Studies" section with an example search term "Heart attack" AND "Los Angeles", a text input field for "PROTOCOL #", and a "Search" button. To the right of the search box is a "Search Help" section with links: "How to search", "How to find results of studies", and "How to read a study record". Further right is a "Locations of Recruiting Studies" section with a pie chart showing the distribution: Non-U.S. only (54%), U.S. only (40%), and Both U.S. and non-U.S. (6%). Below the chart, it states "Total N = 37,706 studies (Data as of March 24, 2016)" and includes a link "See more trends, charts, and maps". At the bottom, there are three columns of links: "For Patients and Families" (How to find studies, See studies by topic, Learn about clinical studies, Learn more), "For Researchers" (How to submit studies, Download content for analysis, About the results database, Learn more), and "For Study Record Managers" (Why register?, How to register your study, FDAAA 801 requirements, Learn more). A "Learn More" section includes links for "Tutorials for using ClinicalTrials.gov", "Glossary of common site terms", "For the press", and "Using our RSS feeds".

Display 1. Looking TSVAl values for REGID - Step 1

- Then when you find your study, give a click on the title.

The screenshot shows the ClinicalTrials.gov search results page. At the top, there is a search bar with the example text "Heart attack" AND "Los Angeles". Below the search bar, there are navigation links: "Advanced Search", "Help", "Studies by Topic", and "Glossary". A blue navigation bar contains "Find Studies", "About Clinical Studies", "Submit Studies", "Resources", and "About This Site". The breadcrumb trail reads "Home > Find Studies > Search Results". A summary box indicates "180 studies found for: PROTOCOL #". Below this, there are tabs for "List", "By Topic", "On Map", and "Search Details". There are also options to "Show Display Options", "Download", and "Subscribe to RSS". A checkbox section allows filtering by "Include only open studies" and "Exclude studies with Unknown status". The main table has columns for Rank, Status, and Study. The first entry is Rank 1, Status "Completed", and Study "Adolescent Weight Management Study". Below the study title, it lists conditions and interventions.

Display 2. Looking TSVAl values for REGID - Step 2

- Finally, you will be directed to a window with all the information about the study; in this case, the TSVAl for TSPARMCD = "REGID" will be the "ClinicalTrials.gov Identifier" (NCT002173) as you can see below in Display 3 and the TSVCDVER will be "CT.GOV".

The screenshot shows the study record detail page for "Adolescent Weight Management Study". The breadcrumb trail is "Home > Find Studies > Search Results > Study Record Detail". A summary box indicates "Trial record 1 of 180 for: PROTOCOL #". Below this, there are links for "Previous Study", "Return to List", and "Next Study". The study title "Adolescent Weight Management Study" is prominently displayed. A red banner states "This study has been completed." Below this, the sponsor is listed as "National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)". The information provided by is also "National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)". On the right side, the "ClinicalTrials.gov Identifier" is listed as "NCT002173", along with dates for "First received", "Last updated", and "Last verified", and a link for "History of Changes". At the bottom, there are tabs for "Full Text View", "Tabular View", and "No Study Results Posted", along with links for "Disclaimer" and "How to Read a Study Record".

Display 3. Looking TSVAl values for REGID - Step 3

## HOW TO GET TSVVAL VALUES FOR TSPARMCD=TRT (INVESTIGATIONAL THERAPY OR TREATMENT)

- Go to the **U.S. Food and Drug Administration (FDA) Substance Registration System – Unique Ingredient Identifier (UNII)** website (<http://fdasis.nlm.nih.gov/srs>).
- Then, put the active ingredient of the treatment (usually mentioned in the protocol) and hit “Go” an example looking for “ASPIRIN” is shown in the Display 4 below.

The screenshot shows the search interface of the FDA Substance Registration System. At the top, there are logos for the FDA and the National Library of Medicine. The main heading is "Search Substance Registration System" with a sub-note "Information available for 76,532 substances." Below this, there are radio buttons for "By Name/Code" (selected) and "By UNII". A search box contains the text "ASPIRIN" and a "Go" button. Below the search box, there are three links: "Show examples (including wildcard searches using \*)", "UNII List download (updated Jan 21, 2016)", and "About Substance Registration System".

Display 4. Looking TSVVAL values for TRT - Step 1

- On the next window you will see that the TSVVAL for TSPARMCD=TRT (Investigational Therapy or Treatment) will be the UNII value shown in Display 5 below.

The screenshot shows the search results page for Aspirin. At the top, there is a breadcrumb "Home > Search Results" and a search box with "ASPIRIN" and a "Go" button. Below the search box, there is a link "Go back to previous page." A blue bar with "Search Results" is followed by the following information: "Preferred Substance Name: ASPIRIN" with a "Synonyms and Mappings" button; "UNII: R16CO5Y76E" with a "Search the web for this UNII" button; "Formula: C9H8O4"; "UNII Type: INGREDIENT SUBSTANCE"; and "InChIKey: BSYNRYMUTXBXSQ-UHFFFAOYSA-N" with a "Search the web for this InChIKey" button.

Display 5. Looking TSVVAL values for TRT - Step 2

Note: Not all treatments have a record on the UNII database, but in case you find it, you can populate the **TSVAL** as described above and **TSVCDREF= “UNII”**.

If you do not find the TRT value on the UNII database, discuss with your team about the alternatives you may have to fill this field.

## HOW TO GET TSVAl VALUES FOR TSPARMCD =PCLAS (PHARMACOLOGICAL CLASS OF INVEST. THERAPY)

- Open the National Drug File – Reference terminology link: <http://bioportalbioontology.org/ontologies/NDF-RT>. And select “Classes” from the Menu in the left upper corner, as shown in Display 6 below.

**National Drug File - Reference Terminology**  
 Summary **Classes** Properties Notes Mappings Widgets

**Details**

ACRONYM	NDF-RT
VISIBILITY	Public
BIOPORTAL PURL	<a href="http://purl.bioontology.org/ontology/NDF-RT">http://purl.bioontology.org/ontology/NDF-RT</a>
VIEW OF ONTOLOGY	<a href="http://purl.bioontology.org/ontology/NDF-RT">http://purl.bioontology.org/ontology/NDF-RT</a>
DESCRIPTION	Logically inferred version in OWL of the National Drug File - Reference Terminology (NDF-RT). NDF-RT is the terminology used by FDA and the FedMed collaboration to code these essential pharmacologic properties of medications: Mechanism of Action, Physiologic Effect, Structural Class.
STATUS	
FORMAT	OWL
CONTACT	Steven H. Brown;MD, <a href="mailto:steven.brown@va.gov">steven.brown@va.gov</a>
HOME PAGE	
PUBLICATIONS PAGE	
DOCUMENTATION PAGE	<a href="http://evs.nci.nih.gov/ftp1/NDF-RT/">http://evs.nci.nih.gov/ftp1/NDF-RT/</a>
CATEGORIES	
GROUPS	

**Metrics** ?

NUMBER OF CLASSES:	46044
NUMBER OF INDIVIDUALS:	0
NUMBER OF PROPERTIES:	49
MAXIMUM DEPTH:	13
MAXIMUM NUMBER OF CHILDREN:	5507
AVERAGE NUMBER OF CHILDREN:	5
CLASSES WITH A SINGLE CHILD:	4454
CLASSES WITH MORE THAN 25 CHILDREN:	326
CLASSES WITH NO DEFINITION:	46044

**Visits** Download as CSV

200  
175  
150

Display 6. Looking TSVAl values for PCLAS - Step 1

- Then expand the “+” sign on the left list next to “Pharmaceutical Preparations”.
- And also expand the button for “FDA Established Pharmacologic Classes”.
- Finally, look for your treatment or drug name in that list, if you found the PCLAS for your study; the TSVAl will be the “NUI” number located in the right side of the window as shown in Display 7 below. In case you are unsure, check with your team to make sure which value needs to be selected.

**National Drug File - Reference Terminology**  
 Summary Classes Properties Notes Mappings Widgets

Jump To:

Cellular or Molecular Interactions

Preferred Name	Cellular or Molecular Interactions
ID	<a href="http://evs.nci.nih.gov/ftp1/NDF-RT/NDF-RT.owl#N0000000223">http://evs.nci.nih.gov/ftp1/NDF-RT/NDF-RT.owl#N0000000223</a>
code	C624
Display_Name	Cellular or Molecular Interactions
label	Cellular or Molecular Interactions [MoA]
NUI	N0000000223
prefixIRI	N0000000223
RxNorm_CUI	986459
UMLS_CUI	C1373094
subClassOf	<a href="http://www.w3.org/2002/07/owl#Thing">http://www.w3.org/2002/07/owl#Thing</a>
disjointWith	<a href="#">RxNorm Dose Forms</a> <a href="#">Physiological Effects</a> <a href="#">Therapeutic Categories</a> <a href="#">VA Drug Interactions</a>

Display 7. Looking TSVAl values for PCLAS - Step 2

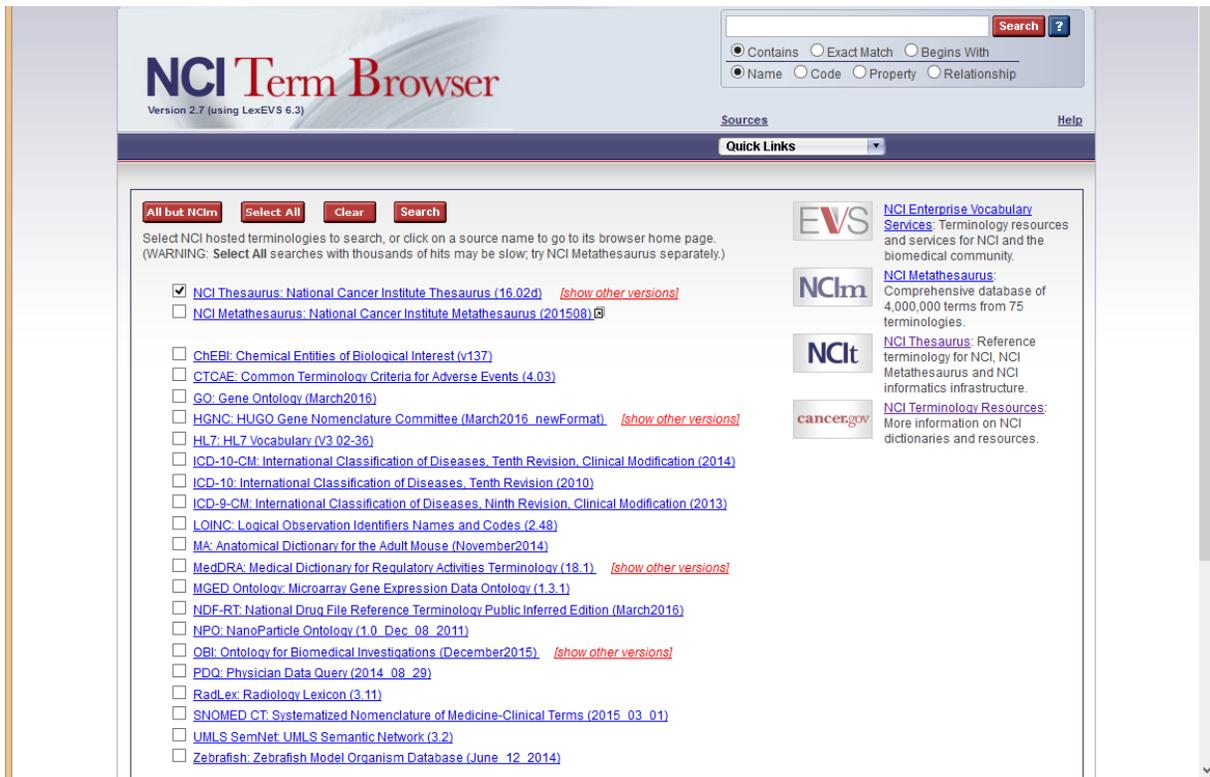
## HOW TO GET TSVAL VALUES FOR TSPARMCD = TDIGRP (DIAGNOSIS GROUP) AND INDIC (TRIAL INDICATION)

- Go to the National Cancer Institute (NCI) Term Browser: <https://ncit.nci.nih.gov/ncitbrowser/pages/vocabulary.jsf>
- Then select “NCI Term Browser” from the right menu as shown in the Display 8 below.



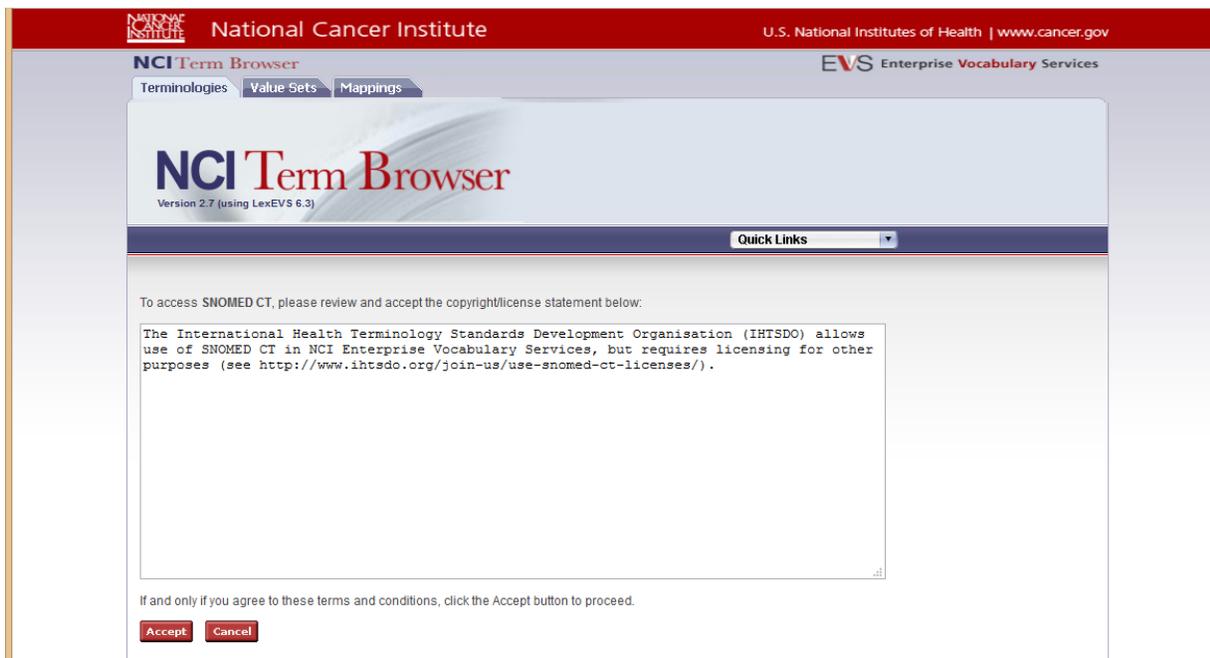
Display 8. Looking TSVAL values for TDIGRP - Step 1

- Then select the “SNOMED CT” link shown at the bottom of the Display 9 below.



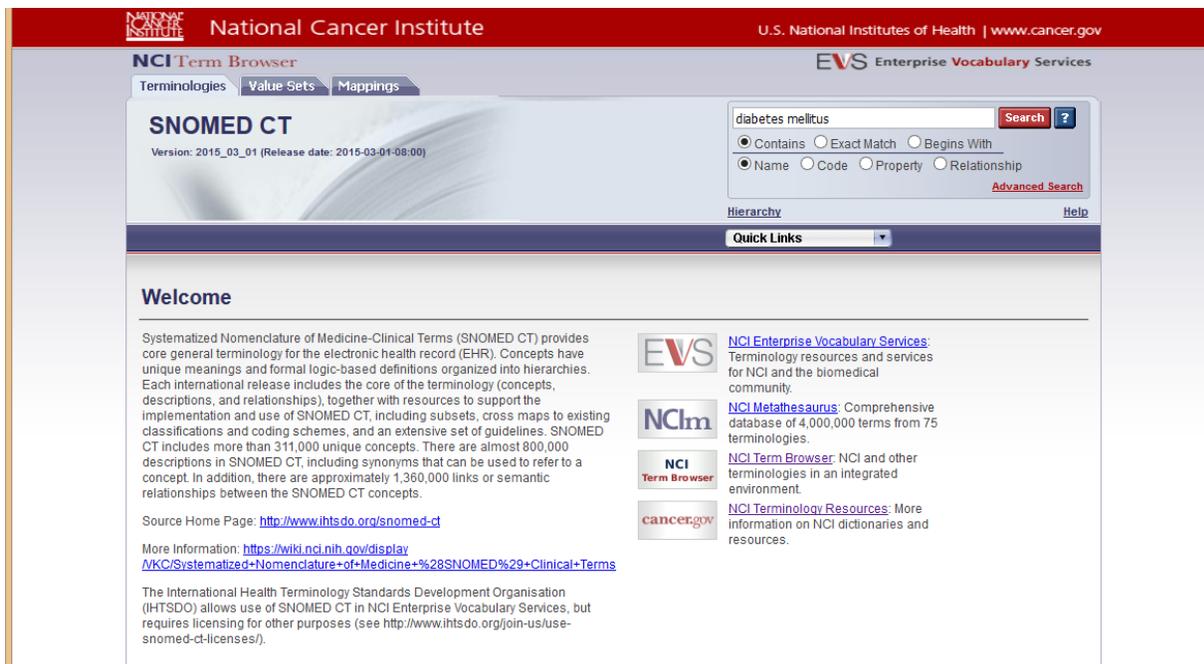
Display 9. Looking TSVAL values for TDIGRP - Step 2

- Then accept the terms and conditions for that page as shown in the Display 10 below.



Display 10. Looking TSVAl values for TDIGRP - Step 3

- Now you can type the “Diagnosis” or the “Indication” for your study in the “Search” box on the right upper corner and select the one better fits to your study. An example for “Diabetes Mellitus” is given on Display 11 below.



Display 11. Looking TSVAl values for TDIGRP - Step 4

- Now select the term that better fits to your study **“Indication”** or **“Diagnosis”**, then the TSVAl will be **“73211009”** (SNOMED Clinical Terms Code) for this example, as shown in Display 12 below.



Display 12. Looking TSVAl values for TDIGRP - Step 5

## USEFUL TIPS WHILE DEVELOPING SDTM SPECS

- Make sure to understand the SDTM IG basic concepts and the difference between them like: EPOCH, ELEMENT, ARM, etc. and review the examples before creating your own specs; in the SDTM IG they provided a lot of helpful information that can fit into your study.
- Make sure to use the correct CDISC – SDTM Terminology file. When in doubt do not hesitate to consult with your team, is always better to ask than do re-work!
- The TS and the TI domain are filled from the protocol directly, (make sure to review the minimum parameters required by the corresponding SDTM IG version; for TS in SDTM IG version 3.2 they are located in the Appendix C1: Trial Summary Codes.

## MAIN DIFFERENCE BETWEEN TDM SPECS ON SDTM IG V3.1.2 AND V3.1.3

Basically all the TDM domains remain the same in both version, except for the TS domain, since in v3.1.3, there are four new variables that were added as well as additional required parameters and controlled terminology associated with additional parameters. This gives a more “complicated” touch to the TS specs development. The new variables are:

- TSVAlNF – Parameter Null Flavor. This new variable is required when there is no value on the TSVAl field to explain why the value is not present. There is associated controlled terminology for this field.
- TSVAlCD – Parameter Value Code. This variable could have more than one unique value, please check the SDTM IG examples for further reference. This is for example, the value of the code associated with VALUE itself within a code list, not the code to identify a code list.
- TSVAlDREF – Name of the Reference Terminology. There is not much information in the SDTM IG v3.1.2 about this new variable, but the majority of the sources can be found with the external links referenced before. This is essentially the name of the source for the code list/values used.
- TSVAlDVER – Version of the Reference Terminology. Be careful when getting TSVAlDREF values and take a note of the version you are using to fill this field.

## CONCLUSION

As you can see in this paper, developing the TDM specs for the first time is not a simple process, you have to use a lot of resources and have clear many concepts at the same time, and this can drive you really crazy if you do not know where to look.

I hope this summary of tips and tricks will be helpful to all of you and your teams when doing this task in your organization. Keep in mind that just the experience can help you to fill the gaps you may have at the beginning, do not give up! And do not be afraid to ask questions to the experts, also keep a constant communication with your team; that way the whole process will be easier.

## REFERENCES

- Clinical Trials.Gov website: [www.clinicaltrials.gov](http://www.clinicaltrials.gov)
- U.S. Food and Drug Administration. Substance Registration System – Unique Ingredient Identifier website: <http://fdasis.nlm.nih.gov/srs>
- National Cancer Institute (NCI) Term Browser: <http://ncit.nci.nih.gov/ncitbrowser/pages/vocabulary.isf>
- National Drug File – Reference terminology website: <http://bioportal.bioontology.org/ontologies/NDF-RT>

## ACKNOWLEDGMENTS

Thanks to Christine McNichol for helping me during this tough learning process and also for being there always to help responding questions and reviewing my documents with so many patience!

## RECOMMENDED READING

- CDISC SDTM Implementation Guide (version corresponding to your study)

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Ruth (Marisol) Rivera  
Enterprise: Chiltern, Inc.  
E-mail: [ruth.riverabarragan@chiltern.com](mailto:ruth.riverabarragan@chiltern.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.