

# SEND Implementation Wiki - CT Fundamentals

## Table of Contents

<i>What is CT?</i> .....	2
<i>Where Do I Find It?</i> .....	2
<i>Using CT Files</i> .....	2
<i>Fields</i> .....	3
<i>Dealing with Updates</i> .....	3
<i>Updating the OpenCDISC Validator Configuration</i> .....	4
<i>Mapping Considerations</i> .....	5
<i>Finding a Match for Your Term</i> .....	5
<i>Getting Your Terms/Synonyms Added</i> .....	6
<i>Mapping Design</i> .....	6
<i>Changing Internal Lexicons</i> .....	6
<i>Links</i> .....	6

## What is CT?

Controlled Terminology, or CT, is a collection of harmonized lists of preferred terms where one term in a list represents a single concept. A single concept may have many synonyms; CT's job is to ensure that synonymous terms map to the same preferred term, such that different organizations effectively all call the same concept by the same name.

For example, the adrenal gland tissue can be specified in multiple ways, as "gland, adrenal", "Adrenal Gland", "adrenals", etc. CT provides a common term, "GLAND, ADRENAL" to which all submissions must map.

CT is a critical piece to data standardization in ensuring consistency of meaning of terms across organizations.

CT is developed for SEND by the SEND CT subteam in conjunction with CDISC CT, who ensures harmonization between clinical and nonclinical sides, and NCI-EVS, who provides the database and semantic term management.

## Where Do I Find It?

- [CDISC SEND](#) - This page has links to the current SEND Controlled Terminology
- [NCI Thesaurus](#) - An online dictionary hosted and maintained by NCI-EVS, which can be used to search for terms, view synonyms and preferred terms across standards, and so on.

As releases are published, the CDISC SEND site is updated. It is recommended to subscribe to the page's RSS feed (start by clicking the Subscribe to our RSS Feed button on the right of the CT page) so that you can be apprised of updates.

## Using CT Files

Each release of a CT package comes with a few key files.

- *SEND Terminology* - This file contains the complete SEND Terminology file. It comes in several extensions, such as xls (ideal for viewing) and txt (ideal for electronic assimilation). The xls file has AutoFilter turned on, so you can more easily search for lists, terms, synonyms, etc.

*SEND Terminology Changes* - This file contains the changes made since the last package. It also comes in a couple extensions. Assuming you are updating as each package comes out, this becomes an invaluable resource to determine what sort of changes you need to effect in your systems.

## Fields

SEND Terminology is defined with several key columns:

- *Code* - Also known as the CUI Code, this is a code unique to the term concept that generally stays the same across versions. For example, the "%" term from the UNIT codelist has a Code of C25613. Suppose in a later version, the preferred term is changed to "PERCENTAGE". The Code will remain C25613. The same code value is assigned to both --TEST and --TESTCD for the same concept enabling systems to ensure correct --TESTCD value is matched with a specific --TEST value. This field is critical for systems to link the same term across versions. Note that for rows referencing a list (usually highlighted and with a null Codelist Code), this Code represents the unique code for the list as opposed to a term. Note also that you can search the [NCI Thesaurus](#) with this value to find the term and get more information. The information in this thesaurus is available for download in various formats from the [NCI Thesaurus FTP site](#).
- *Codelist Code* - This field provides a similar identifier as above, but referencing the list to which the term applies. For example, for the "%" term, the Codelist Code of C71620 is that of the UNIT codelist.
- *Codelist Extensible (Yes/No)* - Only populated for lists, this specifies whether the codelist is extensible, meaning that terms may be added if not already in the list. When "No", it means that you must map to a term in the list, or it will be out of SEND compliance. When "Yes", you should make every effort to use an existing term, as it is most likely already present, but you may use additional terms if the concept you need is not presently in the list.
- *Codelist Name* - This provides the human readable name of the list to which the row applies. It is helpful to autofilter on this column to view a particular list.
- *CDISC Submission Value* - This field gives the preferred value that should be used for submission purposes. For instance, you may collect in units of "mg/mL", but the preferred label for this unit is "g/L". This field would specify the "g/L" preferred label.
- *CDISC Synonym(s)* - This field provides most common synonyms for the term. When mapping, it is helpful to search this column for the term you use if you do not find it in the CDISC Submission Value field. In the mg/mL example above, mg/mL is listed among the synonyms for g/L.
- *CDISC Definition* - This field gives a definition of the term.

*NCI Preferred Term* - This field gives the preferred term as specified by NCI. This value can be used to directly find the term in the [NCI Thesaurus](#) to get more information.

## Dealing with Updates

The SEND Terminology Changes spreadsheet provides a list of the changes between the current version and the one immediately before it.

The changes of which to take special note are changes to existing terms and deletions. These can mean special (possibly manual) updates to your systems, especially if the CUI Code changes. Types of changes:

- *New terms* - This is the most common form of change, either in the form of new lists or new terms. These are also the least impactful, as the main consideration is just whether to use the new terms. If you were previously extending the list to provide a term that now has an official mapping, be sure to have that extended term mapped to the new official term (and submission value).
- *Synonym or definition change* - this generally has no impact
- *Preferred term change* - this will not impact anything provided that you base off the Code and that your mapping dictionaries map to codes as opposed to text. If either of these is not true, then you will need to update mappings and so on.
- *Code change* - If the CUI Code behind a term changes (seldom), then you may need to handle delicately. This is a case where having an internal identifier similar to the CUI is a good idea, since you will be shielded from this change. If not, then be careful to remap as needed

*Term deprecation* - In the case that a term is removed, you should make sure that this term was not used by your dictionaries. This can either be due to the term being inappropriate for the list (in which case, it was probably also not used) OR because it is a duplicate of an existing term. In either case, verify that your mappings do not use the term, and in the latter case, that they now map to the right term.

## Updating the OpenCDISC Validator Configuration

If you use the [OpenCDISC Validator](#) to validate your SEND packages, it is a good idea to update this tool's dictionaries with updated CT as well while going through the update process.

The [OpenCDISC Validator](#) stores a copy of the same SEND Terminology file you can obtain from the [NCI site](#), using the SEND Terminology.txt file. To update:

1. Download the SEND Terminology.txt file for the version you want, if you haven't already
2. Place the SEND Terminology.txt file in the config\data subfolder of the validator's install path (e.g., C:\Program Files\opencdisc-validator\config\data), overwriting the previous file

These same steps can be followed to have the tool validate against any version of CT (past versions are available from the [NCI site](#) under Archive.

## Mapping Considerations

When populating a field that has controlled terminology, you will need to map terms that you have used to the controlled terms. This section has some tips for mapping and system design.

### Finding a Match for Your Term

First, familiarize yourself with the list by reviewing the list and seeing what types of terms are within. This will help with some of the mapping.

When trying to find a match for your term, the simplest way to start is to search for your term.

1. First, filter to the list you want. In the SEND Controlled Terminology excel file, filter to the list that you want by clicking the drop-down to the side of the "Codelist Name" header, selecting the list you want.
2. Next, you can search the file for your term. This is generally more effective if you first select the CDISC Submission Value and CDISC Synonym(s) columns before doing a Find (to confine your search to these columns).

In the simple case, your exact term will either already be the submission value or be listed among the synonyms for a term.

If this doesn't work, try looking for pieces of the term. For instance, if you have "adrenals" but didn't find a term, try searching for just "adrenal". If you have a compound word, such as thyroid/parathyroid, try searching for one of the terms.

If this still doesn't work, review the list to see if it is a synonym for something already on the list, just not listed as a synonym. It may help to have someone who is versed in the science behind it to assist with this task.

Another search that you can do is against the [NCI Thesaurus](#). It has all of the information contained in the SEND Controlled Terminology spreadsheet and more.

It is important to understand that a term you use may actually refer to multiple concepts, and thus map to multiple controlled terminology lists. For example, if you use a lab test name "Urine Protein" this is actually referring to two concepts: a test and a specimen analyzed. So your test name would map to both a LBTEST controlled term and an LBSPEC controlled term.

If all else fails, see the Getting Your Terms/Synonyms Added section below to submit a term suggestion.

## Getting Your Terms/Synonyms Added

You may have a term or synonym to an existing term that is not currently represented.

If you have exhausted the CT list in question and cannot find your term OR you have found the term, but your synonym is not present, submit a new term/synonym request through NCI-EVS's [Term Suggestion](#) online form. This will propagate through the proper channels so that your addition/change will be considered and/or added/changed. If the term already exists, then you will receive a response as to which term to which your term should map.

## Mapping Design

When storing or designing systems to map to CT, consider not mapping to the textual submission value and instead to an internal identifier for the term, that will stay consistent over time. The Code field listed in the SEND Controlled Terminology file (aka "CUI Code") can serve this purpose, although this can also change over time (although usually for good reason).

Then, if the submission value changes, you will be immune to the change, as any of your mappings reference the identifier which remains the same. This also eases the task of producing submissions for different versions, as you can reference the code, and then see what the term value was for the code as of a certain version.

## Changing Internal Lexicons

A possibility to smooth the process of mapping can be to change internal lexicons to the preferred terms in CT. Long term, this can save some headache and human translation from raw data to SEND datasets.

However, even if you modify your lexicons to perfectly match a particular version of CT, it is still a good idea to map your terms to their CT equivalents, since preferred terms can change over time (although infrequently).

## Links

- [CDISC SEND](#) - This page has links to the current SEND Controlled Terminology
- [NCI Thesaurus](#) - An online dictionary hosted and maintained by NCI-EVS, which can be used to search for terms, view synonyms and preferred terms across standards, and so on.