Navigating the Labyrinth of Controlled Terminology: Challenges and Opportunities in Integrating and Managing Diverse Terminologies for Broad Use of Metadata

Presented by Kasper Dideriksen, Senior Standards Developer, Novo Nordisk
Martin Gram, Principal Standards Specialist, Novo Nordisk

Meet the Speakers

Kasper Dideriksen

Title: Senior Standards Developer

Organization: Novo Nordisk

Kasper Dideriksen, PhD, is specialized within cardiovascular assessments, safety, COA and image data collection. In his daily work, he is working on translating the newest innovative experimental approaches into the CDISC controlled terminology framework, utilizing his background as a clinical researcher in human physiology to navigate the cross-field between life science and data science. In addition, Kasper is volunteer in the CDISC ECG Controlled Terminology team.

Martin Gram

Title: Principal Standards Specialist

Organization: Novo Nordisk A/S

Martin Gram is a principal standards specialist within Novo Nordisk, where he is utilising his background as a clinical researcher in human physiology to help drive the field of standardisation and data specifications. In addition, Martin is also involved in several cutting-edge initiatives incl. the StudyBuilder project aiming at creating a seamless, reusable metadata flow.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

Agenda

1. Introduction
2. Findings in CDISC CT
3. SPOR RMS vs CDISC vs protocol CT
4. CDISC vs FDA vs Internal CT
5. Possible solutions and the way forward

Introduction

- The use of controlled terminology (CT) has historically been driven by requirements from various regulatory agencies
- Gradually, the use of CT is expanding to improve alignment and increase efficiency
- However, regulatory authorities and standards developing organisations don´t require/recommend the same CT
- This creates challenges when trying to create a ‘one-size-fits-all’ CT management system that supports ‘one-source-of-truth’
- Novo Nordisk is creating the OpenStudyBuilder to manage CTs

SETTING THE STAGE

Navigating a ‘multi terminology landscape’ searching for the ‘one-source-of-truth’ terminology

Internal terminology
Tradition, culture, opinions

EMA
SPOR RMS

FDA
Technical specifications
Guidance for Industry

Other dictionaries
UCUM, SNOMED, MedDRA, WHODrug

Submission data
CDISC CT, NCI

Protocol terminology
TransCelerate CPT,
ICH M11, ICH E3, CDISC

Findings in CDISC CT

CDISC CONTROLLED TERMINOLOGY RULES: Rules for All Codelists

24 Sept 2021

A CDISC terminology concept is defined as the c-code, synonym(s), and definition.

- CDISC terminology concept can be associated with multiple codelists.
- CDISC submission value can be the same or different across multiple codelists.
- CDISC terminology concept used across multiple codelists must have the same NCI C-code, CDISC synonym(s) and CDISC definition.

- Concepts can develop over time but within a package date the concept should remain the same across the CDISC CT deliverables.
- Does it always add value to accept different submission values across codelists?

CDISC Glossary

The definition for a concept is different between Glossary and other packages in 107 cases in the 2023-12-15 package:

C15197:

- **SDTM CT: CASE CONTROL**. A study that compares groups of people with generally similar characteristics, those with the condition under study (case) and those without the condition under study (control)
- **Glossary CT: Case-Control Study**. Retrospective study in which individuals with an outcome (cases) are compared with those who do not have the outcome (controls). The outcome variable (disease, event, experience, biomarker) is chosen first, and the exposure (e.g., treatment) is evaluated in cases vs controls to see whether there is an association between exposure and outcome. [After AMA Manual of Style] See also outcome, observational study, exposure.

C42636:

- **SDTM CT: Dose Form**. The physical form in which active and/or inert ingredient(s) are presented.
- **Glossary CT: Pharmaceutical Dosage Form**. Physical characteristics of a drug product, (e.g., tablet, capsule, or solution) that contains a drug substance, generally-but not necessarily-in association with one or more other ingredients. [21 CFR 314.3 and after IDMP]. See also drug product.

Example: The submission value is different

Identical c-code, synonym, definition and NCI preferred term.

C17998:

Submission value: UNKNOWN

Submission value: U

Definition: Not known, not observed, not recorded, or refused. (NCI)

C28554:

Submission value: DEATH

Submission value: EXPIRED

Submission value: DEAD

Definition: The absence of life or state of being dead. (NCI)

Example: The name submission value is different

Identical c-code, synonym, definition and NCI preferred term.

C127628

Name submission value: Maturing Erythroid **Cells**/Total Cells

Code submission value: MERCECE

Name submission value: Maturing Erythroid/Total Cells

Code submission value: MERCECE

Definition: A relative measurement (ratio or percentage) of the maturing erythroid cells to total cells in a biological specimen.

Example: The name and code submission value are different

Identical c-code, synonym, definition and NCI preferred term.

C123556

Name submission value: Tumor Cells/Total Cells

Code submission value: TUMCECE

Name submission value: Neoplastic Cells/Total Cells

Code submission value: NEOCECE

Definition: A relative measurement (ratio or percentage) of the neoplastic cells to total cells in a biological specimen.

SPOR RMS vs CDISC vs protocol CT

Metadata for medicinal products and associated metadata is used very broad. It is a wish from business to standardize metadata across areas:

- SDTM and ADaM
- Single case submissions for safety
- Protocol intervention table
- etc...

- EMA mandates the use of the ISO IDMP model for medicinal products which uses the Referentials Management Services (SPOR RMS) as dictionary
- SDTM and ADaM use CDISC CT
- Protocol metadata is currently without requirements but ICH M11, Transcelerate and CDISC have published recommendations

SPOR RMS maps to CDISC CT

Examples:

- One term in SPOR RMS is mapped to several terms in CDISC.
- Granularity is different with 530 SPOR terms mapping to 75 CDISC terms (of 189)

SPOR Pharmaceutical Dose Form	CDISC C-code	CDISC Pharmaceutical Dosage Form
Dispersion for injection	C42950	INJECTION, LIPID COMPLEX
Dispersion for injection	C42951	INJECTION, SUSPENSION, LIPOSOMAL
Dispersion for injection	C42988	INJECTION, SUSPENSION, SONICATED
Prolonged-release tablet	C42905	TABLET, DELAYED RELEASE
Prolonged-release tablet	C42927	TABLET, EXTENDED RELEASE
Prolonged-release tablet	C42963	TABLET, MULTILAYER, EXTENDED RELEASE
Prolonged-release tablet	C42997	TABLET, DELAYED RELEASE PARTICLES
Prolonged-release tablet	C61006	TABLET, ORALLY DISINTEGRATING, DELAYED RELEASE

SPOR RMS maps to CDISC CT

Examples:

- The definitions of terms are not the same
- UCUM seems to think that Capsule is not a unit
- CDISC uses the submission value plus the codelist to define the context so a 1:1 mapping is not possible

SPOR RMS Units of Presentation	CDISC Unit	UCUM Unit	CDISC Pharmaceutical Dosage Form	SPOR RMS Pharmaceutical Dose Form
Capsule A unit of presentation used to represent the quantity of product that is found in a single discrete entity where the pharmaceutical dose form is a type of capsule.	CAPSULE A dosing measurement based on the capsule unit.(NCI)	N/A	CAPSULE A solid pharmaceutical dosage form that contains medicinal agent within either a hard or soft soluble container or shell, usually used for the oral administration of medicine. The shells are made of a suitable form of gelatin or other substance. (NCI)	Capsule, hard Solid single-dose preparation contained in a hard shell, the capacity of which can be varied. The shell is made of gelatin or other substances. It consists of two prefabricated cylindrical sections one end of which is rounded and closed, the other being open. The contents of the shell may be a solid or semi-solid preparation, which is filled into one of the sections and closed by slipping the other section over it. Hard capsules are intended for oral use.

SPOR RMS maps to CDISC CT

SPOR RMS Pharmaceutical Dose Form	CDISC Pharmaceutical Dosage Form
Capsule, hard	
Capsule, soft	
Chewable capsule, soft	
Gastro-resistant capsule, hard	
Gastro-resistant capsule, soft	
Inhalation powder, hard capsule	
Inhalation vapour, capsule	
Intrauterine capsule	
Modified-release capsule, hard	
Modified-release capsule, soft	
Oromucosal capsule	
Prolonged-release capsule, hard	
Prolonged-release capsule, soft	CAPSULE, COATED, EXTENDED RELEASE
Prolonged-release capsule, soft	CAPSULE, FILM COATED, EXTENDED RELEASE

SPOR RMS Pharmaceutical Dose Form (continued)	CDISC Pharmaceutical Dosage Form (continued)
	CAPSULE, COATED PELLETS
	CAPSULE, COATED
	CAPSULE, DELAYED RELEASE PELLETS
	CAPSULE, DELAYED RELEASE
	CAPSULE, EXTENDED RELEASE
	CAPSULE, GELATIN COATED
	CAPSULE, HARD, EXTENDED RELEASE
	CAPSULE, IMMEDIATE RELEASE
	CAPSULE, LIQUID FILLED
	CAPSULE, SOFTGEL
	CAPSULE, SOFTGEL, EXTENDED RELEASE
	INHALATION VAPOR, CAPSULE

SPOR RMS maps to CDISC CT

Examples:

- PILL and TABLET have unique definitions in CDISC, but PILL does not exist in SPOR RMS anymore?
- Content curation and versioning of terms follow different processes

SPOR Pharmaceutical Dose Form	SPOR RMS Definition	CDISC Pharmaceutical Dosage Form	CDISC Definition
Tablet	A type of solid pharmaceutical dose form consisting of a mass that is formed by compression of uniform volumes of particulate solids, or by other means such as extrusion or moulding.	TABLET	A solid dosage form containing medicinal substances with or without suitable diluents. (NCI)
NULLIFIED	NULLIFIED	PILL	A dose of medicine in the form of a small pellet. (NCI)

CDISC vs FDA vs Internal CT

When to follow which guidance/standard?

- Decisions based on guidance, opinion and tradition

CDISC vs FDA vs Internal terminology

- **Practical examples of how to navigate in a 'multi terminology landscape'**

- Example 1) FDA vs CDISC - Race and Ethnicity data collection

- Example 2) FDA vs CDISC - NASH data collection

- Example 3) FDA vs Internal terminology - NASH data collection

Example 1: Collection of Race and Ethnicity Data in Clinical Trials and Clinical Studies for FDA-Regulated Medical Products

To follow the FDA guidance, the following Submission Values should be in ETHNICC codelist:

- Chicano

or

- Mexican, Mexican American, Chicano
- **Other** Hispanic or Latino

Note: "MEXICAN", "MEXICAN AMERICAN", "HISPANIC OR LATINO" and "LATIN AMERICAN" exists as separate submission values in the CDISC codelist.

FDA: Collection of Race and Ethnicity Data in Clinical Trials.., 2024 draft

Ethnicity Data Standard

Are you Hispanic or Latino? (One or more categories may be selected.)

a. No, not Hispanic or Latino
b. Yes, Mexican, Mexican American, Chicano
c. Yes, Puerto Rican
d. Yes, Cuban
e. Yes, Other Hispanic or Latino

These categories are part of the Hispanic or Latino category of the OMB standard

FDA: Collection of Race and Ethnicity Data in Clinical Trials, 2016

Ethnicity Data Standard

Are you Hispanic, Latino/a, or of Spanish origin? (One or more categories may be selected)

a. ____No, not of Hispanic, Latino/a, or Spanish origin
b. ____Yes, Mexican, Mexican American, Chicano/a
c. ____Yes, Puerto Rican
d. ____Yes, Cuban
e. ____Yes, Another Hispanic, Latino/a or Spanish origin

These categories roll up to the Hispanic or Latino category of the OMB standard

CDISC vs FDA

- FDA => Create new codes ” Mexican, Mexican American, Chicano” and ”**Other** Hispanic or Latino”
- CDISC => Stick to existing codes “HISPANIC OR LATINO” and “LATIN AMERICAN” etc.
- In Novo Nordisk we choose to follow FDA and create sponsor defined codes for the missing terms

Example 2: Technical Specifications for Submitting Clinical Trial Data Sets for Treatment of Noncirrhotic Nonalcoholic Steatohepatitis (NASH)

For capturing the size (diameter) of biopsy needle used.

To follow the FDA guidance, we need NEED GAUGE to be in UNIT codelist.

Notes:

There is a FATEST= NEEDLE GAUGE (FATESTCD= NDLGAUGE).

CDISC did retire NEEDLE GAUGE from UNIT codelist in CT package 2021-12-17.

Table 1: Biospecimen Domain Test Values

BSTESTCD	BSTEST	BSORRES	*Notes*
DIAMETER	Diameter		This test is extended from the BSTEST codelist.
LENGTH	Length		The result should be reported in millimeters (mm).
NEEDSIZE	Needle Size	e.g., 16	The unit of "NEEDLE GAUGE" would be captured under BSORRESU. This test is extended from the BSTEST codelist.

FDA: Technical Specifications for Submitting Clinical Trial Data Sets for Treatment of Noncirrhotic Nonalcoholic Steatohepatitis (NASH), 2022

CDISC vs FDA

- FDA => Add ”NEEDLE GAUGE” to UNIT and “NEEDSIZE” to BSTEST codelists
- CDISC => Existing FATEST: ”NEEDLE GAUGE”. No “NEEDSIZE” (or “DIAMETER”) in BSTEST
- In Novo Nordisk we choose to follow FDA and create sponsor defined codes for the missing terms

Example 3: Technical Specifications for Submitting Clinical Trial Data Sets for Treatment of Noncirrhotic Nonalcoholic Steatohepatitis (NASH)

Assessment	Collected Responses	FDA guidance
Ishak Fibrosis Score	0 - NO FIBROSIS 1 - FIBROUS EXPANSION OF SOME PORTAL AREAS 2 - FIBROUS EXPANSION OF MOST PORTAL AREAS 3 - FIBROUS EXPANSION OF MOST PORTAL AREAS, WITH OCC PORTAL-TO-PORTAL BRIDGING 4 - FIBROUS EXPANSION OF PORTAL AREAS WITH MARKED BRIDGING 5 - MARKED BRIDGING WITH OCCASIONAL NODULE (INCOMPLETE CIRRHOSIS) 6 - CIRRHOSIS, PROBABLE OR DEFINITE	0-18
Portal Inflammation	NONE MINIMAL (SPRINKLING MATRIX > INFLAMMATION) MILD (MATRIX APPROXIMATELY = INFLAMMATION) MODERATE (INFLAMMATION COVERS MATRIX) SEVERE (INFLAMMATION COVERS MATRIX AND MATRIX IS EXPANDED)	None to minimal, >Minimal

FDA vs Internal terminology

- **ISHAK SCORING SYSTEM:** We use the fibrosis score from 0-6, FDA guides for the necro-inflammatory scores from 0-18.

- **PORTAL INFLAMMATION:** Interest in portal inflammation changes => added the grade 3 and 4.

- For other parameters like **NAFLD TOTAL SCORE, STEATOSIS EXTENT, LOBULAR INFLAMMATION, HEPATOCELLULAR BALLOONING** and **NASH CRN FIBROSIS STAGE** we follow FDA guidance.

Possible solutions and the way forward

- Controlled terminology is constantly evolving and improving, but it is currently not uniform
- Conflicting CT recommendations has to be dealt with if we want to improve standardisation
- The build of a CT management system should be able to include and integrate different CTs to deliver one-source-of-truth for broad use of metadata
- Sponsor defined synonym mappings can be necessary until CT organisations integrate with each other
- Different content hierarchy can be particular difficult to manage

Thank You!

Contact:

Kasper Dideriksen - KPDI@novonordisk.com

Martin Gram - XMGM@novonordisk.com