Conceitos OMOP CDM

Conceitos (*concepts*) utilizados no projeto de aplicativo SAÚDE!

Observação

Caso seja necessário a criação de um novo ***concept*** utilizar um id notar que é necessário o uso de um valor maior que 2.000.000.000.

Fonte https://ohdsi.github.io/CommonDataModel/faq.html

If an source code is not supported by the OMOP Vocabulary, one can create a new records in the CONCEPT table, however the CONCEPT_IDs should start >2000000000 so that it is easy to tell between the OMOP Vocabulary concepts and the site specific concepts. Once those concepts exist CONCEPT_RELATIONSHIPS can be generated to assign them to a standard terminologies, USAGI can facilitate this process as well (THEMIS issue #22).

Entrada padrão para um campo sem resposta

Vocabulário None (subconjunto com concept -> Non-Standard e validity -> Valid)

Localizado no arquivo none_concept_id.csv (1 entrada)

ID	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
0	No matching concept	No matching concept	Undefined	Non-standard	Valid	Metadata	None

Pode ser usado para qualquer campo *concept*

5.6.10 Flavors of NULL

Many vocabularies contain codes about absence of information. For example, of the five gender concepts 8507 “Male,” 8532 “Female,” 8570 “Ambiguous,” 8551 “Unknown,” and 8521 “Other”, only the first two are Standard, and the other three are source concepts with no mapping. In the Standardized Vocabularies, there is no distinction made why a piece of information is not available; it might be because of an active withdrawal of information by the patient, a missing value, a value that is not defined or standardized in some way, or the absence of a mapping record in CONCEPT_RELATIONSHIP. Any such concept is not mapped, which corresponds to a default mapping to the Standard Concept with the concept ID = 0.

Fonte: https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html

Entrada padrão para um campo sem resposta (criação no BD)

Melhor maneira de fazer até o momento encontrada : Pegar os dados já exportados dos Vocabulários do Athena (que já vem completos para os concepts que foram escolhidos). Caso não funcione, usar NULL nas FK (se o SGBD não reclamar).

Fonte: https://forums.ohdsi.org/t/newbie-vocabulary-import/574/25

concept_id	concept_name	domain_id	vocabulary_id	concept_class_id
0	No matching concept	Metadata	None	Undefined

standard_concept	concept_code	valid_start_date	valid_end_date	invalid_reason
	No matching concept	19700101	20991231	

Referência para vocabulários

https://ohdsi-github-io.translate.goog/TheBookOfOhdsi/StandardizedVocabularies.html?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

(capítulo que discute mais sobre vocabulários)

Sobre concepts Standard e Non-standard:

“One concept representing the meaning of each clinical event is designated the Standard. … The others are designated non-standard or source concepts and mapped to the Standard ones. Standard Concepts are indicated through an “S” in the STANDARD_CONCEPT field. And only these Standard Concepts are used to record data in the CDM fields ending in "_CONCEPT_ID".”

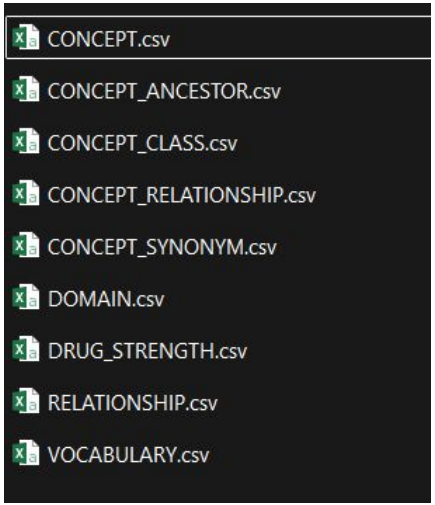
Tabelas Auxiliares

As tabelas mostradas abaixo, já vem populadas automaticamente ao baixar os vocabulários (É PRECISO ESTAR LOGADO, O BOTÃO DOWNLOAD ACIMA É O CORRETO) no site Athena.

Dessa forma, ao selecionar os vocabulários desejados, não precisamos nos preocupar em popular as tabelas CONCEPT, CONCEPT_CLASS,

DOMAIN, VOCABULARY, que estão no nosso modelo inicial, basta exportar para o BD desses arquivos .csv. A tabela CONCEPT_SYNONYM precisará ser populada com as traduções dos conceitos que mostraremos aos usuários.

O trabalho a seguir é então escolher quais conceitos mostraremos para o usuário escolher (dentro do escopo do projeto SAÚDE!)



gender_concept_id (Person)

Vocabulário Gender (subconjunto com concept -> Standard e validity -> Valid)

Localizado no arquivo gender_concept_id.csv (2 entradas)

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
8532	F	FEMALE	Gender	Standard	Valid	Gender	Gender
8507	M	MALE	Gender	Standard	Valid	Gender	Gender

`8551 - OTHER` *

Pela referência encontrada no slide 5, usaria apenas concepts Standard e Valid

Contextualização gender_concept_id (Person)

How to populate gender_concept_id

The term “gender_concept_id” is outdated and really should be “sex_concept_id”. Since changing gender_concept_id is a huge lift for developers and package maintainers, this change will be implemented at the next major release. In the meantime, what value should be used to populate gender_concept_id?

The ratified convention

PERSON.gender_concept_id should hold the standard concept representing the patient’s assigned sex at birth

Fonte: https://ohdsi.github.io/Themis/populate_gender_concept_id.html#convention-type

5.6.1 Gender

Gender in the OMOP CDM and Standardized Vocabularies denotes the biological sex at birth. Often, questions are posed how to define alternative genders. These use cases have to be covered through records in the OBSERVATION table, where the self-defined gender of a person is stored (if the data asset contains such information).

Fonte: https://ohdsi-github-io.translate.goog/TheBookOfOhdsi/StandardizedVocabularies.html?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

race_concept_id (Person)

```
add_concept(8515, "Asian", "Race", "2", "Race", "Race", "Asiático")
add_concept(8527, "White", "Race", "5", "Race", "Race", "Branco")
add_concept(38003572, "American Indian", "Race", "1.01", "Race", "Race", "Indígena Americano")
add_concept(38003598, "Black", "Race", "3.01", "Race", "Race", "Preto")
```

Para a primeira versão esses 3 conceitos parecem ser o suficiente mesmo.

Sugestão: ao invés de “black or african american” utilizar o concept 38003598 “black”. Parece ser uma tradução mais fiel do campo.

+ Indígena americano (“American Indian”, 38003572) -> não sei se é preciso para o continente americano como um todo

  Id = 0 caso o usuário não queira identificar esse campo. Descrito melhor no slide 3.

race_concept_id (Person)

Vocabulário Race (subconjunto com concept -> Standard e validity -> Valid)

Localizado no arquivo race_concept_id.csv

Exemplo das 4 primeiras linhas da tabela (existem 50 em hierarquia)

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
38003600	3.03	African	Race	Standard	Valid	Race	Race
38003599	3.02	African American	Race	Standard	Valid	Race	Race
38003573	1.02	Alaska Native	Race	Standard	Valid	Race	Race
38003572	1.01	American Indian	Race	Standard	Valid	Race	Race
8657	1	American Indian or Alaska Native	Race	Standard	Valid	Race	Race

Contextualização race_concept_id (Person)

How to handle multiple races or ethnicities per person

Data sources might have more than one race value per a person, have more than one source with differing race values or have a source value semantically equivalent to “multiple race” (I.e. multi-racial, > 1 race, etc.)

The ratified convention

Where do the data go?

•If your data has only one race source value, then map this to PERSON.race_concept_id *Example:* There is one race source value, ‘Black or African American’, for a person. Then PERSON.race_concept_id = 8516, ‘Black or African American’

•If your data has > 1 race source value 1. Use 44814659, ‘Multiple Race’, to populate PERSON.race_concept_id 2. In the Observation table, populate OBSERVATION.observation_concept_id = 4013886, ‘Race’. THEN populate OBSERVATION.value_as_concept_id with the concept_id for the person’s race. Create as many records for a person as the source data have.

What date do I put into the OBSERVATION.observation_date for the records?

•The OBSERVATION.observation_date represents the date in which the fact was recorded. IF your source data do not have a date associated with the record or the visit in which the fact was recorded, THEN use the date of the most recent visit record for a person. Logic for this decision: We never use default dates far in the past or future because this will make our Observation Period erroneously long. The same goes for birth date. It is against CDM rules to leave this field NULL. So, since we don’t know when the fact was recorded, but do know when the person had their most recent visit with the health system, we use this date. It’s not perfect, but neither are race data. And this is the best approximation taking into consideration the requirements of the CDM and limitations of the source data.

What do I do with flavors of NULL?

•If your data include race source values such as: ‘unknown’, ‘unspecified’, ‘patient refused to answer’, etc.; don’t bring these data into the CDM. *Example:* If a person only has one race source value and it is a flavor of NULL, then PERSON.race_concept_id = 0

•IF your data have one valid race source value and one flavor of NULL race source value, THEN create only one record in the Person table. *Example:* There are two race source values for a person: ‘American Indian or Alaskan Native’ and ‘Unknown’. PERSON.race_concept_id = 8657

Fonte: https://ohdsi.github.io/Themis/multiple_races_per_person.html

5.6.2 Race and Ethnicity

These follow the definitions of how the US government defines this. Ethnicity is a differentiation of Hispanic or non-Hispanic populations, which can have any race. Race is divided into the common 5 top races, which have ethnicities as their hierarchical descendants. Mixed races are not included

Fonte: https://ohdsi-github-io.translate.goog/TheBookOfOhdsi/StandardizedVocabularies.html?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

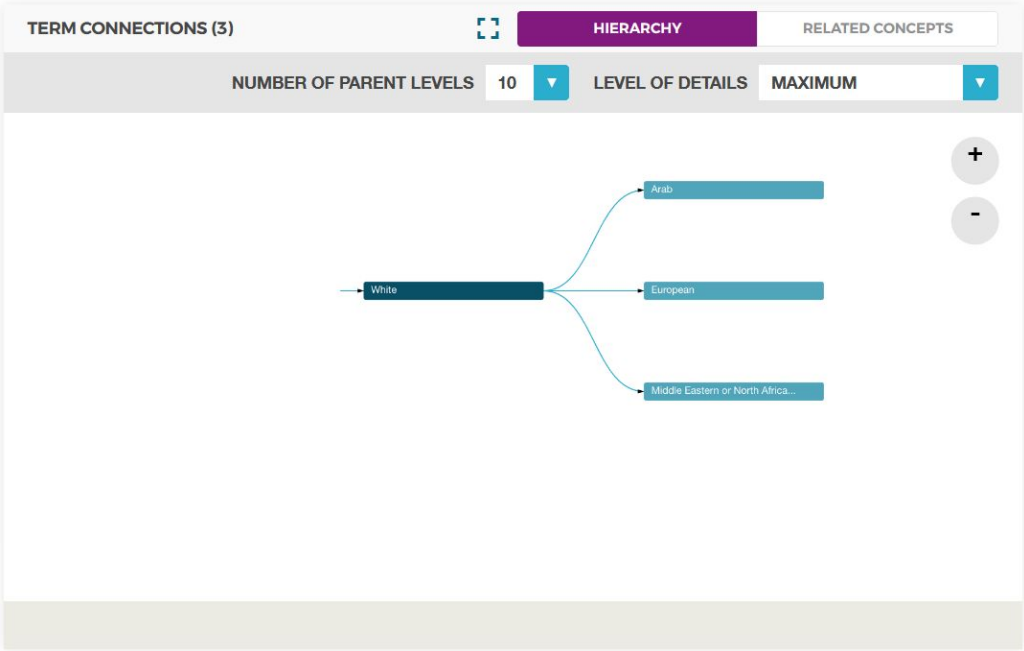
Sugestão: não aceitamos múltiplas entradas nesses campos por agora

2 - Para lidar com múltiplas entradas *race* ele sugere o uso de um vocabulário Non-Standard no race_concept_id, que seria “Multiple_Race”. Parece contraditório com 1. Acredito que podemos usar Non-Standard para *concepts* que indicam que o concept_id Standard estará localizado em outra entrada/tabela (nesse caso os *concepts* Standard de *race* estarão na tabela Observation)

Contextualização race_concept_id (Person)

Os *concepts* de race_concept_id satisfazem certa hierarquia entre eles, como mostrado no exemplo do *concept White*.

Fonte: https://athena.ohdsi.org/search-terms/terms/8527/graph?fullscreen=false&levels=10&standardsOnly=false&zoomLevel=4



ethinicity_concept_id (Person)

Vocabulário Ethinicity (subconjunto com concept -> Standard e validity -> Valid)

Localizado no arquivo ethinicity_concept_id (2 entradas)

Sugestão: É mais estadunidense, id = 0 em todos.

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
38003563	Hispanic	Hispanic or Latino	Ethnicity	Standard	Valid	Ethnicity	Ethnicity
38003564	Not Hispanic	Not Hispanic or Latino	Ethnicity	Standard	Valid	Ethnicity	Ethnicity

Contextualização ethinicity_concept_id (Person)

User guide

This field captures Ethnicity as defined by the Office of Management and Budget (OMB) of the US Government: it distinguishes only between “Hispanic” and “Not Hispanic”. Races and ethnic backgrounds are not stored here.

ETL Conventions

Only use this field if you have US-based data and a source of this information. Do not attempt to infer Ethnicity from the race or ethnic background of the Person. Accepted ethnicity concepts

Fonte: https://ohdsi.github.io/CommonDataModel/cdm54.html#person

5.6.2 Race and Ethnicity

These follow the definitions of how the US government defines this. Ethnicity is a differentiation of Hispanic or non-Hispanic populations, which can have any race.

Fonte: https://ohdsi-github-io.translate.goog/TheBookOfOhdsi/StandardizedVocabularies.html?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

speciality_concept_id (Provider)

Vocabulário Provider (subconjunto com concept -> Standard e validity -> Valid)

Localizado no arquivo speciality_concept_id.csv (6 entradas)

Sugestão: não colocaria Counselor e Service Provider por enquanto. O que conceitualmente parece mais próximo do ACS parece ser o Allied Health

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
33003	OMOP5117445	Service Provider	Provider	Standard	Valid	Provider	Provider
32578	OMOP4822447	Counselor	Provider	Standard	Valid	Provider	Provider
32581	OMOP4822446	Nurse	Provider	Standard	Valid	Provider	Provider
32577	OMOP4822444	Physician	Physician Specialty	Standard	Valid	Provider	Provider
32580	OMOP4822445	Allied Health Professional	Provider	Standard	Valid	Provider	Provider
33005	OMOP5117448	Psychiatry or Neurology	Physician Specialty	Standard	Valid	Provider	Provider

speciality_concept_id (Provider)

Vários vocabulários tem campos para provider (10 vocabulários com 2383 *concepts*, com 741 Standard) -> existe hierarquia entre eles

Vocabulários:

- Standard -> ABMS - HES Specialty - Medicare Specialty - NUCC - Provider
- Non-standard -> ABMS - CDISC - HES Specialty - Medicare Specialty - NUCC - Nebraska Lexicon - Read - SNOMED - Supplier

Contextualização speciality_concept_id (Provider)

User guide

This field either represents the most common specialty that occurs in the data or the most specific concept that represents all specialties listed, should the provider have more than one. This includes physician specialties such as internal medicine, emergency medicine, etc. and allied health professionals such as nurses, midwives, and pharmacists.

ETL Conventions

If a Provider has more than one Specialty, there are two options: 1. Choose a concept_id which is a common ancestor to the multiple specialties, or, 2. Choose the specialty that occurs most often for the provider. Concepts in this field should be Standard with a domain of Provider. Accepted Concepts.

Fonte: https://ohdsi.github.io/CommonDataModel/cdm54.html#person

5.6.7 Providers and Specialties

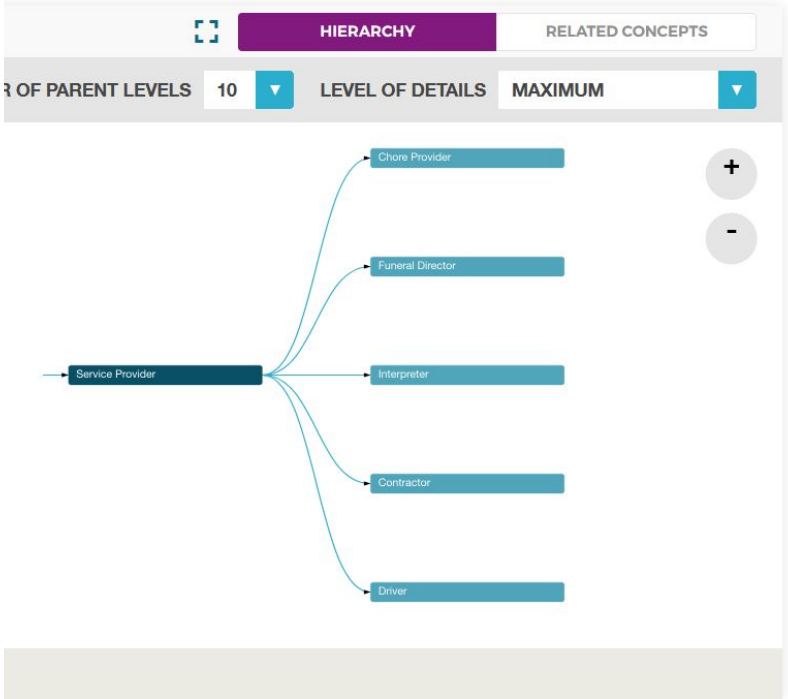
Any human provider is defined in the provider domain. These can be medical professionals such as doctors and nurses, but also non-medical providers like optometrists or shoemakers. Specialties are descendants of the provider “Physician.” Care Sites cannot carry a specialty, even though they are often defined by the specialty of their main staff (“Surgical department”).

Fonte: https://ohdsi-github-io.translate.goog/TheBookOfOhdsi/StandardizedVocabularies.html?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

Contextualização speciality_concept_id (Provider)

As categorias para especialidades são altamente hierarquizadas. Então para a opção médico existem muitas subcategorias, por exemplo. Para ver se o conceito faz sentido, é necessário ver se os conceitos abaixo deles estão conectados com a ideia pensada. Por exemplo, a figura ao lado mostra os níveis do conceito “Service Provider”, é possível notar que se trata de um tipo de profissional não tão ligado com a saúde diretamente.

Fonte:
https://athena.ohdsi.org/search-terms/terms/33003/graph?fullscreen=false&levels=10&standardsOnly=false&zoomLevel=4



country_concept_id (Location)

Vocabulário OSM (subconjunto com concept -> Standard | domain -> Geography | validity -> Valid | Class -> 2nd level)

OSM tem vários níveis, quanto maior o nível, mais específico parece ser a localização. Importante é o Brazil (id = 41892130)

Localizado no arquivo country_concept_id.csv (19 entradas, NÃO temos todos os países)

41987173	365331	Italy	2nd level	Standard	Valid	Geography	OSM
42030997	52822	Sweden	2nd level	Standard	Valid	Geography	OSM
42032171	2323309	Netherlands	2nd level	Standard	Valid	Geography	OSM
42020824	1311341	Spain	2nd level	Standard	Valid	Geography	OSM
41892130	59470	Brazil	2nd level	Standard	Valid	Geography	OSM

country_concept_id (Location)

Vários vocabulários tem concepts para location (5 vocabulários com 204,738 *concepts*!, com 203,563 sendo Standard)

Bastante confuso, pois parece que os conceitos Standard e Valid de diferentes países estão em diferentes Vocabulários. (Os outros países estão no Vocabulário SNOMED / Standard / Geography)

Vocabulários:

- Standard -> OSM - SNOMED - US Census
- Non-standard -> CDISC - MeSH - OSM - SNOMED - US Census

Contextualização country_concept_id (Location)

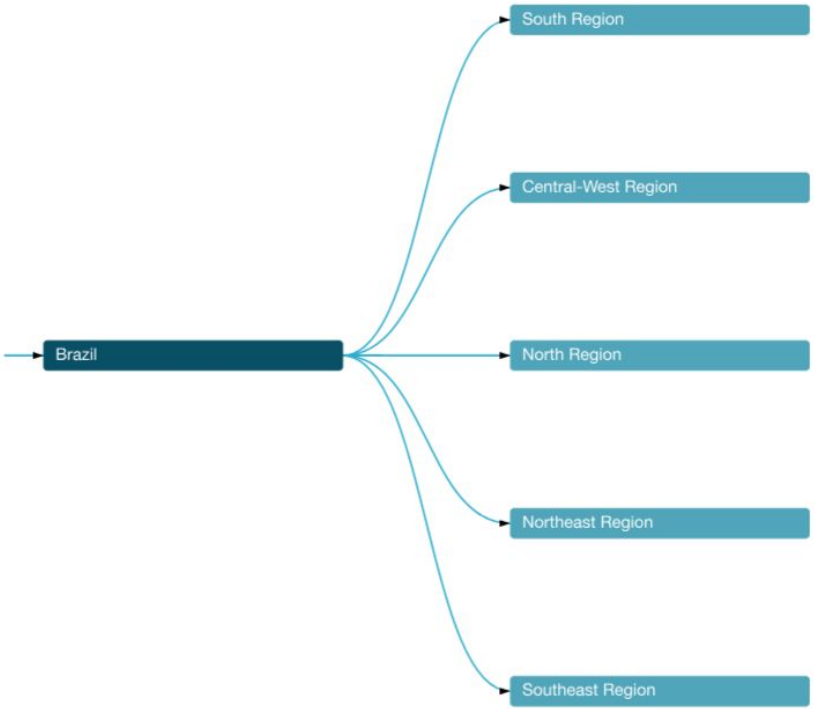
User guide

The Concept Id representing the country. Values should conform to the Geography domain.

Fonte:
https://ohdsi.github.io/CommonDataModel/cdm54.html#person

É possível utilizar algumas parte mais específica do Brasil caso necessário e ilustrado ao lado

Fonte:
https://athena.ohdsi.org/search-terms/terms/41892130/graph?fullscreen=false&levels=10&standardsOnly=false&zoomLevel=4



place_of_service_concept_id (Care_site)

Vocabulário Visit (subconjunto com concept -> Standard | validity -> Valid)

Localizado no arquivo place_of_service_concept_id.csv (20 entradas). É utilizado para definir o principal “serviço” fornecido por aquela localização.

Existe outro conjunto do domínio Place of service (Vocabulário Nebraska Lexicon, Non-standard 330 entradas), que pode confundir. Não é o padrão para esta entrada. Essa confusão pode surgir ao não ler atentamente a descrição em sobre ETL (Fonte: https://github.com/OHDSI/Vocabulary-v5.0/wiki/General-Structure,-Download-and-Use)

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
581478	OMOP4822457	Ambulance Visit	Visit	Standard	Valid	Visit	Visit
581476	OMOP4822459	Home Visit	Visit	Standard	Valid	Visit	Visit
9201	IP	Inpatient Visit	Visit	Standard	Valid	Visit	Visit
32036	OMOP4822461	Laboratory Visit	Visit	Standard	Valid	Visit	Visit

Contextualização place_of_service_concept_id (Care_site)

How to populate the place_of_service_concept_id in the CARE_SITE table

Issue summary

If a care site has multiple types of visits or care associated with it, it is unclear which place of service concept should be used to appropriately represent the setting of care.

The ratified convention

Choose the concept in the visit domain that best represents the setting in which healthcare is provided in the Care Site. If most visits in a Care Site are Inpatient, then the place_of_service_concept_id should represent Inpatient. If information is present about a unique Care Site (e.g. Pharmacy) then a Care Site record should be created.

It is also possible to set a specific visit_concept_id in the VISIT_OCCURRENCE table that represents the exact setting of care for that visit, which may differ from the place_of_service_concept_id in the CARE_SITE table. For example, a patient may go to an inpatient hospital for an outpatient procedure. In this case, the hospital entry in the care_site table should identify it as an inpatient hospital but the specific visit_concept_id for the outpatient procedure should identify that it was an outpatient encounter.

Link to DQD check

No. The isPrimaryKey check will make sure the care site ids are not duplicated in the CARE_SITE table but there is no explicit check for the same care_site_source_value and different place_of_service_concept_ids.

Fonte: https://ohdsi.github.io/Themis/mapping_place_of_service_concept.html

User Guide

This is a high-level way of characterizing a Care Site. Typically, however, Care Sites can provide care in multiple settings (inpatient, outpatient, etc.) and this granularity should be reflected in the visit.

ETL Conventions

Choose the concept in the visit domain that best represents the setting in which healthcare is provided in the Care Site. If most visits in a Care Site are Inpatient, then the place_of_service_concept_id should represent Inpatient. If information is present about a unique Care Site (e.g. Pharmacy) then a Care Site record should be created. Accepted Concepts. For information about how to populate this field please see the THEMIS Conventions.

Fonte: https://ohdsi.github.io/CommonDataModel/cdm54.html#care_site

Contextualização place_of_service_concept_id (Care_site)

5.6.6 Visits and Services

Visits concepts define the nature of healthcare encounters. In many source systems they are called Place of Service, denoting some organization or physical structure, such as a hospital. In others, they are called services. These also differ between countries, and their definition is hard to obtain. Care sites are often specializing on one of few visits (XYZ Hospital), but still should not be defined by them (even in XYZ hospital patients might encounter non-hospital visits).

Fonte: https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html

language_concept_id (Concept_synonym)

Vocabulário SNOMED (subconjunto com concept -> Standard | domain -> Language)

Por enquanto, só vamos precisar do Portuguese language. (id = 4181536)

Localizado no arquivo language_concept_id.csv (839 entradas)

ID ▼	CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
4181536	297504001	Portuguese language	Qualifier Value	Standard	Valid	Language	SNOMED

language_concept_id (Concept_synonym)

Três vocabulários tem concepts para language (3 vocabulários com 881 *concepts*, com 840 sendo Standard). SNOMED é o principal

Vocabulários:

- Standard -> Language - SNOMED
- Non-standard -> SNOMED - MeSH