

***PHENOTYPING:***  
**Herramientas OMOP para**  
**definir problemas de salud II -**  
***CodelistGenerator***

Agustina Giuliadori, [agiuliadori@idiapjgol.org](mailto:agiuliadori@idiapjgol.org)  
Marzo 2025

# ***Introducción a CodelistGenerator***



Es una librería de R que puede ser muy útil en el proceso de **FENOTIPADO**.

Nos permite realizar búsquedas sistemáticas de conceptos a partir de palabras claves.

Utiliza la información de las tablas de vocabulario estandarizado de OMOP.

## **Contenido de la presentación:**

1. Tablas de vocabulario en OMOP
2. CodelistGenerator

# ***Vocabulario en O.MOP***

Phenotyping: CodelistGenerator

# ***1. Vocabulario en OMOP***

## **Características de OMOP CDM**

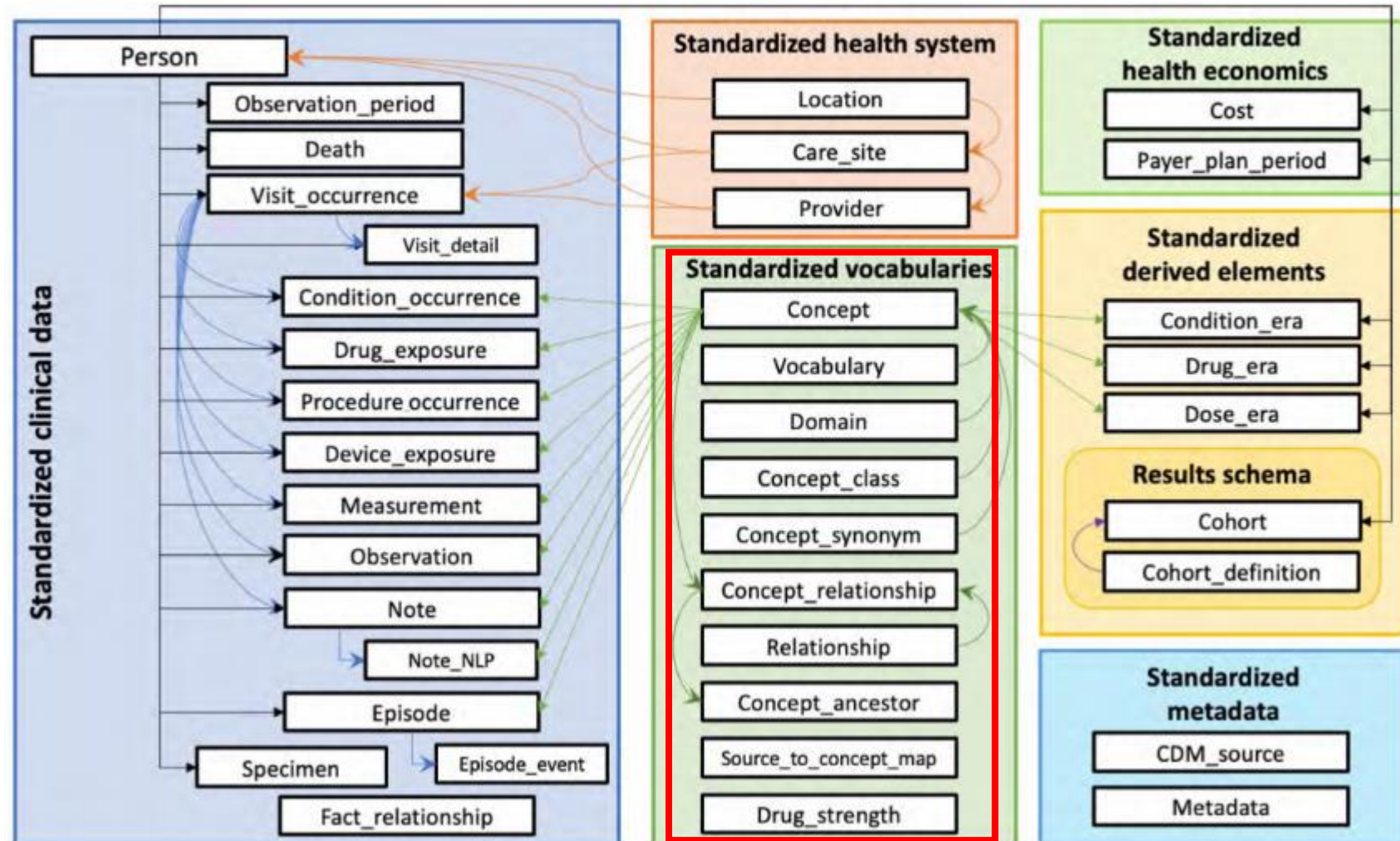
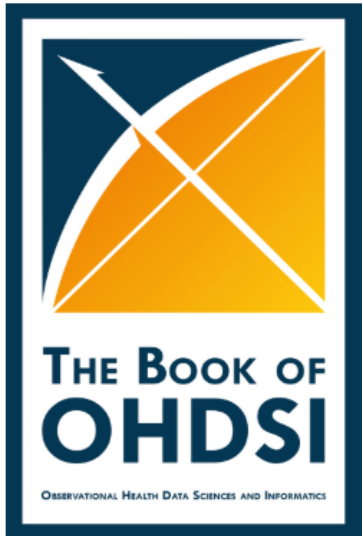
- **Uso de vocabularios estandarizados:** permite garantizar que los registros de salud tengan una representación uniforme y consistente. Hay un **repositorio online** disponible para acceder a los vocabularios estandarizados.
- **Datos normalizados y representados mediante conceptos:** se identifican por medio de identificadores únicos (CONCEPT\_ID) provenientes de vocabularios estandarizados.
- **Mantenimiento de códigos fuentes:** Aunque todos los códigos se convierten a vocabularios estándar, el modelo también guarda los códigos originales de la fuente.

# ***1. Vocabulario en OMOP***

Los **Vocabularios Estandarizados** contienen todos los conjuntos de códigos, terminologías, vocabularios, nomenclaturas, clasificaciones, etc. que permiten:

- Generar la base de datos OMOP (estandarizada) desde la FUENTE.
- Buscar, consultar y extraer los datos transformados (estandarizados).
- Navegar entre jerarquías y relaciones entre diferentes conceptos.
- Interpretar el significado de los datos.

# *1. Vocabularios en OMOP*



# *1. Vocabularios en OMOP*

## Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

- CONCEPTOS: representan entidades clínicas únicas

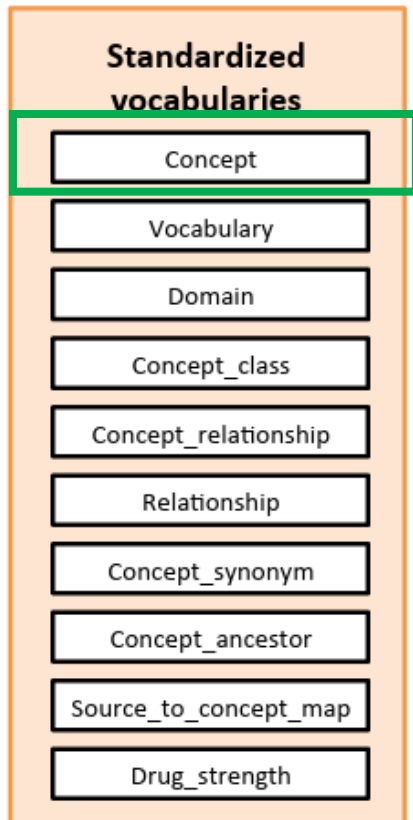
- RELACIONES: conectan conceptos (ej: debido a, mapeado desde)

- JERARQUÍA: ir desde un concepto general (padre) a más específicos (hijos).

- MAPEO: mapeo entre códigos fuentes y estándar.

# ***1.1. CONCEPT***

Todas las instancias del cdm utilizan los mismos identificadores y vocabularios, lo que permite la interoperabilidad en OMOP.



Columna	Descipción
concept_id	Identificador único del concepto
concept_name	Nombre del concepto
domain_id	Ej: "Condition", "Drug", "Measurement", etc.
Vocabulary_id	Ej: "SNOMED", "RxNorm", "LOINC"
Concept_class_id	Ej: Tipo de concepto Ej: "Procedure"
Standard_concept	"S" Standard, "C" classification
Concept_code	Código del concepto en su vocabulario.
Valid start/end date	Fecha de validez



## 1.1. CONCEPT

Todas las instancias del cdm utilizan los mismos identificadores y vocabularios, lo que permite la interoperabilidad en OMOP.

cdm\$concept

### Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

```
> cdm$concept
# Source: table<main.concept> [?? x 10]
# Database: DuckDB v1.1.1 [agiuliodori@windows 10 x64:R 4.4.0/C:/Users/agiuliodori/AppData/Local/Temp/RtmpUzazu9/file42047ce744c2ckdb]
  concept_id concept_name      domain_id vocabulary_id concept_class_id standard_concept concept_code valid_start_date
  <int>      <chr>              <chr>      <chr>          <chr>          <chr>          <chr>      <date>
1  45756805 Pediatric Cardiology    Provider ABMS          Physician Speci... S          OMOP4821938 1970-01-01
2  45756804 Pediatric Anesthesiology Provider ABMS          Physician Speci... S          OMOP4821939 1970-01-01
3  45756803 Pathology-Anatomic / Path... Provider ABMS          Physician Speci... S          OMOP4821940 1970-01-01
4  45756802 Pathology - Pediatric    Provider ABMS          Physician Speci... S          OMOP4821941 1970-01-01
5  45756801 Pathology - Molecular Gen... Provider ABMS          Physician Speci... S          OMOP4821942 1970-01-01
6  45756800 Microbiology      Provider ABMS          Physician Speci... S          OMOP4821943 1970-01-01
7  45756799 Pathology - Hematology    Provider ABMS          Physician Speci... S          OMOP4821944 1970-01-01
8  45756798 Pathology - Forensic      Provider ABMS          Physician Speci... S          OMOP4821945 1970-01-01
9  45756797 Pathology - Clinical      Provider ABMS          Physician Speci... S          OMOP4821946 1970-01-01
10 45756796 Pathology - Chemical     Provider ABMS          Physician Speci... S          OMOP4821947 1970-01-01
# i more rows
```

# 1. 2. CONCEPT\_RELATIONSHIP

<https://athena.ohdsi.org>

Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

SEARCH

DOWNLOAD

Agustina Giuliodori Picco

?

←

Bronchitis

DETAILS

Domain ID	Condition
Concept Class ID	Disorder
Vocabulary ID	SNOMED
Concept ID	256451
Concept code	32398004
Validity	Valid
Concept	Standard

LANGUAGE	SYNONYM	CONCEPT
English	Bronchitis (disorder)	
Spanish	bronquitis	
Spanish	bronquitis (trastorno)	

TERM CONNECTIONS (76)

RELATIONSHIP

RELATES TO

CONCEPT ID

VOCABULARY

Active possibly\_equivalent\_to inactive (SNOMED)

Acute lower respiratory tract infection

40316061

SNOMED

Acute lower respiratory tract infection

40345696

SNOMED

Active was\_a inactive (SNOMED)

Bronchitis NOS

3538386

SNOMED

Bronchitis NOS

40316590

SNOMED

Bronchitis unspecified

4148123

SNOMED

Associated finding of (SNOMED)

FH: Bronchitis

4169827

SNOMED

H/O: bronchitis

4077220

SNOMED

Concept replaces

Bronchitis NOS

40316590

SNOMED

Has Module

SNOMED CT core

40642539

SNOMED

Has associated morphology (SNOMED)

Inflammatory morphology

4253008

SNOMED

# 1. 2. CONCEPT\_RELATIONSHIP

## Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

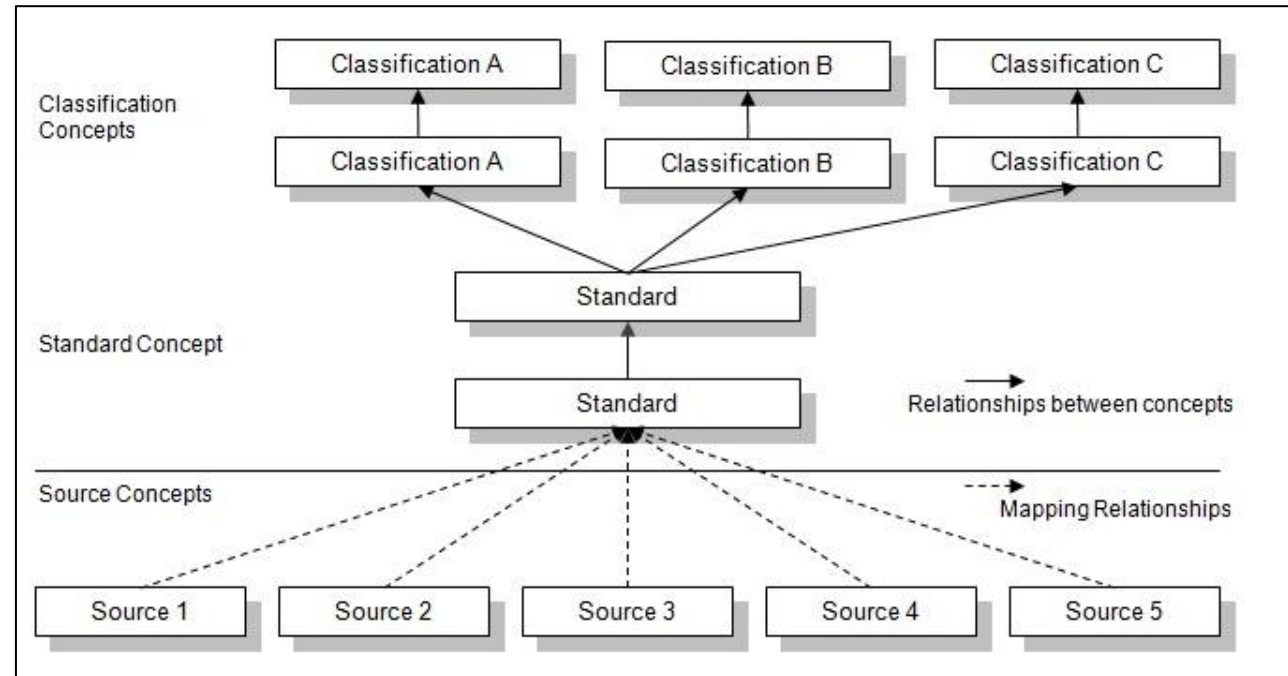
Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength



<https://www.ohdsi.org/data-standardization/>

# 1.2. CONCEPT\_RELATIONSHIP

## Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

cdm\$concept\_relationship

```
> cdm$concept_relationship
# Source:   table<main.concept_relationship> [?? x 6]
# Database: DuckDB v1.1.1 [agiuliodori@windows 10 x64:R 4.4.0/C:\Users\agiuliodori\AppData\Local\ckdb]
```

	concept_id_1 <int>	concept_id_2 <int>	relationship_id <chr>	valid_start_date <date>	valid_end_date <date>	invalid_reason <chr>
1	35804314	912065	Has modality	2021-01-26	2099-12-31	NA
2	35804314	42542145	Has accepted use	2019-08-29	2099-12-31	NA
3	35804314	42542145	Is current in	2019-08-29	2099-12-31	NA
4	35804327	35803584	Has context	2019-05-27	2099-12-31	NA
5	35804327	42542145	Has accepted use	2019-08-29	2099-12-31	NA
6	35804327	42542145	Is historical in	2019-08-29	2099-12-31	NA
7	35804329	35803584	Has context	2019-05-27	2099-12-31	NA
8	35804329	42542145	Has accepted use	2019-08-29	2099-12-31	NA
9	35804329	42542145	Is historical in	2019-08-29	2099-12-31	NA
10	35804338	35803584	Has context	2019-05-27	2099-12-31	NA

# 1 more rows

# 1. 2. CONCEPT\_RELATIONSHIP

cdm\$concept\_relationship

## Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

```
cdm$concept_relationship %>%
```

```
  filter(concept_id_1 == 256451) %>%
```

Bronchitis standard concept\_id

```
  left_join(cdm$concept %>% select(concept_id, concept_name),
    by = c("concept_id_2" = "concept_id"))
```

```
# Source:   SQL [?? x 4]
```

```
# Database: postgres [lperez@10.80.192.21:5432/omop232t]
```

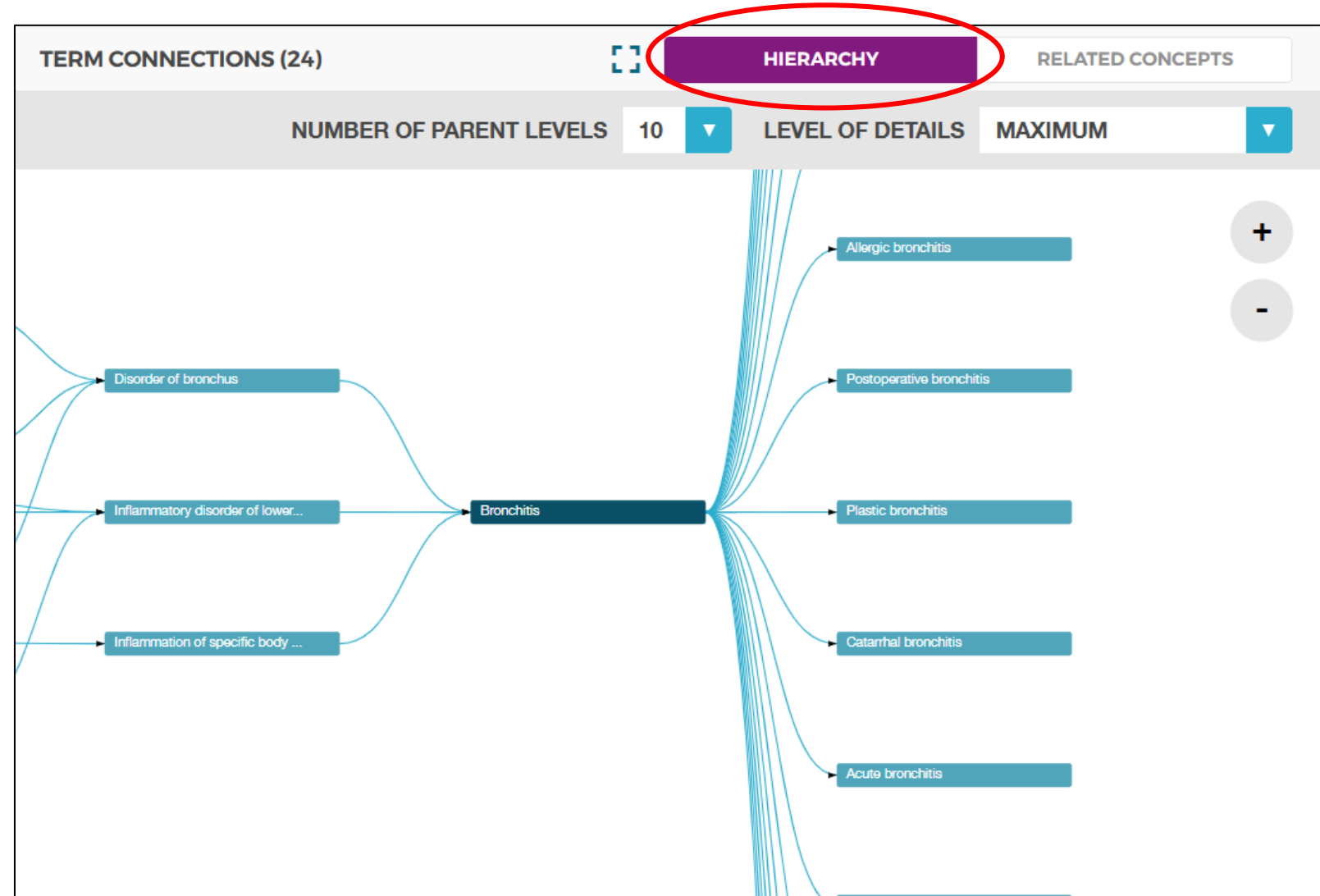
	concept_id_1 <int>	concept_id_2 <int>	relationship_id <chr>	concept_name <chr>
1	256451	4077220	Value mapped from	H/O: bronchitis
2	256451	3538386	Concept was_a from	Bronchitis NOS
3	256451	3538386	Mapped from	Bronchitis NOS
4	256451	40642537	Has status	Defined
5	256451	40642539	Has Module	SNOMED CT core
6	256451	256722	Subsumes	Bronchopneumonia
7	256451	256896	Subsumes	Tuberculosis of bronchus
8	256451	4077220	Asso finding of	H/O: bronchitis
9	256451	4148123	Concept was_a from	Bronchitis unspecified
10	256451	4148123	Mapped from	Bronchitis unspecified

```
# i more rows
```

# 1.3. CONCEPT\_ANCESTOR

**Standardized vocabularies**

- Concept
- Vocabulary
- Domain
- Concept\_class
- Concept\_relationship
- Relationship
- Concept\_synonym
- Concept\_ancestor**
- Source\_to\_concept\_map
- Drug\_strength



# 1.3. CONCEPT\_ANCESTOR

cdm\$concept\_ancestor

## Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

```
cdm$concept_ancestor %>%
```

```
  filter(ancestor_concept_id == 256451) %>%
```

Bronchitis standard concept\_id

```
  left_join(cdm$concept %>% select(concept_id, concept_name),
    by = c("descendant_concept_id" = "concept_id"))
```

```
# Source:   SQL [?? x 5]
```

```
# Database: postgresql [lperez@10.80.192.21:5432/omop232t]
```

	ancestor_concept_id	descendant_concept_id	min_levels_of_separation	max_levels_of_separation	concept_name
	<int>	<int>	<int>	<int>	<chr>
1	256451	4233924	1	1	Pneumococcal br...
2	256451	4112677	2	3	Acute bronchiti...
3	256451	439298	3	3	Bronchitis and ...
4	256451	46269725	2	4	Bronchopneumoni...
5	256451	4135197	2	2	Hypostatic bron...
6	256451	37206131	2	2	Bronchitis co-o...
7	256451	4049964	3	3	Acute bacterial...
8	256451	4148204	2	2	Acute tracheobr...
9	256451	37116690	2	2	Chronic suppura...
10	256451	46269709	2	2	Bronchopneumoni...

```
# 1 more rows
```

# 1. 4. SOURCE\_TO\_CONCEPT\_MAP

CÓDIGOS FUENTE



CÓDIGOS ESTÁNDAR

Standardized  
vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength



Bronchitis, not specified as acute or chronic

## DETAILS

Domain ID	Condition
Concept Class ID	3-char billing code
Vocabulary ID	ICD10CM <span>?</span>
Concept ID	35208013
Concept code	J40
Validity	Valid
Concept	Non-standard
Valid start	01-Jan-2007
Valid end	31-Dec-2099

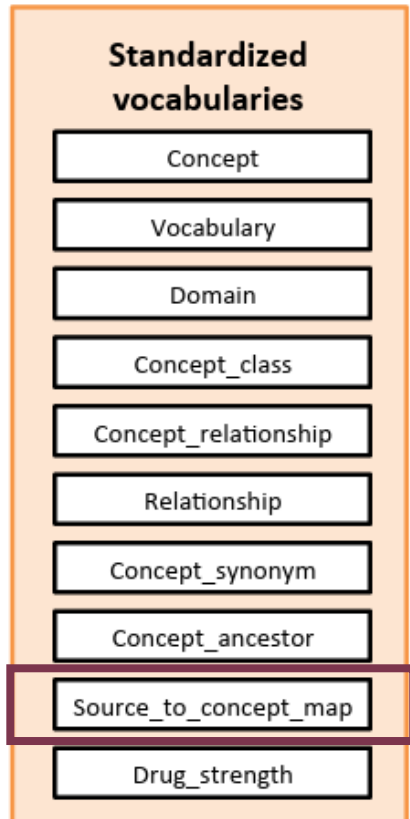
## TERM CONNECTIONS (1)

RELATIONSHIP	RELATES TO	CONCEPT ID	VOCABULARY
Non-standard to Standard map (OMOP)	Bronchitis	256451	SNOMED



# 1. 4. SOURCE\_TO\_CONCEPT\_MAP

## SOURCE\_TO\_CONCEPT\_MAP



**CÓDIGOS FUENTE**

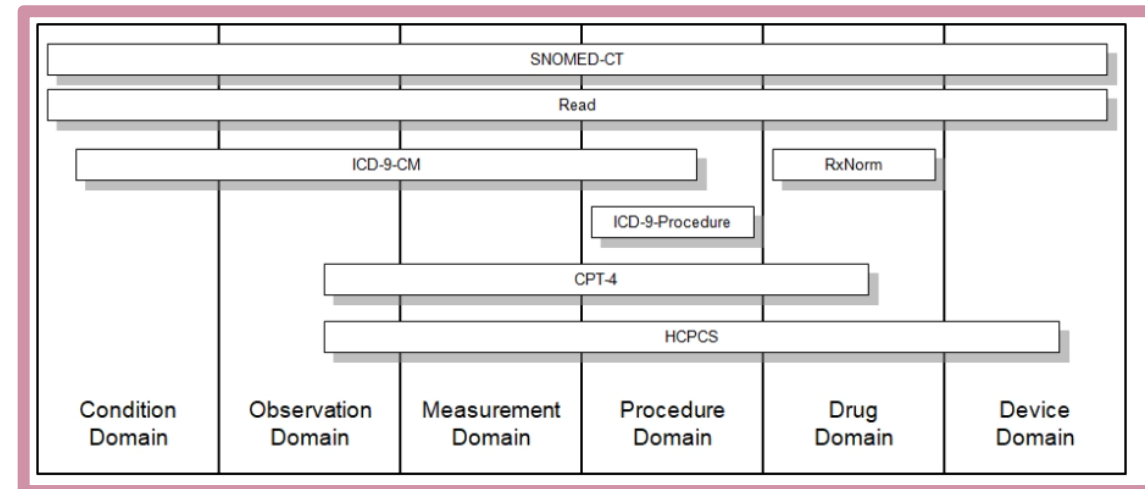


**CÓDIGOS ESTÁNDAR**

Ej: SIDIAP utiliza:

- **ICD10-CM**
- **ATC-7**

**OMOP CDM** utiliza como vocab. estándar:



# 1. 4. SOURCE\_TO\_CONCEPT\_MAP

## SOURCE\_TO\_CONCEPT\_MAP

CÓDIGOS FUENTE



CÓDIGOS ESTÁNDAR

cdm\$source\_to\_concept\_map

### Standardized vocabularies

Concept

Vocabulary

Domain

Concept\_class

Concept\_relationship

Relationship

Concept\_synonym

Concept\_ancestor

Source\_to\_concept\_map

Drug\_strength

```
> cdm$source_to_concept_map %>% filter(source_code == "J40")
# Source:   SQL [1 x 9]
# Database: postgres [lperez@10.80.192.21:5432/omop232t]
  source_code source_concept_id source_vocabulary_id source_code_description target_concept_id target_vocabulary_id valid_start_date
  <chr>          <int> <chr>          <chr>          <int> <chr>          <date>
1 J40              0 ICD10          bronchitis, not classi...    256451 SNOMED          1970-01-01
# i 2 more variables: valid_end_date <date>, invalid_reason <chr>
```

```
> cdm$source_to_concept_map %>% filter(target_concept_id == 256451)
# Source:   SQL [2 x 9]
# Database: postgres [lperez@10.80.192.21:5432/omop232t]
  source_code source_concept_id source_vocabulary_id source_code_description target_concept_id target_vocabulary_id
  <chr>          <int> <chr>          <chr>          <int> <chr>
1 J40              0 ICD10          bronchitis, not classified ...    256451 SNOMED
2 490              0 ICD9CM          Bronchitis not specified as...
```

# *CodelistGenerator*

Phenotyping: CodelistGenerator

## 2. CodelistGenerator vs ATLAS

ATLAS

Home

Data Sources

Search

Concept Sets

Cohort Definitions

Characterizations

Cohort Pathways

Incidence Rates

Profiles

Estimation

Prediction

Reusables

Jobs

Configuration

Feedback

Apache 2.0  
open source software  
provided by  
OHDSI  
join the journey

English

Search

bronchitis

Advanced Options

View record count for: Evidence network counts

Filter: Search...

Showing 1 to 50 of 715 entries

Show columns

Copy

CSV

Show 50 entries

Vocabulary

SNOMED (193)

Nebraska Lexicon (158)

LOINC (96)

Read (61)

ICD10 code (59)

Class

Disorder (156)

Clinical Finding (125)

Read (61)

ICD10 code (59)

Domain

Condition (470)

Observation (90)

Drug (77)

Measurement (71)

Standard Concept

Non-Standard (530)

Standard (155)

Classification (30)

Invalid Reason

Valid (515)

Invalid (200)

Has Records

false (566)

true (149)

Has Descendant Records

false (565)

true (150)

Has Person Count

false (649)

true (66)

Has Descendant Person Count

false (647)

true (68)

	Id	Code	Name	Class	RC	DRC	PC	DPC	Domain	Vocabulary
	256451	32398004	Bronchitis	Disorder	48,582,260	245,336,740	26,896,600	109,386,550	Condition	SNOMED
	260139	10509002	Acute bronchitis	Disorder	121,682,030	123,472,210	59,614,780	60,658,730	Condition	SNOMED
	255841	63480004	Chronic bronchitis	Disorder	7,128,460	55,906,030	2,279,690	12,178,910	Condition	SNOMED
	4028970	13617004	Tracheobronchitis	Disorder	9,320	13,833,240	7,850	8,027,880	Condition	SNOMED
	4312602	85915003	Laryngotracheobronchitis	Disorder	43,330	13,819,790	29,230	8,017,380	Condition	SNOMED
	258780	185086009	Emphysematous bronchitis	Disorder	10,160,460	41,950,620	2,265,440	7,603,280	Condition	SNOMED
	4137525	425748003	Acute exacerbation of chronic bronchitis	Disorder	17,660	31,815,390	5,050	5,347,300	Condition	SNOMED
	46270376	293241000119100	Acute exacerbation of chronic obstructive bronchitis	Disorder	31,790,160	31,790,160	5,337,840	5,337,840	Condition	SNOMED
	261889	61937009	Simple chronic bronchitis	Disorder	4,767,360	4,767,360	1,641,250	1,641,250	Condition	SNOMED
	4208807	312371005	Acute infective bronchitis	Disorder	3,370	1,724,910	2,010	999,970	Condition	SNOMED
	4309350	84409004	Fetid chronic bronchitis	Disorder	30	2,009,400	10	637,740	Condition	SNOMED
	4049964	233598009	Acute bacterial bronchitis	Disorder	26,910	918,840	20,570	612,790	Condition	SNOMED
	4105601	29591002	Purulent bronchitis	Disorder	11,980	1,411,290	1,680	495,380	Condition	SNOMED
	257905	74417001	Mucopurulent chronic bronchitis	Disorder	1,393,170	1,393,170	491,850	491,850	Condition	SNOMED
	4058712	16146001	Viral bronchitis	Disorder	4,400	801,450	3,290	386,880	Condition	SNOMED
	4052543	233601004	Acute viral bronchitis	Disorder	5,660	796,320	4,570	383,130	Condition	SNOMED
	4051332	233599001	Acute mycoplasmal bronchitis	Disorder	483,660	483,660	331,120	331,120	Condition	SNOMED
	4112521	195728004	Acute bronchitis due to rhinovirus	Disorder	350,590	350,590	166,530	166,530	Condition	SNOMED
	4195736	79479005	Respiratory syncytial virus bronchitis	Disorder	530	331,750	350	155,690	Condition	SNOMED
	4110484	195727009	Acute respiratory syncytial virus bronchitis	Disorder	331,220	331,220	155,340	155,340	Condition	SNOMED
	4112826	195953005	Mixed simple and mucopurulent chronic bronchitis	Disorder	616,200	616,200	145,880	145,880	Condition	SNOMED
	4112015	195720006	Acute streptococcal bronchitis	Disorder	142,270	142,350	90,540	90,610	Condition	SNOMED
	46273460	10625791000119101	Bronchitis due to chemical	Disorder	20	139,550	10	72,830	Condition	SNOMED

## ***2. CodelistGenerator vs ATLAS***



Ventajas de **CodelistGenerator** comparado con ATLAS:

- Búsqueda de códigos sistemática y eficiente.
- Mayor precisión y consistencia en la búsqueda.
- Permite expandir la búsqueda incluyendo términos sinónimos y jerárquicos.
- Escalabilidad: ideal cuando necesitamos múltiples listas de códigos.

¿Qué información necesita **CodelistGenerator** para hacer la búsqueda?

**vocabularios estandarizados**



## ***2. CDM for CodelistGenerator***

### **Crear el CDM:**

1. Desde conexión a BD OMOP
2. Desde Vocabularios (descargados de Athena)

```
library(DBI)
library(CDMConnector)

db <- DBI::dbConnect()

cdm <- CDMConnector::cdmFromCon(
  con = db,
  cdmSchema = "omop",
  writeSchema = "results"
)
```

## 2. CDM for CodelistGenerator



### Crear el CDM:

1. Desde conexión a BD OMOP
2. Desde Vocabularios (descargados de Athena)

SEARCH

DOWNLOAD

Agustina Giuliodori Picco

Show all

SHOW HISTORY

DOWNLOAD VOCABULARIES

<input checked="" type="checkbox"/>	ID (CDM V4.5)	CODE (CDM V5)	NAME	REQUIRED	LATEST UPDATE
<input checked="" type="checkbox"/>		SNOMED	Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO)		01-Feb-2024
<input checked="" type="checkbox"/>	2	ICD9CM	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS)		01-Oct-2014
<input checked="" type="checkbox"/>	3	ICD9Proc	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS)		01-Oct-2014
<input checked="" type="checkbox"/>	4	CPT4	Current Procedural Terminology version 4 (AMA)	EULA required	06-May-2024
<input checked="" type="checkbox"/>	5	HCPCS	Healthcare Common Procedure Coding System (CMS)		01-Jul-2024
<input checked="" type="checkbox"/>	6	LOINC	Logical Observation Identifiers Names and Codes (Regenstrief Institute)		27-Feb-2024
<input checked="" type="checkbox"/>	7	NDFRT	National Drug File - Reference Terminology (VA)		06-Aug-2018
<input checked="" type="checkbox"/>	8	RxNorm	RxNorm (NLM)		06-May-2024
<input checked="" type="checkbox"/>	9	NDC	National Drug Code (FDA and manufacturers)		25-Aug-2024
	10	GPI	Medi-Span Generic Product Identifier (Wolters Kluwer Health)	License required	14-Dec-2017

## 2. *CDM for CodelistGenerator*



### Crear el CDM:

1. Desde conexión a BD OMOP
2. Desde Vocabularios (descargados de Athena)

Base de datos vacía

Escribir tablas en mi bd

Crear CDM

```
library(DBI)
library(CDMConnector)

db <- DBI::dbConnect( duckdb::duckdb("vocab.duckdb"))

concept_athena <- read_delim("CONCEPT.csv")
dbWriteTable(db, "concept", concept_athena , overwrite = TRUE)

cdm <- CDMConnector::cdmFromCon(con = db, cdm_schema = "main", write_schema = "main")
```

Repetir para todas  
las tablas de vocab.



## 2. CDM for CodelistGenerator



### Crear el CDM:

1. Desde conexión a BD OMOP
2. Desde Vocabularios (descargados de Athena)

```

# OMOP CDM reference (duckdb) of cdm
• omop tables: person, observation_period, concept, vocabulary, concept_relationship, concept_synonym, concept_ancestor,
drug_strength
• cohort tables: -
• achilles tables: -
• other tables: -
    
```

```

> cdm$concept
# Source:   table<main.concept> [?? x 10]
# Database: DuckDB v1.1.1 [agiuliodori@windows 10 x64:R 4.4.0/C:\Users\agiuliodori\Documents\SEXUAL_HEALTH\Code\Athena_broad\vocab.d
uckdb]

```

	concept_id	concept_name	domain_id	vocabulary_id	concept_class_id	standard_concept	concept_code	valid_start_date
	<dbl>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<dbl>
1	45756805	Pediatric Cardiology	Provider	ABMS	Physician Speci...	S	OMOP4821938	19700101
2	45756804	Pediatric Anesthesiology	Provider	ABMS	Physician Speci...	S	OMOP4821939	19700101
3	45756803	Pathology-Anatomic / Path...	Provider	ABMS	Physician Speci...	S	OMOP4821940	19700101
4	45756802	Pathology - Pediatric	Provider	ABMS	Physician Speci...	S	OMOP4821941	19700101
5	45756801	Pathology - Molecular Gen...	Provider	ABMS	Physician Speci...	S	OMOP4821942	19700101
6	45756800	Microbiology	Provider	ABMS	Physician Speci...	S	OMOP4821943	19700101
7	45756799	Pathology - Hematology	Provider	ABMS	Physician Speci...	S	OMOP4821944	19700101
8	45756798	Pathology - Forensic	Provider	ABMS	Physician Speci...	S	OMOP4821945	19700101
9	45756797	Pathology - Clinical	Provider	ABMS	Physician Speci...	S	OMOP4821946	19700101
10	45756796	Pathology - Chemical	Provider	ABMS	Physician Speci...	S	OMOP4821947	19700101

```

# i more rows
    
```

## 2. Funciones CodelistGenerator



**CodelistGenerator::getVocabVersion( cdm )**

```
> getVocabVersion(cdm)
[1] "v5.0 30-AUG-24"
```

**CodelistGenerator::getVocabularies( cdm )**

```
> getvocabularies(cdm)
[1] "ABMS" "ATC" "CDM" "CMS Place of Service" "Concept Class"
[6] "Condition Status" "Condition Type" "Cost" "Cost Type" "Currency"
[11] "Death Type" "Device Type" "Domain" "Drug Type" "Episode"
[16] "Ethnicity" "Gender" "HCPCS" "ICD10CM" "ICD9CM"
[21] "ICD9Proc" "Korean Revenue Code" "Language" "LOINC" "Meas Type"
[26] "Medicare Specialty" "Metadata" "NDC" "None" "Note Type"
[31] "NUCC" "Obs Period Type" "Observation Type" "OMOP Extension" "OSM"
[36] "Plan" "Plan Stop Reason" "Procedure Type" "Race" "Relationship"
[41] "Revenue Code" "RxNorm" "RxNorm Extension" "SNOMED" "SOPT"
[46] "SPL" "Sponsor" "Type Concept" "UB04 Point of Origin" "UB04 Pri Typ of Adm"
[51] "UB04 Pt dis status" "UB04 Typ bill" "UCUM" "US Census" "Visit"
[56] "Visit Type" "Vocabulary"
```

**CodelistGenerator::getRelationshipId( cdm )**

```
> getRelationshipId(cdm)
[1] "Asso finding of" "Asso with finding" "Due to of" "Finding asso with" "Followed by" "Follows"
[7] "Has asso finding" "Has due to" "Has manifestation" "Has pathology" "Has temp finding" "Is a"
[13] "Manifestation of" "Mapped from" "Maps to" "Occurs after" "Occurs before" "Pathology of"
[19] "Subsumes" "Temp related to"
```

## ***2. Funciones CodelistGenerator***



### **CodelistGenerator::getDomains( cdm )**

```
> getDomains(cdm)
[1] "Procedure"           "Meas Value"         "Device"              "Race"                "Drug"
[6] "Provider"            "Metadata"           "Ethnicity"           "Gender"              "Type Concept"
[11] "Geography"           "Relationship"        "Specimen"            "Route"               "Currency"
[16] "Observation"         "Plan"               "Condition"           "Visit"               "Condition status"
[21] "Cost"                "Episode"            "Unit"                "Measurement"          "Note"
[26] "Plan Stop Reason"    "Revenue Code"       "Spec Anatomic Site"  "Language"             "Meas value operator"
[31] "Payer"               "Sponsor"
```

### **CodelistGenerator::getDoseForm( cdm )**

```
> getDoseForm(cdm)
[1] "Augmented Topical Cream"           "Augmented Topical Gel"
[3] "Augmented Topical Lotion"          "Augmented Topical Ointment"
[5] "Auto-Injector"                     "Bar Soap"
[7] "Bath Salts"                         "Buccal Film"
[9] "Buccal Tablet"                      "Cartridge"
[11] "Cement"                             "Chewable Bar"
[13] "Chewable Extended Release Oral Tablet" "Chewable Tablet"
[15] "Chewing Gum"                       "Collodion"
[17] "Cutaneous patch"                   "Delayed Release Oral Capsule"
[19] "Delayed Release Oral Granules"      "Delayed Release Oral Tablet"
[21] "Dental Pin"                         "Disintegrating Oral Tablet"
```

## ***2. Funciones CodelistGenerator***



Obtener **códigos candidatos** para mis definiciones basado en palabras claves de inclusión / exclusión:

```
CodelistGenerator::getCandidateCodes(  
    cdm,  
    keywords,  
    exclude = NULL,  
    domains = "Condition",  
    standardConcept = "Standard",  
    searchInSynonyms = FALSE,  
    searchNonStandard = FALSE,  
    includeDescendants = TRUE,  
    includeAncestor = FALSE  
)
```

Concept\_synonym

Concept\_ancestor

## ***2. Funciones CodelistGenerator***



Obtener **códigos candidatos** para mis definiciones basado en palabras claves de inclusión / exclusión:

```
CodelistGenerator::getCandidateCodes(  
  cdm,  
  keywords = "Bronchitis",  
  searchInSynonyms = F,  
  includeDescendants = F)
```

Limiting to domains of interest  
Getting concepts to include  
Search completed. Finishing up.  
✓ 79 candidate concepts identified  
Time taken: 0 minutes and 1 seconds

```
CodelistGenerator::getCandidateCodes(  
  cdm,  
  keywords = "Bronchitis",  
  searchInSynonyms = T,  
  includeDescendants = F)
```

Limiting to domains of interest  
Getting concepts to include  
Adding concepts using synonyms  
Search completed. Finishing up.  
✓ 80 candidate concepts identified  
Time taken: 0 minutes and 2 seconds

## ***2. Funciones CodelistGenerator***



Comparar las dos listas resultantes de la búsqueda:

**CodelistGenerator::compareCodelists**(codes, codes\_with\_synonyms)



	concept_id	concept_name	codelist
76	37311903	Obstructing Aspergillus tracheobronchitis	Both
77	3655117	Bronchitis caused by vapors	Both
78	4320830	Mycoplasmal tracheobronchitis	Both
79	439298	Bronchitis and pneumonitis caused by chemical fumes	Both
80	40482019	Inflammation of bronchus caused by Aspergillus	Only codelist 2

## 2. *Funciones CodelistGenerator*



Uso de la tabla concept\_synonym:

```
cdm$concept_synonym %>% filter(grepl("bronquitis", concept_synonym_name))
```

	concept_id	concept_synonym_name	language_concept_id	concept_name
75	40395539	Laryngotracheobronchitis (disorder)	4180186	Laryngotracheobronchitis
76	40482019	Aspergillus bronchitis	4180186	Inflammation of bronchus caused by Aspergillus
77	4268306	Infectious canine tracheobronchitis (disorder)	4180186	Infectious canine tracheobronchitis
78	3012549	Avian infectious bronchitis; Conventional virus neutralizatio...	4180186	Infectious bronchitis virus serotype [Identifier] in Isolate by ...
79	3021363	Avian infectious bronchitis; DNA NUCLEIC ACID PROBE; DN...	4180186	Infectious bronchitis virus RNA [Presence] in Specimen by Pr...
80	3028275	3 Self-Sustaining Sequence Replication; 3SR SR; Amplif; Am...	4180186	Infectious bronchitis virus RNA [Presence] in Specimen by N...

## 2. *Funciones CodelistGenerator*



Uso de la tabla concept\_synonym:

```
cdm$concept_synonym %>% filter(grepl("bronquitis", concept_synonym_name))
```

	concept_id	concept_synonym_name	language_concept_id	concept_name
75	40395539	Laryngotracheobronchitis (disorder)	4180186	Laryngotracheobronchitis
76	40482019	Aspergillus bronchitis	4180186	Inflammation of bronchus caused by Aspergillus
77	4268306	Infectious canine tracheobronchitis (disorder)	4180186	Infectious canine tracheobronchitis
78	3012549	Avian infectious bronchitis; Conventional virus neutralizatio...	4180186	Infectious bronchitis virus serotype [Identifier] in Isolate by ...
79	3021363	Avian infectious bronchitis; DNA NUCLEIC ACID PROBE; DN...	4180186	Infectious bronchitis virus RNA [Presence] in Specimen by Pr...
80	3028275	3 Self-Sustaining Sequence Replication; 3SR SR; Amplif; Am...	4180186	Infectious bronchitis virus RNA [Presence] in Specimen by N...



## 2. Funciones CodelistGenerator

1. Obtener códigos candidatos para definir “Chlamydia” buscando condiciones:



```
CodeListGenerator::getCandidateCodes(
  cdm = cdm,
  keywords = c("Chlamydia"),
  exclude = c("neonatal"),
  domains = c("Condition"),
  standardConcept = "Standard",
  searchInSynonyms = T,
  searchNonStandard = F,
  includeDescendants = T,
  includeAncestor = F
)
```

```
# A tibble: 81 x 6
  concept_id found_from concept_name domain_id vocabulary_id
  <dbl> <chr> <chr> <chr> <chr>
1 36716495 From initial search Non-ulcerative sexually transmitted infection caused... Condition SNOMED
2 44811657 From initial search Reactive arthritis after chlamydia trachomatis infec... Condition SNOMED
3 45757283 From initial search Venereal disease caused by chlamydia trachomatis Condition SNOMED
4 608803 From initial search Bilateral chlamydial conjunctivitis Condition SNOMED
5 439686 From initial search Urethritis caused by Chlamydia trachomatis Condition SNOMED
6 194288 From initial search Chlamydial infection of anus and rectum Condition SNOMED
7 4096951 From initial search Chlamydial pelvic inflammatory disease Condition SNOMED
8 198076 From initial search Chlamydial peritonitis Condition SNOMED
9 4070540 From initial search Congenital chlamydial pneumonia Condition SNOMED
10 372550 From initial search Chlamydial conjunctivitis Condition SNOMED
# i 71 more rows
```

## 2. Funciones CodelistGenerator

2. Obtener códigos candidatos para definir “Chlamydia” buscando condiciones y medidas:



```
CodeListGenerator::getCandidateCodes (
  cdm = cdm,
  keywords = "Chlamydia",
  exclude = "neonatal",
  domains = c("Condition", "Measurement" ),
  standardConcept = "Standard",
  searchInSynonyms = T,
  searchNonStandard = F,
  includeDescendants = T,
  includeAncestor = F
)
```

```
# A tibble: 663 x 6
  concept_id found_from      concept_name      domain_id vocabulary_id
  <dbl> <chr>      <chr>      <chr>      <chr>
1  3026384 From initial search Chlamydia sp Ag [Presence] in Stool Measurement LOINC
2  3036452 From initial search Chlamydia sp DNA [Presence] in Throat by N... Measurement LOINC
3  3010096 From initial search Chlamydia trachomatis Ag [Presence] in Blo... Measurement LOINC
4  3014693 From initial search Chlamydia trachomatis Ag [Presence] in Uri... Measurement LOINC
5  3050305 From initial search Chlamydia sp Ab [Presence] in Cerebral spi... Measurement LOINC
6  3049931 From initial search Chlamydia trachomatis rRNA [Presence] in B... Measurement LOINC
7  3047699 From initial search Chlamydia sp IgG Ab [Presence] in Body flu... Measurement LOINC
8  3046789 From initial search Chlamydia trachomatis D+E+F+G+H+I+J+K IgM ... Measurement LOINC
9  3047131 From initial search Chlamydia sp Ag [Presence] in Urethra Measurement LOINC
10 3042746 From initial search Chlamydia trachomatis IgG Ab [Presence] in... Measurement LOINC
# i 653 more rows
```

## 2. *Funciones CodelistGenerator*



Obtener lista de conceptos desde una cohorte previamente definida:

**CodelistGenerator::codesFromCohort** (path\_to\_json, cdm)

chlamydia_diagnosis.json: Bloc de notas	syphilis_diagnosis.json: Bloc de notas	gonorrhea_diagnosis.json: Bloc de notas
<pre> {   "ConceptSets": [     {       "id": 0,       "name": "Chl",       "expression": {         "items": [           {             "concept": {               "CONCEPT_CLASS_ID": "Disorder",               "CONCEPT_CODE": "235861001",               "CONCEPT_ID": 40,               "CONCEPT_NAME": "Abscess gonococcal",               "DOMAIN_ID": "Condition",               "INVALID_REASON": "v",               "INVALID_REASON_CAPTION": "Valid",               "STANDARD_CONCEPT": "S",               "STANDARD_CONCEPT_CAPTION": "Standard",               "VOCABULARY_ID": "SNOMED",               "VALID_START_DATE": "2002-01-31",               "VALID_END_DATE": "2099-12-31"             },             "includeDescendants": true,             "includeMapped": false           }         ]       }     }   ],   "include": {     "includeDescendants": true,     "includeMapped": false   } } </pre>	<pre> {   "ConceptSets": [     {       "id": 3,       "name": "syphilis",       "expression": {         "items": [           {             "concept": {               "CONCEPT_CLASS_ID": "Disorder",               "CONCEPT_CODE": "235861001",               "CONCEPT_ID": 40,               "CONCEPT_NAME": "Abscess gonococcal",               "DOMAIN_ID": "Condition",               "INVALID_REASON": "v",               "INVALID_REASON_CAPTION": "Valid",               "STANDARD_CONCEPT": "S",               "STANDARD_CONCEPT_CAPTION": "Standard",               "VOCABULARY_ID": "SNOMED",               "VALID_START_DATE": "2002-01-31",               "VALID_END_DATE": "2099-12-31"             },             "includeDescendants": true,             "includeMapped": false           }         ]       }     }   ],   "include": {     "includeDescendants": true,     "includeMapped": false   } } </pre>	<pre> {   "ConceptSets": [     {       "id": 2,       "name": "gonorrhea",       "expression": {         "items": [           {             "concept": {               "CONCEPT_CLASS_ID": "Disorder",               "CONCEPT_CODE": "235861001",               "CONCEPT_ID": 4341651,               "CONCEPT_NAME": "Abscess gonococcal",               "DOMAIN_ID": "Condition",               "INVALID_REASON": "v",               "INVALID_REASON_CAPTION": "Valid",               "STANDARD_CONCEPT": "S",               "STANDARD_CONCEPT_CAPTION": "Standard",               "VOCABULARY_ID": "SNOMED",               "VALID_START_DATE": "2002-01-31",               "VALID_END_DATE": "2099-12-31"             },             "includeDescendants": true,             "includeMapped": false           }         ]       }     }   ],   "include": {     "includeDescendants": true,     "includeMapped": false   } } </pre>

31 437221 36716495 ...  
080594 4127392 ...  
9 3657843 4294449 ...

## ***2. Funciones CodelistGenerator***



Obtener conceptos mapeados:

```
CodelistGenerator::getMappings(
  candidateCodelist = data.frame(concept_id = c(437221)),
  cdm = cdm,
  nonStandardVocabularies = c("ICD10CM")
)
```



	standard_concept_id	standard_concept_name	standard_vocabulary_id	non_standard_concept_id	non_standard_concept_name	non_standard_concept_code	non_standard_vocabulary_id
1	437221	Chlamydia trachomatis infection	SNOMED	35205604	Sexually transmitted chlamydial infection of other sites	A56.8	ICD10CM

## ***2. Funciones CodelistGenerator***



Otras funciones específicas para fármacos:

**CodelistGenerator:: getDrugIngredientCodes(cdm = cdm, c("acetaminophen"))**

	concept_id	concept_name	domain_id	vocabulary_id	concept_class_id	standard_concept	concept_code
1	40231937	acetaminophen 325 MG / pseudoephedrine hydrochloride 3...	Drug	RxNorm	Clinical Drug	S	1049179
2	40231862	acetaminophen 300 MG / oxycodone hydrochloride 10 MG ...	Drug	RxNorm	Branded Drug Comp	S	1049190
3	40231861	acetaminophen 300 MG / oxycodone hydrochloride 10 MG ...	Drug	RxNorm	Branded Drug	S	1049191
4	40231866	acetaminophen 300 MG / oxycodone hydrochloride 2.5 MG ...	Drug	RxNorm	Branded Drug Comp	S	1049195

```
CodelistGenerator:: getATCCodes(
  cdm = cdm,
  level = c("ATC 1st"),
  Name = c("Dermatologicals")
)
```

## ***2. Funciones CodelistGenerator***



Resumir el uso de códigos en nuestra base de datos:

```
chlamydia_code_use <- CodelistGenerator::summariseCodeUse(  
  chlamydia_cs,  
  byYear = TRUE,  
  bySex = TRUE,  
  ageGroup = list(c(0,17),c(18,150)),  
  cdm= cdm_sidiap ) %>%  
omopgenerics::suppress(minCellCount = 5)
```

```
CodelistGenerator::tableCodeUse(chlamydia_code_use)
```

## 2. Funciones CodelistGenerator

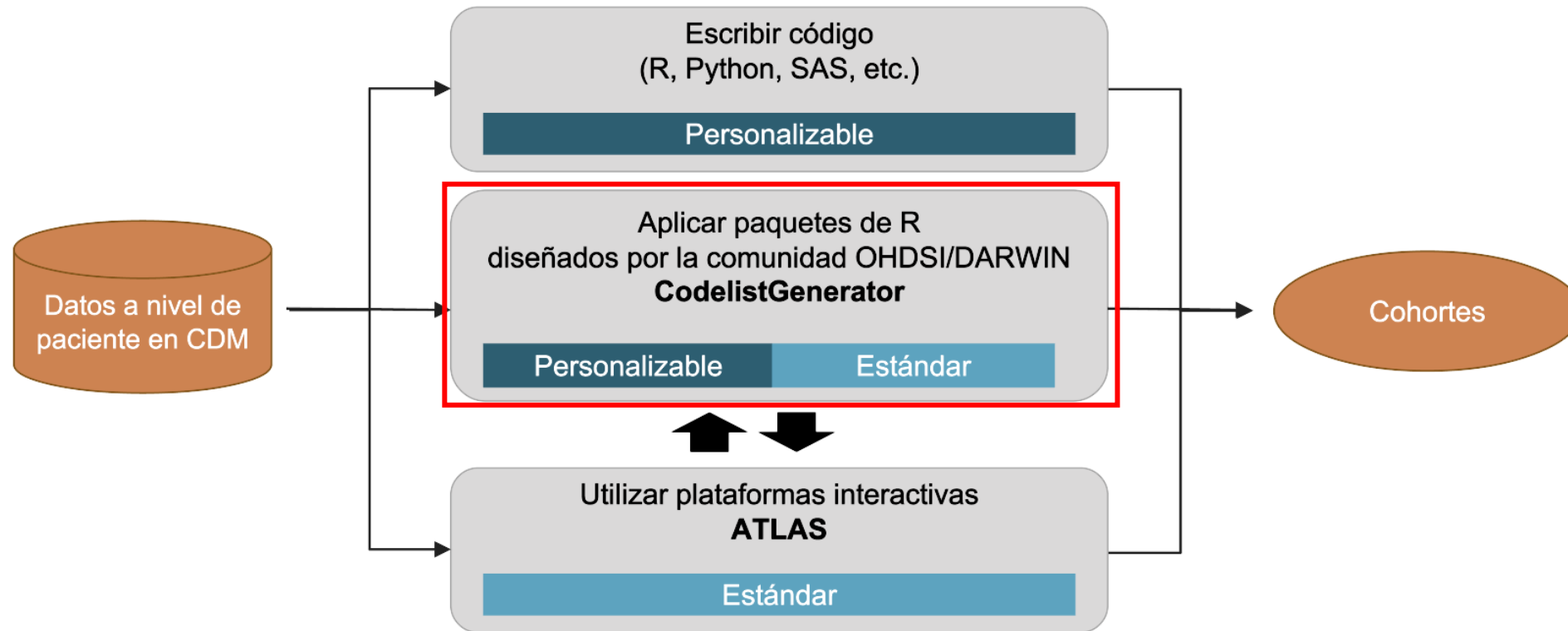


Resumir el uso de códigos en nuestra base de datos:

**CodelistGenerator::tableCodeUse(chlamydia\_code\_use)**

Codelist name	Year	Sex	Age group	Standard concept name	Standard concept ID	Source concept name	Source concept id	Domain id	Record count	Person count
chlamydia	overall	overall	overall	overall	-	overall	overall	overall	2,016	1,694
				Trachoma	380640	Unspecified trachoma	3539433	condition	6	6
				Chlamydial infection	438066	Chlamydial infection, unspecified	35205642	condition	231	196
				Ornithosis	438967	(AIDS with viral/chlamydial dis NOS) or (ornithosis) or (trachoma)	40319627	condition	8	8
				Chlamydial conjunctivitis	372550	Chlamydial conjunctivitis	372550	condition	7	5
				Genitourinary chlamydia infection	79732	Chlamydial infection of pelviperitoneum and other genitourinary organs	4089885	condition	189	164
				Chlamydial vulvovaginitis	4127408	Chlamydial vulvovaginitis	45532846	condition	7	6
				Chlamydia trachomatis infection	437221	Sexually transmitted chlamydial infection of other sites	35205604	condition	736	661
				Lymphogranuloma venereum	194287	Chlamydial lymphogranuloma (venereum)	35205600	condition	18	16
				Chlamydial pharyngitis	23137	Chlamydial pharyngitis	40376619	condition	19	19
				Pneumonia caused by Chlamydiae	254066	Pneumonia due to chlamydia	44833609	condition	11	11
				Chlamydial infection of anus and rectum	194288	Chlamydial infection of anus and rectum	35205602	condition	37	33
				Chlamydial infection of lower genitourinary tract	72245	Chlamydial infection of lower genitourinary tract	1567316	condition	734	681
				Inflammation of pelvis caused by Chlamydia trachomatis	37312673	Female chlamydial pelvic inflammatory disease	4060952	condition	12	10
				Trachoma, active stage	375736	Active stage of trachoma	35205639	condition	<5	<5

## ***2. Cohortes a partir de codelists***





## ***2. Cohortes a partir de codelists***

**CDMConnector** nos permite crear cohortes a partir de listas de códigos.



```
cdm <- cdm %>%  
  CDMConnector::generateConceptCohortSet(  
    name = "cohort",  
    conceptSet = list(  
      "chlamydia" = list1_concepts_id,  
      "gonorrhea" = list2_concepts_id))  
    limit = "first",  
    requiredObservation = c(365, 0),  
    end = "observation_period_end_date",  
    overwrite = TRUE  
  )
```

*Gracias*

Agustina Giuliodori  
[agiuliodori@idiapjgol.org](mailto:agiuliodori@idiapjgol.org)  
Real-World Epidemiology (RWEpi) Group