



April Olympians: Vocabulary techniques for ETL

Alexander Davydov, Tanya Skugarevskaya
Odysseus Data Services

Anna Ostropolets, Dmitry Dymshyts
Johnson & Johnson Innovative Medicine



Overview of topics



Dmytro Dymshyts



Tanya Skugarevskaya



Anna Ostropolets



Alexander Davydov

How to map your
source concepts

How to store your
vocabularies/mappings

How to add your
mappings/vocabularies
to Athena

What's the future

Mapping principles
Mapping tools

Source_to_concept_map
Concept,
concept_relationship

Community contribution
Mapping knowledge base



What is a concept mapping?

Same term can be coded differently

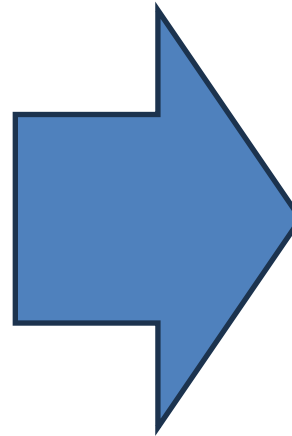
Need for the universal representation

Standard concept representing all its non-standard counterparts

Source entities:

- F20.0|Paranoid Schizophrenia (ICD10CM)
- 295.30|Paranoid type schizophrenia, unspecified (ICD9CM)
- 10033872|Paranoid schizophrenia (MedDRA)
- Schizophrenia (Medical notes)

Maps to relationship



Target concept

64905009|Paranoid schizophrenia (SNOMED)



Mapping principles

- Target standard concepts
 - Standard are the ones we're agreed to use in CDM and studies
- Domain and vocabulary consistency
 - lab data maps to measurement domain
 - Calcium from labs maps to Calcium measurement, Calcium from drugs maps to Calcium Gluconate Injection
 - cancer stage data maps to Measurement domain, Cancer Modifier vocabulary
- Usefulness
 - don't spend time mapping ambiguous or senseless in current setting source concepts
 - for example, amongst lab tests names, the term 'blood', or concepts such as 'documents reviewed', 'patient refused', etc.
- if equivalent mapping is impossible
 - map uphill (to less granular concept)
 - if possible, map to several concepts to fully represent the source concept information, for example:



○



Identify source concepts

Scenario 1. Vocabulary that exists in Athena

You know which vocabulary is used or the source tells you

diag_date	diag_time	diagnosis_cd	diagnosis_cd_type
2012-09-24	04:00:00	73020	ICD9
2012-09-13	04:00:00	K5900	ICD10
2012-09-22	04:00:00	29562	ICD9
2012-09-08	07:51:00	V426	ICD9
2012-09-13	15:11:21	2810	ICD9
2012-09-13	04:00:00	Z00129	ICD10
2012-09-04	18:29:13	2722	ICD9
2012-09-14	10:09:35	78605	ICD9
2012-09-25	04:00:00	4139	ICD9
2021-11-10	21:28:05	Z789	ICD10
2021-11-29	15:45:00	R7301	ICD10
2021-11-10	12:22:12	Z1231	ICD10
2021-11-08	05:00:00	Z1231	ICD10

diagnosis_cd -> concept_code: 73020 ->730.20

Diagnosis_cd_type -> vocabulary_id: ICD9 -> ICD9CM



Source Concept: Unspecified osteomyelitis, site unspecified (ICD9CM)



Maps to

Target concept: Osteomyelitis (SNOMED)

What to do:

Identify source concepts and use Maps and Maps to value relationships in OHDSI vocabulary to get to standard concepts

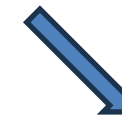


Identify source concepts

Scenario 2. Vocabulary absent in Athena

ABC standard_disease_code	ABC standard_disease_name	ABC icd10_level4_code	ABC icd10_level4_name
8850569	Superficial skin infection	L089	Local infection of skin and subcutaneous tissue, unspecified
5351003	chronic gastritis	K295	Chronic gastritis, unspecified
5580037	chronic enteritis	K529	Noninfective gastroenteritis and colitis, unspecified
0091009	acute gastroenteritis	A099	Gastroenteritis and colitis of unspecified origin
3814010	otitis media with effusion	H659	Nonsuppurative otitis media, unspecified
8847603	seasonal allergic conjunctivitis	H101	Acute atopic conjunctivitis
8837491	vaginitis	N760	Acute vaginitis
2809009	iron-deficiency anemia	D509	Iron deficiency anaemia, unspecified
3000010	anxiety neurosis	F411	Generalized anxiety disorder
1539002	colon cancer	C189	Malignant neoplasm of colon / Colon, unspecified
3898008	inner ear hearing loss	H905	Sensorineural hearing loss, unspecified
6221003	uterine cervical dysplasia	N879	Dysplasia of cervix uteri, unspecified
9100006	forehead abrasion	S002	Other superficial injuries of eyelid and periorcular area
5280017	stomatitis	K121	Other forms of stomatitis
8825128	deep vein thrombosis	I802	Phlebitis and thrombophlebitis of other deep vessels of lower

standard_disease_code



ICD10_level4_code



OHDSI concept using concept table



Standard concept via mapping

What to do:

Try to find a crosswalk to the Vocabulary that exists in Athena, then follow Scenario # 1

If not, do lexical match



Identify source concepts

Scenario 3. Text entries

ABC ptid	note_date	ABC neoplasm_histology_key	ABC neoplasm_type	ABC neoplasm_qualifier	ABC neoplasm_temporal_status	ABC direction	ABC neopla	ABC neoplasm_dx_date_type	ABC histology	ABC histology_qualifier	ABC histology_temporal_status
	2013-03-07	030e01e5f65468be1bba9d29c0680f8e	colon	actual	current	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-01	55cb44c7a0254cf98499402f6a88d7b2	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	adenocarcinoma	actual	current
	2013-03-18	c25e5b7a18d32ff522bd0fc3ed104007	ovarian	actual	history	[NULL]	[NULL]	[NULL]	carcinoma	actual	history
	2013-03-28	576ee8ec6ebc36f771a987fc2da69fa	prostatic	actual	current	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-22	050189bf6330975ea8915a07a52d44ce	mammary	actual	current	right	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-05	3820bd911da8632f24c8afe81bc9fcae	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	melanoma	actual	current
	2013-03-08	0cb8e9b83b1473f7b2c0ac3f5dc4d7c4	colon	actual	history	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-11	072817cf3ebbbe7af8fe9f2991ef2dd	bladder	actual	current	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-10	7621a0584fd82b2210b4d2cb1687ed98	prostatic	actual	history	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-30	d5c49bfd9b62c514f0c758ce4f564223	mammary	actual	history	right	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-28	0637791b317c2248106acc518e03a4f7	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]	lymphoma, non-hodgkin	actual	current
	2013-03-09	0797c0f4b6223f3eb4186da8a3bda8a2	lung	actual	current	[NULL]	[NULL]	[NULL]	squamous epithelial cells	actual	current
	2013-03-07	094d1fa02ac9cf14c77be1fdbd583d19	mammary	actual	current	right	[NULL]	[NULL]	[NULL]	[NULL]	[NULL]
	2013-03-22	0d49b5045496056d45bbba343475a209	cecum	actual	current	[NULL]	201211dd	incomplete	adenocarcinoma	actual	current

What to do:

Identify source concept, it can be tricky: in this case it's made of 3 columns:

direction + neoplasm type + histology: right mammary carcinoma

→ Mapping using Lexical match





Mapping tools

- Usagi

Google

usagi ohdsi



1st result:

[Home](#) › [Software](#) › [USAGI for vocabulary mapping](#)

USAGI for vocabulary mapping

Usagi is a software tool used to help in the process of mapping codes from a source system into terminologies, preferably standard ones, stored in the Observational Medical Outcomes Partnership (OMOP) Vocabulary (<http://www.ohdsi.org/data-standardization/vocabulary-resources/>). The word Usagi is Japanese for rabbit and was named after the first mapping exercise it was used for; mapping source codes used in a Japanese dataset into OMOP Vocabulary concepts.

For more information, check out Usagi on our wiki page: www.ohdsi.org/web/wiki



<https://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:usagi>



Nicely written documentation

Usagi

NOTICE APRIL 2021: this page is not updated anymore, the most recent documentation can be found on Github

<http://ohdsi.github.io/Usagi>



Introduction

Usagi is a software tool created by the Observational Health Data Sciences and Informatics (OHDSI) team and is used to help in the process of mapping codes from a source system into the standard terminologies stored in the Observational Medical Outcomes Partnership (OMOP) Vocabulary (<http://www.ohdsi.org/data-standardization/vocabulary-resources/>). The word Usagi is Japanese for rabbit and was named after the first mapping exercise it was used for; mapping source codes used in a Japanese dataset into OMOP Vocabulary concepts.

Mapping source codes into the OMOP Vocabulary is valuable for two main reasons:

1. When converting a raw dataset into the OMOP Common Data Model (CDM) (<http://www.ohdsi.org/data-standardization/the-common-data-model/>), translating source specific codes into standard concepts (i.e. RxNorm or SNOMED) translates the source data into a "common language" other CDMs follow.
2. Having source codes tied into the OMOP Vocabulary concepts allow a researcher to leverage the power of finding relevant source codes leveraging classification terminologies in the OMOP Vocabulary (e.g. find all antipsychotic medications or find all condition codes related to heart failure).

Scope and purpose

Source codes that needs mapping are loaded into the Usagi (if the codes are not in English additional translations columns are needed). A term similarity approach is used to connect source codes to Vocabulary concepts. However these code connections need to be manually reviewed and Usagi provides an interface to facilitate that.

Table of Contents

- ♦ Usagi
 - ♦ Introduction
 - ♦ Scope and purpose
 - ♦ Process Overview
 - ♦ Installation and support
 - ♦ Using the Application Functions
 - ♦ Importing Source Codes into Usagi
 - ♦ Reviewing Source Code to OMOP Vocabulary Concept Maps
 - ♦ Export the Usagi Map Created
 - ♦ Updating an Usagi mapping
 - ♦ Menu Options





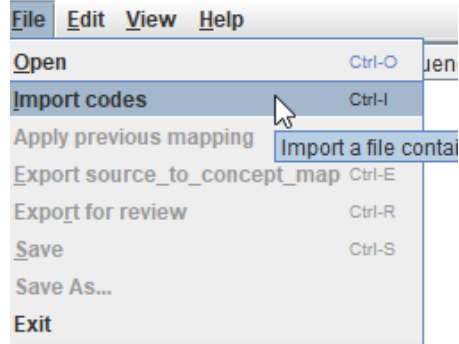
How to use usagi

- Install usagi
 - Download and unzip the OHDSI vocabulary
 - Built an index
 - Import source codes and map them
-



Import codes

Usagi v1.4.3



Import codes from onc_cond_top20.xlsx

source_code	source_vocabulary_id	frequency
mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	3822339
prostatic neoplasm	JNJ_OPTUM_EHR_ONCNHC	2485966
adenocarcinoma	JNJ_OPTUM_EHR_ONCNHC	1710429
carcinoma	JNJ_OPTUM_EHR_ONCNHC	1684122
lung neoplasm	JNJ_OPTUM_EHR_ONCNHC	1666235
colon neoplasm	JNJ_OPTUM_EHR_ONCNHC	1488257
squamous cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	1476213
skin neoplasm	JNJ_OPTUM_EHR_ONCNHC	1450629
ductal carcinoma	JNJ_OPTUM_EHR_ONCNHC	1359209
basal cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	1309949
melanoma	JNJ_OPTUM_EHR_ONCNHC	838303
in situ ductal carcinoma	JNJ_OPTUM_EHR_ONCNHC	784648
bladder neoplasm	JNJ_OPTUM_EHR_ONCNHC	762419
lymphoma	JNJ_OPTUM_EHR_ONCNHC	745707
left mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	732599
right mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	724802
lung non-small cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	599197
multiple myeloma	JNJ_OPTUM_EHR_ONCNHC	558437
malignant mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	538308
skin basal cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	523844

Column mapping

Source code column	source_code
Source name column	source_code
Source frequency column	frequency
Auto concept ID column	
Additional info column	

Filters

- Filter by user selected concepts / ATC code
- Filter standard concepts
- Include source terms
- Filter by concept class: [dropdown]
- Filter by vocabulary: [dropdown]
- Filter by domain: [dropdown]

Cancel Import

le breast 135489 Primary malignant neoplas... Condition Disorder SNOMED 93884005 S

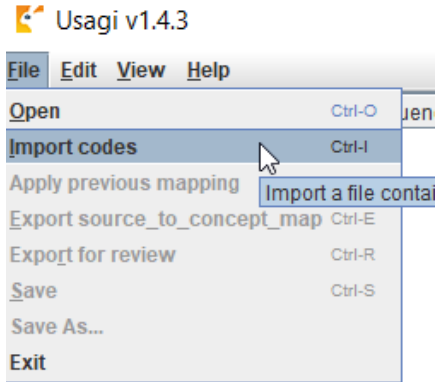


Mapping without filters of target

Source code	So...	Frequency	Match score	Con...	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept
adenocarcinoma	ad...	1710429	1.00	458...	Adenocarcinoma	Meas Value	Answer	LOINC	LA11900-0	S
ductal carcinoma	du...	1359209	1.00	422...	Infiltrating duct carcinoma	Observation	Morph Abnorma...	SNOMED	82711006	S
squamous cell carcinoma	sq...	1476213	1.00	363...	Squamous cell carcinoma	Meas Value	Answer	LOINC	LA23735-6	S
multiple myeloma	mu...	558437	1.00	360...	Multiple myeloma	Meas Value	Answer	LOINC	LA31350-4	S
melanoma	me...	838303	1.00	458...	Melanoma	Meas Value	Answer	LOINC	LA14279-6	S
lymphoma	ly...	745707	1.00	458...	Lymphoma	Meas Value	Answer	LOINC	LA15685-3	S
carcinoma	car...	1684122	1.00	458...	Carcinoma	Meas Value	Answer	LOINC	LA15448-6	S
basal cell carcinoma	ba...	1309949	1.00	402...	Basal cell carcinoma	Observation	Morph Abnorma...	SNOMED	1338007	S
lung non-small cell carcinoma	lun...	599197	0.89	402...	Non-small cell carcinoma	Observation	Morph Abnorma...	SNOMED	128632008	S
skin basal cell carcinoma	ski...	523844	0.87	411...	Basal cell carcinoma of skin	Condition	Disorder	SNOMED	254701007	S
prostatic neoplasm	pro...	2485966	0.84	432...	Metastatic malignant neoplasm	Condition	Disorder	SNOMED	128462008	S
bladder neoplasm	bla...	762419	0.77	200...	Neoplasm of bladder	Condition	Disorder	SNOMED	126885006	S
colon neoplasm	col...	1488257	0.76	403...	Neoplasm	Observation	Morph Abnorma...	SNOMED	108369006	S
malignant mammary neoplasm	ma...	538308	0.75	439...	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S
in situ ductal carcinoma	in ...	784648	0.74	371...	Non-infiltrating intraductal carcinoma	Observation	Morph Abnorma...	SNOMED	1162814007	S
skin neoplasm	ski...	1450629	0.73	444...	Neoplasm of skin	Condition	Disorder	SNOMED	126488004	S
lung neoplasm	lun...	1666235	0.70	412...	Neoplasm of lung	Condition	Disorder	SNOMED	126713003	S
mammary neoplasm	ma...	3822339	0.61	413...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
left mammary neoplasm	left...	732599	0.54	413...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
right mammary neoplasm	rig...	724802	0.54	609...	Right retinal primary malignant neoplasm	Condition	Disorder	SNOMED	35306100011...	S



How to filter the target concepts



Import codes from onc_cond_top20.xlsx

source_code	source_vocabulary_id	frequency
mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	3822339
prostatic neoplasm	JNJ_OPTUM_EHR_ONCNHC	2485966
adenocarcinoma	JNJ_OPTUM_EHR_ONCNHC	1710429
carcinoma	JNJ_OPTUM_EHR_ONCNHC	1684122
lung neoplasm	JNJ_OPTUM_EHR_ONCNHC	1666235
colon neoplasm	JNJ_OPTUM_EHR_ONCNHC	1488257
squamous cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	1476213
skin neoplasm	JNJ_OPTUM_EHR_ONCNHC	1450629
ductal carcinoma	JNJ_OPTUM_EHR_ONCNHC	1359209
basal cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	1309949
melanoma	JNJ_OPTUM_EHR_ONCNHC	838303
in situ ductal carcinoma	JNJ_OPTUM_EHR_ONCNHC	784648
bladder neoplasm	JNJ_OPTUM_EHR_ONCNHC	762419
lymphoma	JNJ_OPTUM_EHR_ONCNHC	745707
left mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	732599
right mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	724802
lung non-small cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	599197
multiple myeloma	JNJ_OPTUM_EHR_ONCNHC	558437
malignant mammary neoplasm	JNJ_OPTUM_EHR_ONCNHC	538308
skin basal cell carcinoma	JNJ_OPTUM_EHR_ONCNHC	523844

Column mapping

Source code column: source_code

Source name column: source_code

Source frequency column: frequency

Auto concept ID column: [dropdown]

Additional info column: [dropdown]

Filters

Filter by user selected concepts / ATC code

Filter standard concepts

Include source terms

Filter by concept class: [dropdown]

Filter by vocabulary: ICDO3, SNOMED

Filter by domain: Condition

Cancel Import



Mapping when target is filtered by domain and vocabulary

Source code	Sourc...	Frequency	Match score	Con...	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard co...
adenocarcinoma	aden...	1710429	1.00	404...	Malignant adenomatous neoplasm	Condition	Disorder	SNOMED	443961001	S
bladder neoplasm	bladd...	762419	0.77	200...	Neoplasm of bladder	Condition	Disorder	SNOMED	126885006	S
right mammary neoplasm	right ...	724802	0.54	609...	Right retinal primary malignant neoplasm	Condition	Disorder	SNOMED	3530610001...	S
left mammary neoplasm	left m...	732599	0.54	413...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
mammary neoplasm	mam...	3822339	0.61	413...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
ductal carcinoma	ductal...	1359209	0.78	373...	Infiltrating duct carcinoma of prostate	Condition	Disorder	SNOMED	823017009	S
multiple myeloma	multip...	558437	1.00	437...	Multiple myeloma	Condition	Disorder	SNOMED	109989006	S
squamous cell carcinoma	squa...	1476213	1.00	430...	Squamous cell carcinoma	Condition	Disorder	SNOMED	402815007	S
skin basal cell carcinoma	skin b...	523844	0.87	411...	Basal cell carcinoma of skin	Condition	Disorder	SNOMED	254701007	S
in situ ductal carcinoma	in situ...	784648	0.71	400...	Ductal carcinoma in situ of breast	Condition	Disorder	SNOMED	109889007	S
prostatic neoplasm	prost...	2485966	0.84	432...	Metastatic malignant neoplasm	Condition	Disorder	SNOMED	128462008	S
lung non-small cell carcinoma	lung n...	599197	0.80	371...	Non-small cell adenocarcinoma of lung	Condition	Disorder	SNOMED	1255725002	S
lymphoma	lymph...	745707	1.00	432...	Malignant lymphoma	Condition	Disorder	SNOMED	118600007	S
carcinoma	carcin...	1684122	1.00	367...	Malignant epithelial neoplasm	Condition	Disorder	SNOMED	722688002	S
lung neoplasm	lung n...	1666235	0.70	412...	Neoplasm of lung	Condition	Disorder	SNOMED	126713003	S
melanoma	mela...	838303	0.76	411...	Malignant melanoma of iris	Condition	Disorder	SNOMED	255012009	S
skin neoplasm	skin n...	1450629	0.73	444...	Neoplasm of skin	Condition	Disorder	SNOMED	126488004	S
basal cell carcinoma	basal ...	1309949	0.88	429...	Basal cell carcinoma of upper extremity	Condition	Disorder	SNOMED	402521004	S
malignant mammary neoplasm	malig...	538308	0.75	439...	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S
colon neoplasm	colon	1488257	0.72	413	Neoplasm of colon	Condition	Disorder	SNOMED	126838000	S



Sort by Match score and approve mappings with high Match score

Source code	Sourc...	Frequency	Match score	Conc...	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard co...
adenocarcinoma	aden...	1710429	1.00	4048...	Malignant adenomatous neoplasm	Condition	Disorder	SNOMED	443961001	S
multiple myeloma	multip...	558437	1.00	437233	Multiple myeloma	Condition	Disorder	SNOMED	109989006	S
squamous cell carcinoma	squa...	1476213	1.00	4300...	Squamous cell carcinoma	Condition	Disorder	SNOMED	402815007	S
lymphoma	lymph...	745707	1.00	432571	Malignant lymphoma	Condition	Disorder	SNOMED	118600007	S
carcinoma	carcin...	1684122	1.00	3671...	Malignant epithelial neoplasm	Condition	Disorder	SNOMED	722688002	S
basal cell carcinoma	basal ...	1309949	0.88	4297...	Basal cell carcinoma of upper extremity	Condition	Disorder	SNOMED	402521004	S
skin basal cell carcinoma	skin b...	523844	0.87	4112...	Basal cell carcinoma of skin	Condition	Disorder	SNOMED	254701007	S
prostatic neoplasm	prost...	2485966	0.84	432851	Metastatic malignant neoplasm	Condition	Disorder	SNOMED	128462008	S
lung non-small cell carcinoma	lung n...	599197	0.80	3716...	Non-small cell adenocarcinoma of lung	Condition	Disorder	SNOMED	1255725002	S
ductal carcinoma	ductal...	1359209	0.78	3731...	Infiltrating duct carcinoma of prostate	Condition	Disorder	SNOMED	823017009	S
bladder neoplasm	bladd...	762419	0.77	200680	Neoplasm of bladder	Condition	Disorder	SNOMED	126885006	S
melanoma	mela...	838303	0.76	4111...	Malignant melanoma of iris	Condition	Disorder	SNOMED	255012009	S
malignant mammary neoplasm	malig...	538308	0.75	439392	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S
skin neoplasm	skin n...	1450629	0.73	444209	Neoplasm of skin	Condition	Disorder	SNOMED	126488004	S
colon neoplasm	colon ...	1488257	0.72	4130...	Neoplasm of colon	Condition	Disorder	SNOMED	126838000	S
in situ ductal carcinoma	in situ...	784648	0.71	4001...	Ductal carcinoma in situ of breast	Condition	Disorder	SNOMED	109889007	S
lung neoplasm	lung n...	1666235	0.70	4128...	Neoplasm of lung	Condition	Disorder	SNOMED	126713003	S
mammary neoplasm	mam...	3822339	0.61	4130...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
left mammary neoplasm	left m...	732599	0.54	4130...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S
right mammary neoplasm	riht	724802	0.54	609064	Right retinal primary malignant neoplasm	Condition	Disorder	SNOMED	3530610001	S

U

Replace concept Add concept

Flag EQUAL Approve

ymshyts Vocabulary version: v5.0 29-FEB-24



Sort by Frequency

Let's map the most frequent concept

Status	Source code	Sourc...	Frequency	Match score	Conc...	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard co...	Parents	Children	Assigned To	Equivalence	Comment	Status Provenance
Unchecked	mammary neoplasm	mam...	3822339	0.61	4130...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S	1	1				
Unchecked	prostatic neoplasm	prost...	2485966	0.84	432851	Metastatic malignant neoplasm	Condition	Disorder	SNOMED	128462008	S	7	40				
Approved	adenocarcinoma	aden...	1710429	1.00	4048...	Malignant adenomatous neoplasm	Condition	Disorder	SNOMED	443961001	S	2	888		EQUAL		dymshyts (2024-0...
Approved	carcinoma	carcin...	1684122	1.00	3671...	Malignant epithelial neoplasm	Condition	Disorder	SNOMED	722688002	S	2	2663		EQUAL		dymshyts (2024-0...
Unchecked	lung neoplasm	lung n...	1666235	0.70	4128...	Neoplasm of lung	Condition	Disorder	SNOMED	126713003	S	5	19				
Unchecked	colon neoplasm	colon ...	1488257	0.72	4130...	Neoplasm of colon	Condition	Disorder	SNOMED	126838000	S	5	19				
Approved	squamous cell carcinoma	squa...	1476213	1.00	4300...	Squamous cell carcinoma	Condition	Disorder	SNOMED	402815007	S	4	716		EQUAL		dymshyts (2024-0...
Unchecked	skin neoplasm	skin n...	1450629	0.73	444209	Neoplasm of skin	Condition	Disorder	SNOMED	126488004	S	6	18				
Unchecked	ductal carcinoma	ductal...	1359209	0.78	3731...	Infiltrating duct carcinoma of prostate	Condition	Disorder	SNOMED	823017009	S	2	2				
Unchecked	basal cell carcinoma	basal ...	1309949	0.88	4297...	Basal cell carcinoma of upper extremity	Condition	Disorder	SNOMED	402521004	S	2	10				
Unchecked	melanoma	mela...	838303	0.76	4111...	Malignant melanoma of iris	Condition	Disorder	SNOMED	255012009	S	3	0				
Unchecked	in situ ductal carcinoma	in situ...	784648	0.71	4001...	Ductal carcinoma in situ of breast	Condition	Disorder	SNOMED	109889007	S	3	13				
Unchecked	bladder neoplasm	bladd...	762419	0.77	200680	Neoplasm of bladder	Condition	Disorder	SNOMED	126885006	S	7	18				
Approved	lymphoma	lymph...	745707	1.00	432571	Malignant lymphoma	Condition	Disorder	SNOMED	118600007	S	7	384		EQUAL		dymshyts (2024-0...
Unchecked	left mammary neoplasm	left m...	732599	0.54	4130...	Neoplasm of paramammary lymph nodes	Condition	Disorder	SNOMED	127258002	S	1	1				
Unchecked	right mammary neoplasm	right ...	724802	0.54	609064	Right retinal primary malignant neoplasm	Condition	Disorder	SNOMED	3530610001...	S	3	1				
Unchecked	lung non-small cell carcinoma	lung n...	599197	0.80	3716...	Non-small cell adenocarcinoma of lung	Condition	Disorder	SNOMED	1255725002	S	1	0				
Approved	multiple myeloma	multipl...	558437	1.00	437233	Multiple myeloma	Condition	Disorder	SNOMED	109989006	S	8	55		EQUAL		dymshyts (2024-0...
Unchecked	malignant mammary neoplasm	malig...	538308	0.75	439392	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S	1	2172				
Unchecked	skin basal cell carcinoma	skin b...	523844	0.87	4112...	Basal cell carcinoma of skin	Condition	Disorder	SNOMED	254701007	S	1	35				

Source code	Source term	Frequency
mammary neoplasm	mammary neoplasm	3822339

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children	Mapping Type	Creation Provenance
4130844	Neoplasm of paramammary l...	Condition	Disorder	SNOMED	127258002	S	1	1	MAPS_TO	<auto> (2024-04-09)

MAPS_TO Remove concept

Search

Query:

Use source term Query:

Filters

Filter by user selected concepts / ATC code Filter by concept class:

Filter standard concepts Filter by vocabulary:

Include source terms Filter by domain:

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.61	Neoplasm of paramammary ...	4130844	Neoplasm of paramammary ...	Condition	Disorder	SNOMED	127258002	S	1	1
0.61	Papillary neoplasm	40482256	Papillary neoplasm	Observation	Morph Abnormality	SNOMED	442172002	S	1	0
0.56	Benign neoplasm of parama...	4240035	Benign neoplasm of parama...	Condition	Disorder	SNOMED	92268005	S	2	0
0.53	Urinary neoplasm screening	4064917	Screening for malignant neo...	Measurement	Procedure	SNOMED	171178007	S	2	1
0.52	Mammoplasty NEC	2006479	Other mammoplasty	Procedure	4-dig billing code	ICD9Proc	85.89	S	2	0
0.51	Fatty neoplasm	4111953	Lipomatous tumor	Condition	Disorder	SNOMED	254827004	S	1	7
0.51	Primary malignant inflammat...	37018660	Primary malignant inflammat...	Condition	Disorder	SNOMED	96291000119105	S	3	0
0.51	Primary malignant neoplasm	439392	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S	1	2172
0.48	Malignant proliferation of pri...	37157802	Primary malignant neoplastic...	Observation	Qualifier Value	SNOMED	1234914003	S	1	0
0.46	Functioning pituitary neoplasm	40488429	Functioning pituitary neoplasm	Condition	Disorder	SNOMED	448148000	S	1	5
0.45	Micropapillary neoplasm patt...	40482825	Micropapillary tumor pattern	Condition	Clinical Finding	SNOMED	443715007	S	1	0
0.45	Primary malignant neoplasm...	4155171	Primary malignant neoplasm...	Condition	Disorder	SNOMED	371996000	S	4	161
0.45	Primary malignant neoplasm...	4311500	Primary malignant neoplasm...	Condition	Disorder	SNOMED	93990005	S	3	3
0.45	Primary malignant neoplasm...	4247338	Primary malignant neoplasm...	Condition	Disorder	SNOMED	93926005	S	3	5
0.44	Primary malignant neoplasm...	36715801	Primary malignant neoplasm...	Condition	Disorder	SNOMED	721567004	S	4	12
0.44	Primary malignant neoplasm...	134290	Primary malignant neoplasm...	Condition	Disorder	SNOMED	372002009	S	3	301
0.44	Primary malignant neoplasm...	4162253	Primary malignant neoplasm...	Condition	Disorder	SNOMED	372137005	S	3	605
0.44	Primary malignant neoplasm...	30346	Primary malignant neoplasm...	Condition	Disorder	SNOMED	93922007	S	2	14
0.44	Primary malignant neoplasm...	140950	Primary malignant neoplasm...	Condition	Disorder	SNOMED	371990006	S	4	152

MAPS_TO Replace concept Add concept

Comment: Flag EQUAL Approve



Chose concept from the list of suggestions for 'mammary neoplasm'

Query

Use source term
 Query:

Filters

Filter by user selected concepts / ATC code
 Filter standard concepts
 Include source terms

Filter by concept class:
 Filter by vocabulary:
 Filter by domain:

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.61	Neoplasm of paramammary lymph nodes	4130844	Neoplasm of paramammary...	Condition	Disorder	SNOMED	127258002	S	1	1
0.56	Benign neoplasm of paramammary lymph nodes	4240035	Benign neoplasm of param...	Condition	Disorder	SNOMED	92268005	S	2	0
0.51	Fatty neoplasm	4111953	Lipomatous tumor	Condition	Disorder	SNOMED	254827004	S	1	7
0.51	Primary malignant inflammatory neoplasm of femal...	37018660	Primary malignant inflamma...	Condition	Disorder	SNOMED	96291000119105	S	3	0
0.51	Primary malignant neoplasm	439392	Primary malignant neoplasm	Condition	Disorder	SNOMED	372087000	S	1	2172
0.46	Functioning pituitary neoplasm	40488429	Functioning pituitary neopla...	Condition	Disorder	SNOMED	448148000	S	1	5
0.45	Micropapillary neoplasm pattern	40482825	Micropapillary tumor pattern	Condition	Clinical Finding	SNOMED	443715007	S	1	0
0.45	Primary malignant neoplasm of lip	4155171	Primary malignant neoplas...	Condition	Disorder	SNOMED	371996000	S	4	161
0.45	Primary malignant neoplasm of rib	4311500	Primary malignant neoplas...	Condition	Disorder	SNOMED	93990005	S	3	3
0.45	Primary malignant neoplasm of nose	4247338	Primary malignant neoplas...	Condition	Disorder	SNOMED	93926005	S	3	5
0.44	Primary malignant neoplasm of placenta	36715801	Primary malignant neoplas...	Condition	Disorder	SNOMED	721567004	S	4	12
0.44	Primary malignant neoplasm of palate	134290	Primary malignant neoplas...	Condition	Disorder	SNOMED	372002009	S	3	301
0.44	Primary malignant neoplasm of heart	316644	Primary malignant neoplas...	Condition	Disorder	SNOMED	93825008	S	4	92
0.44	Primary malignant neoplasm of breast	4162253	Primary malignant neoplas...	Condition	Disorder	SNOMED	372137005	S	3	605
0.44	Primary malignant neoplasm of neck	30346	Primary malignant neoplas...	Condition	Disorder	SNOMED	93922007	S	2	14
0.44	Primary malignant neoplasm of gum	140950	Primary malignant neoplas...	Condition	Disorder	SNOMED	371990006	S	4	152
0.43	Primary malignant neoplasm of eye	374874	Primary malignant neoplas...	Condition	Disorder	SNOMED	371986009	S	3	257
0.43	Primary malignant neoplasm of ovary	200051	Primary malignant neoplas...	Condition	Disorder	SNOMED	93934004	S	8	197
0.43	Primary malignant neoplasm of patella	4247849	Primary malignant neoplas...	Condition	Disorder	SNOMED	93950007	S	2	0

MAPS_TO

Replace concept Add concept

Comment:

Flag EQUAL Approve

Approved / total: 0 / 20 0.0% of total frequency

Author: dymshyts Vocabulary version: v5.0 29-FEB-24



Search by keywords

MAPS_TO Remove concept

Search

Query

Use source term

Query: Filters

Filter by user selected concepts / ATC code Filter by concept class:

Filter standard concepts Filter by vocabulary:

Include source terms Filter by domain:

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.76	Breast neoplasm screening	4147961	Screening for malignant neo...	Measurement	Procedure	SNOMED	268547008	S	1	2
0.71	Neoplasm of breast	81251	Neoplasm of breast	Condition	Disorder	SNOMED	126926005	S	4	32
0.68	Breast neoplasm screening status	4086756	Breast neoplasm screening ...	Observation	Clinical Finding	SNOMED	243878006	S	1	9
0.67	Malignant neoplasm of breast	4112853	Malignant tumor of breast	Condition	Disorder	SNOMED	254837009	S	9	23
0.65	Malignant neoplasm of male breast	4157447	Malignant neoplasm of male...	Condition	Disorder	SNOMED	372095001	S	2	6
0.61	Malignant neoplasm of breast in remission	37172171	Malignant neoplasm of brea...	Condition	Disorder	SNOMED	45061000087102	S	2	2
0.61	Malignant neoplasm of female breast	4157332	Malignant neoplasm of fema...	Condition	Disorder	SNOMED	372064008	S	7	14
0.60	Benign neoplasm of breast	72576	Benign tumor of breast	Condition	Disorder	SNOMED	269485000	S	4	117
0.59	Primary malignant neoplasm of breast	4162253	Primary malignant neoplas...	Condition	Disorder	SNOMED	372137005	S	3	605
0.58	Primary malignant neoplasm of male breast	135489	Primary malignant neoplas...	Condition	Disorder	SNOMED	93884005	S	3	6
0.57	Benign neoplasm of male breast	4242803	Benign neoplasm of male br...	Condition	Disorder	SNOMED	92206006	S	2	3
0.57	Malignant neoplasm of breast in partial remission	37168468	Malignant neoplasm of brea...	Condition	Disorder	SNOMED	1281812003	S	2	0
0.57	Benign neoplasm of left breast	37208024	Benign neoplasm of left bre...	Condition	Disorder	SNOMED	350821000119101	S	2	6
0.56	Benign neoplasm of right breast	37208028	Benign neoplasm of right br...	Condition	Disorder	SNOMED	351381000119101	S	2	6
0.55	Malignant neoplasm of ectopic site of male breast	4091471	Malignant neoplasm of ecto...	Condition	Disorder	SNOMED	188168005	S	2	1
0.55	Malignant neoplasm of breast in full remission	37175247	Malignant neoplasm of brea...	Condition	Disorder	SNOMED	44721000087109	S	2	0
0.55	No history of malignant neoplasm of breast	45763684	No history of malignant tum...	Observation	Context-dependent	SNOMED	700146008	S	2	0
0.55	Primary malignant neoplasm of female breast	137809	Primary malignant neoplas...	Condition	Disorder	SNOMED	93796005	S	2	13
0.55	Primary malignant neoplasm of left male breast	37208047	Primary malignant neoplas...	Condition	Disorder	SNOMED	354491000119109	S	2	0

MAPS_TO Replace concept Add concept

Comment:

Approved / total: 0 / 20 0.0% of total frequency Author: dymshyts Vocabulary version: v5.0 29-FEB-24



Choose concept and approve it

Search
Query

Use source term
 Query:

Filters

- Filter by user selected concepts / ATC code
- Filter standard concepts
- Include source terms
- Filter by concept class:
- Filter by vocabulary:
- Filter by domain:

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.71	Neoplasm of breast	81251	Neoplasm of breast	Condition	Disorder	SNOMED	126926005	S	4	32
0.67	Malignant neoplasm of breast	4112853	Malignant tumor of breast	Condition	Disorder	SNOMED	254837009	S	9	23
0.65	Malignant neoplasm of male breast	4157447	Malignant neoplasm of mal...	Condition	Disorder	SNOMED	372095001	S	2	6
0.61	Malignant neoplasm of breast in remission	37172171	Malignant neoplasm of bre...	Condition	Disorder	SNOMED	45061000087102	S	2	2
0.61	Malignant neoplasm of female breast	4157332	Malignant neoplasm of fem...	Condition	Disorder	SNOMED	372064008	S	7	14
0.60	Benign neoplasm of breast	72576	Benign tumor of breast	Condition	Disorder	SNOMED	269485000	S	4	117
0.59	Primary malignant neoplasm of breast	4162253	Primary malignant neoplas...	Condition	Disorder	SNOMED	372137005	S	3	605
0.58	Primary malignant neoplasm of male breast	135489	Primary malignant neoplas...	Condition	Disorder	SNOMED	93884005	S	3	6
0.57	Benign neoplasm of male breast	4242803	Benign neoplasm of male ...	Condition	Disorder	SNOMED	92206006	S	2	3
0.57	Malignant neoplasm of breast in partial remission	37168468	Malignant neoplasm of bre...	Condition	Disorder	SNOMED	1281812003	S	2	0
0.57	Benign neoplasm of left breast	37208024	Benign neoplasm of left br...	Condition	Disorder	SNOMED	350821000119101	S	2	6
0.56	Benign neoplasm of right breast	37208028	Benign neoplasm of right b...	Condition	Disorder	SNOMED	351381000119101	S	2	6
0.55	Malignant neoplasm of ectopic site of male breast	4091471	Malignant neoplasm of ect...	Condition	Disorder	SNOMED	188168005	S	2	1
0.55	Malignant neoplasm of breast in full remission	37175247	Malignant neoplasm of bre...	Condition	Disorder	SNOMED	44721000087109	S	2	0
0.55	Primary malignant neoplasm of female breast	137809	Primary malignant neoplas...	Condition	Disorder	SNOMED	93796005	S	2	13
0.55	Primary malignant neoplasm of left male breast	37208047	Primary malignant neoplas...	Condition	Disorder	SNOMED	354491000119109	S	2	0
0.54	Malignant neoplasm of central part of female breast	4092511	Malignant neoplasm of cen...	Condition	Disorder	SNOMED	188151006	S	2	4
0.54	Benign neoplasm of female breast	4239883	Benign neoplasm of femal...	Condition	Disorder	SNOMED	92102001	S	2	11

MAPS_TO

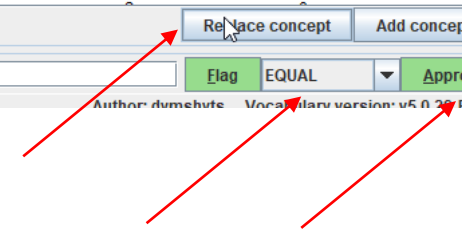
comment:

Approved / total: 5 / 20 24.4% of total frequency

Author: dumebute Vocabulary version: v5.0.20 FER

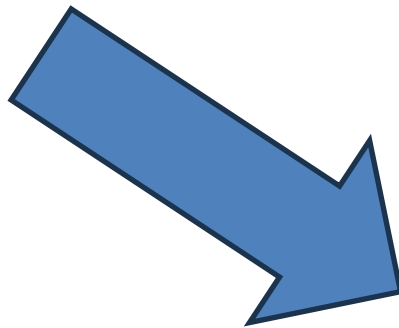
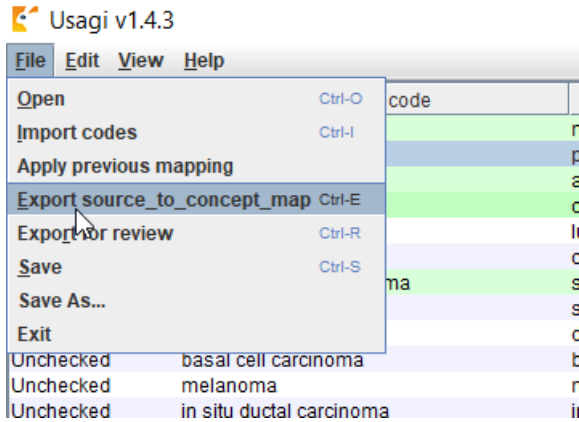
Replace concept Add concept

Flag EQUAL Approve





Source_to_concept_map export



source_code	source_code	source_vocabulary	source_code_description	target_concept_id	target_vocabulary	valid_start	valid_end_date	invalid_reason
carcinoma	0 jnjonco		carcinoma	36716620	SNOMED	1/1/1970	12/31/2099	
mammary neoplasm	0 jnjonco		mammary neoplasm	4130844	SNOMED	1/1/1970	12/31/2099	
squamous cell carcinoma	0 jnjonco		squamous cell carcinoma	4300118	SNOMED	1/1/1970	12/31/2099	
adenocarcinoma	0 jnjonco		adenocarcinoma	40484156	SNOMED	1/1/1970	12/31/2099	
multiple myeloma	0 jnjonco		multiple myeloma	437233	SNOMED	1/1/1970	12/31/2099	
lymphoma	0 jnjonco		lymphoma	432571	SNOMED	1/1/1970	12/31/2099	



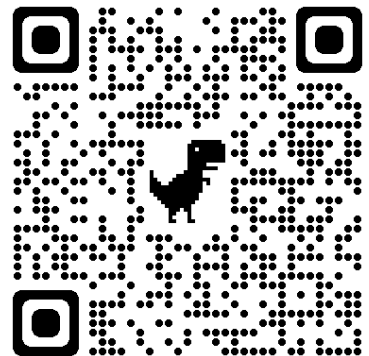
I have my mappings. What next?

Now, we have mappings of source codes to standard concepts that we need to put in our CDM

There are two pathways of doing that during ETL:

A) source_to_concept_map table

B) concept and concept_relationship tables





Scenario 1: source_to_concept_map

Conveniently store your source concepts and mapping in one table

Easy to add/remove mappings

Source_code	Source_concept_id	Source_vocabulary_id	Source_code_description	Target_concept_id	Target_vocabulary_id	Valid_start_date	Valid_end_date	Invalid_reason
Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Varchar(50)	Integer	Varchar(20)	Varchar(255)	Integer	Varchar(20)	Date	Date	Varchar(1)
mammary neoplasm	2100000000	otum_onc_condition		81251	SNOMED	31/01/2002	12/31/2099	



Only exists as _source_concept_id/source_value

Codes and mappings are not visible/cannot be used in standard tools (such as Atlas)

Source codes without mapping exist only as source codes



Scenario 2: concept

Store your source concepts

Concept_id	Concept_name	Domain_id	Vocabulary_id	Concept_class_id	Standard_concept	Concept_code	Valid_start_date	Valid_end_date	Invalid_reason
Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No
2100000000	mammary neoplasm	Condition	otum_onc_condition	Disorder		mammary neoplasm	01/01/1970	12/31/2099	

New 2 billion concepts are visible in Atlas/by other tools and packages, even those without mapping
Need to assign domain, vocabulary and other attributes

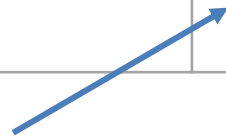
Can create standard concepts – we recommend talking to us first



Scenario 2: concept_relationship

Store your mappings

Concept_id_1	Concept_id_2	Relationship_id	Valid_start_date	Valid_end_date	Invalid_reason
Yes	Yes	Yes	Yes	Yes	No
2100000000	81251	Maps to	01/01/1970	12/31/2099	
81251	2100000000	Mapped from	01/01/1970	12/31/2099	



Mappings are bi-directional, are visible in Atlas, can be shared
Can store relationships other than Maps to (such as Maps to value)



Comparison of two approaches

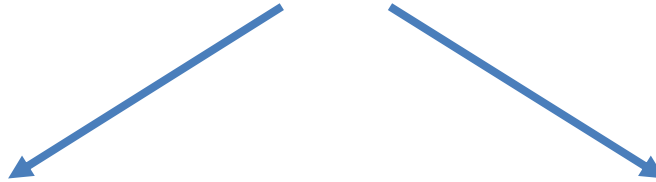
Parameter	STCM	Concept/Concept_relationship
General description	Creation of a mapping from source_code to standard_concept_id	Creation of a custom concept (with concept_id>2bil) and mapping it to a standard_concept_id
Loading	Easy: table is designed for insertion of records	Harder: Requires alteration of vocabulary tables from Athena
Unmapped codes	Live only as source_codes	Live as custom concepts without mapping, i.e. as event_source_concept_id and can be queried by Atlas
Maps_to_value relationships	Does not support	Supports
ETL logic	Requires additional querying of the STCM	Allows for a usual lookup of a concept in the Concept table
Research use	Little: data is not visible in Atlas or the Concept table	Some: Custom concepts can be used in network studies through Atlas
Conclusion	Easy to implement, but not sharable	More complex, but sharable and usable in network research



I want to share my work with the community!
How do I do that?

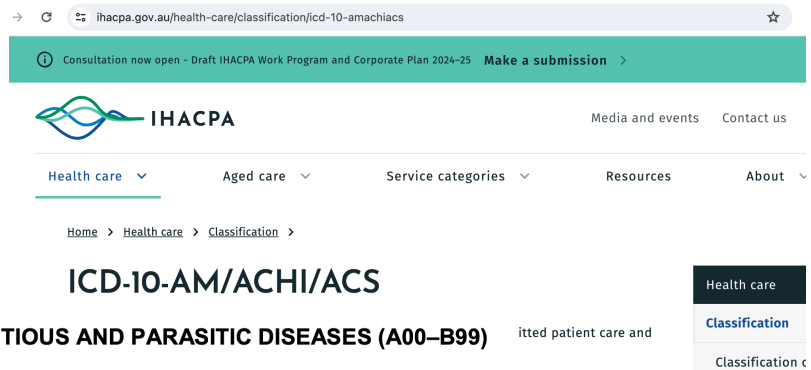


I have



Vocabularies and/or mappings

- Is a dictionary - has codes and structure
- Has external reference



Chapter 1 CERTAIN INFECTIOUS AND PARASITIC DISEASES (A00–B99)

- A00-B99 Certain infectious and parasitic diseases
- A00 Cholera
- A00.0 Cholera due to *Vibrio cholerae* 01, biovar cholerae
- A00.1 Cholera due to *Vibrio cholerae* 01, biovar eltor
- A00.9 Cholera, unspecified
- A01 Typhoid and paratyphoid fevers

Free text strings

- Collection of strings, names are oftentimes poorly standardized
- Has no official vocab name or reference

source_code	frequency
mammary neoplasm	3822339
prostatic neoplasm	2485966
adenocarcinoma	1710429
carcinoma	1684122
lung neoplasm	1666235
colon neoplasm	1488257
squamous cell carcinoma	1476213



I have

Vocabularies and/or mappings

Free text strings

- Is a dictionary - has codes and structure
- Has external reference

Mapping knowledge base is on its way!

Please talk to Melanie Philofsky

ihacpa.gov.au/health-care/classification/icd-10-amachiacs

Consultation now open - Draft IHACPA Work Program and Corporate Plan 2024-25 [Make a submission](#)

IHACPA Media and events Contact us

Health care Aged care Service categories Resources About

Home > Health care > Classification

ICD-10-AM/ACHI/ACS

Health care
Classification
Classification c

Chapter 1 CERTAIN INFECTIOUS AND PARASITIC DISEASES (A00–B99)

- A00-B99** Certain infectious and parasitic diseases
- A00** Cholera
- A00.0** Cholera due to *Vibrio cholerae* 01, biovar cholerae
- A00.1** Cholera due to *Vibrio cholerae* 01, biovar eltor
- A00.9** Cholera, unspecified
- A01** Typhoid and paratyphoid fevers

Healthcare Systems

2024 OKR slides



Melanie Philofsky



I have

Vocabularies and/or mappings

Free text strings

- Collection of strings, names are oftentimes poorly standardized
- Has no official vocab name or reference

Add your source vocabulary with mappings to Athena

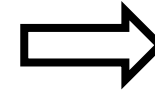
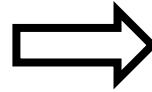
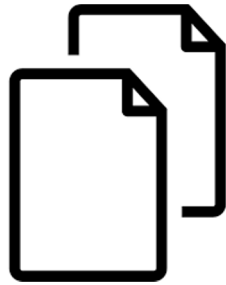
Add mappings for unmapped concepts in Athena

Fill the templates!





Community contribution: pipeline



Issue on GitHub

Template

- Content
- Meta-data
- Checklist (QA)

Submission
Review

Release





Example of templates

One template for everything related to adding a vocabulary: template_4_adding_vocabulary

Spreadsheet 1: Fill concept_manual with your concepts (very similar to concept in Tanya's part)

A	B	C	D	E	F	G	H	I
concept_name	concept_code	vocabulary_id	domain_id	concept_class_id	standard_concept	valid_start_date	valid_end_date	invalid_reason
concept name of the code you want to add max 255 characters	concept code you want to add max 50 characters	vocabulary_id of an existing OHDSI Vocabulary max 20 characters	domain of your source concept, has to be an existing OHDSI domain max 20 characters	has to be an existing OHDSI class max 20 characters	empty only	date from the source or default (1/1/1970)	12/31/99	empty only
Mammary neoplasm	12345	Optum EHR Oncology	Condition	Disorder		1/1/70		

Spreadsheet 2: Fill concept_relationship_manual with your mappings (very similar to concept_relationship in Tanya's part)

A	B	C	D	E	F	G
concept_code_1	vocabulary_id_1	relationship_id	concept_id_2	concept_code_2	vocabulary_id_2	relationship_valid_start_date
codes you want to add max 50 characters	short name (id) of your vocabulary max 20 characters	Maps to only	concept_id in the OHDSI Vocabularies you map to [USE EITHER concept_id_2 OR concept_code_2+vocabulary_id_2] max 50 characters	concept_code in the OHDSI Vocabularies you map to max 50 characters	OHDSI Vocabulary you map to max 20 characters	usually current date
12345	Optum EHR Oncology	Maps to	81251		SNOMED	1/1/70



Example of templates

One template for everything related to adding a vocabulary: `template_4_adding_vocabulary`

Spreadsheet 1: Fill `concept_manual` with your concepts (very similar to concept in Tanya's part)

Spreadsheet 2: Fill `concept_relationship_manual` with your mappings (very similar to `concept_relationship` in Tanya's part)

<code>confidence</code>	<code>predicate_id</code>	<code>mapping_source</code>	<code>mapping_justification</code>	<code>mapping_tool</code>
confidence in the new mapping 0-1	indicates type of mapping: <code>exactMatch</code> for exact match, <code>narrowMatch</code> for mapping from a broad to a granular concept, <code>broadMatch</code> for mapping from granular specific concept to a less specific broad concept	indicate where the mapping is coming from (eg., <code>ManualMapping</code> or name of the resource you are getting the mappings from)	how the mapping was arrived at: <code>ManualMappingCuration</code> , <code>LexicalMatching</code> , <code>DataDriven</code>	fill with the name of the tool if such was used