

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP

OMOP Standard Vocabulary
Version 4.4 2013-01-05
Release Notes

Christian Reich, Patrick Ryan, Don Torok, Sergei Vereshagin,
Mark Khayter and Emily Welebob

05-January-2014

License

© 2009-2014 Observational Medical Outcomes Partnership (OMOP).

Licensed under the Apache License, Version 2.0 (the "License"); you may not use this document except in compliance with the License. You may obtain a copy of the License at <http://omop.org/publiclicense>.

Unless required by applicable law or agreed to in writing, documentation and software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. Any redistributions of this work or any derivative work or modification based on this work should be accompanied by the following source attribution: "This work is based on work by the Observational Medical Outcomes Partnership (OMOP) and used under license from the RUF at <http://omop.org/publiclicense>".

Any scientific publication that is based on this work should include a reference to <http://omop.org>.

Release

This document reflects all the changes to the Vocabulary introduced since the Version 4.4 of 2013-10-15. In addition to executing standard updates (see Appendix for latest content), the following changes were made:

List of changes

1. REVISED MAPPING FROM SPECIAL ICD-9-CM TO SNOMED-CT THAT DO NOT DEFINE A CURRENT DIAGNOSIS

ICD-9 codes such as 539.09 "Other complications of gastric band procedure" are now correctly mapped to the equivalent complication etc., or to concept 0. In addition, codes like V04.6 "Need for prophylactic vaccination and inoculation against mumps alone" are now mapped with a mapping type "OBSERVATION".

2. FIXED VALID DATES FOR DRG AND MS-DRG

DRG and MS-DRG (both vocabulary_id=40) had lost in a previous vocabulary update all validity information (valid_start_date, valid_end_date, invalid_reason). This was restored and updated.

Note: DRG (concept_class='DRG') were changed to MS-DRG (concept_class='MS-DRG') on 1-Oct-2007. All DRG concepts are deprecated (invalid_reason='D') on this date. However all other changes within DRGs and MS-DRGs are upgrades (invalid_reason='U') and records in the concept_relationship table (relationship_id= 299) connect them to their successors.

3. ADDRESSING THE PROBLEM OF DEPRECATED CONCEPTS WITH NO SUCCESSOR

Some vocabularies have a high turn-over of concepts, leading to a high number of deprecated concepts. E.g., SNOMED-CT (vocabulary_id=1) has 94,627 such concepts, without providing any information about a successor concept. Any mapping to valid concepts will become obsolete when the concept is deprecated by the source vendor of these vocabulary data. However, it is our goal to have for each source_code a current active concept. Some of these concepts get deprecated because they are duplicates. We therefore introduced a method to detect and upgrade such duplicates.

vocabulary_id	relationship_id	relationship_name	Count
1	311	Concept replaced by (SNOMED)	17,472
8	309	Concept replaced by (RxNorm)	264

4. ANCESTORS WITH RECORDS IN THE CONCEPT_ANCESTOR TABLE TO SELF

In the concept_ancestor table, concepts have records pointing to themselves as descendants, making it easy to generate a list of an ancestor and all its descendants with a single query. However, in the past, only those ancestors had a record to themselves that also had other, proper descendants. Concepts with no descendants were not listed in the concept_ancestor table at all. This rule is now changed, and all concepts now also point to themselves as descendants.

5. FDB GENSEQNO NOW UP TO DATE

Drug concepts maintained by First Databank (FDB) use GenSeqNo (vocabulary_id=53) as identifiers. These had not been updated in 3 years. This is now fixed, and they are up to date (see Appendix).

6. MAPPING OF ERRONEOUS 11-DIGIT NDCs NO LONGER RE-MAPPED

11-digit NDCs coming from various sources (FDA, FDB, RxNorm) without mapping to an RxNorm Clinical Drug or Branded Drug are no longer being re-mapped. In the past, we tried to map these NDCs using a heuristic that applied finding a valid mapping using only the first 9 digits and applying that mapping if all possible 9-digit maps led to the same target_concept_id. It remains unclear whether these 11-digit NDCs were correctly mapped but the source applied a wrong source_code_description, or whether these NDCs are what the source_code_description is depicting. However, these NDCs are now mapped to concept_id=0.

7. INTRODUCING A NEW VOCABULARY FOR OPCS-4

The Office of Population Censuses and Surveys Classification (OPCS)-4 database is a procedural classification for the coding of operations, procedures and interventions mostly for inpatient procedures. These codes are used in the HES database (see below). OPCS-4 codes were added as Standard Vocabulary concepts (vocabulary_id=55).

8. INTRODUCING A NEW VOCABULARY FOR GEMSCRIPT

The Clinical Practice Research Datalink (CPRD) database is no longer coding drugs in the Multilex (vocabulary_id=22) system. Instead, it introduced a new system called Gemscript (vocabulary_id=56). All Gemscript codes are added to the vocabulary and mapped in the source_to_concept_map to RxNorm (vocabulary_id=8) through the provided Multilex mapping.

9. INTRODUCING A NEW VOCABULARY FOR HES SPECIALTY

The Hospital Episode Statistics (HES) database provides codes for its provider specialties. All codes have now records in the concept table (vocabulary_id=57) as well as records in the source_to_concept_map. The concepts that have an equivalent CMS Specialty (vocabulary_id=48) concept are not Standard Vocabulary (concept_level=0), while those that are complementing the existing CMS specialty concept set became Standard concepts (concept_level=2).

10. REVISION OF UCUM

The Unified Code for Units of Measure (UCUM, vocabulary_id=11) concepts are used to encode lab test result units, such as g/dL. UCUM concepts were revised with the following improvements:

- Errors in the UCUM notation of units (in the concept_code field) were corrected.
- Descriptions in the concept_name were standardized and corrected.
- Additional concepts were added.
- Duplicates were deprecated (invalid_reason='U') and records in the concept_relationship table introduced that would upgrade these.
- New source_to_concept_map tables were produced for units used in the MarketScan, GE and CPRD databases
- The distinction into Standard UCUM and Custom UCUM was abolished (there is no such thing as a standard UCUM). All currently valid concepts now have concept_class "UCUM"

- A new class was introduced "UCUM canonical". These are units can be used as a reference for commensurable units, i.e. units that can be transformed to the canonical (mostly through multiplication). The relationship between the canonical and the ordinary UCUM units will be added at a later time.

Appendix: Source Vocabulary Versions and Sources

The following table contains the location for all external source vocabularies used for the OMOP Standard Vocabularies

Vocabulary ID	Vocabulary Name	Source	Date used for Vocabulary
1	SNOMED-CT	UMLS distribution, http://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html	31-Jul-2013
2	ICD-9-CM	Diagnosis and Procedure Codes: Abbreviated and Full Code Titles, Version 31, http://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/codes.html	1-Oct-2013
3	ICD-9-Procedure		
4	CPT-4	UMLS distribution, http://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html	8-May-2013
5	HCPCS	UMLS distribution, http://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html	28-Aug-2013
6	LOINC	LOINC distribution, http://loinc.org/downloads/loinc	29-Jun-2013
7	NDF-RT	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
8	RxNorm	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
9	NDC	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
		FDB distribution package	20-Jun-2013
		FDA NDC distribution, http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM070838.zip	26-Nov-2013
10	GPI	Medi-Span distribution package	4-Nov-2013
11	UCUM	The Unified Code for Units of Measure, Version 1.8.2, http://aurora.regenstrief.org/~ucum/ucum.html#section-Alphabetic-Index	3-Jan-2014
12	HL7 Administrative Sex	HL7, Appendix A: Data Definition Tables, http://www.hl7.org/special/committees/vocab/v26_appendix_a.pdf	27-Nov-2007
13	CDC Race	http://www.cdc.gov/nchs/data/dvs/Race_Ethnicity_CodeSet.pdf	27-Jun-2000
14	CMS Place of Service	Place of Service Codes for Professional Claims Database, http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/downloads/Website_POS_database.pdf	1-Nov-2009
15	MedDRA	MedDRA Version 16.1, http://www.meddrasso.com/subscriber_download.asp	13-Jun-2013
16	Multum	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013

OMOP STANDARD VOCABULARY 4.4 2014-01-05

Vocabulary ID	Vocabulary Name	Source	Date used for Vocabulary
17	Read	UMLS distribution, http://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html	3-Jan-2014
18	OXMIS	Codes extracted from GPRD database, courtesy of GSK	1-Mar-2009
19	FDB Indication	FDB US distribution package	20-Jun-2013
20	FDB ETC	FDB US distribution package	20-Jun-2013
21	WHO ATC	FDB US distribution package	20-Jun-2013
22	Multilex	FDB UK distribution package	5-Jun-2013
28	VA Product	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
31	SMQ	Standardised MedDRA Queries (SMQs) http://www.meddrasso.com/secure/smq/SMQ_Spreadsheet_14_0_English_update.xlsx	23-Jan-2013
32	VA Class	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
34	ICD-10-CM	2011 ICD-10-PCS and GEMs, Code Descriptions, http://www.cms.gov/Medicare/Coding/ICD10/2012-ICD-10-CM-and-GEMs.html	21-Nov-2013
35	ICD-10-PCS	2012 ICD-10-PCS and GEMs, Code Descriptions, http://www.cms.gov/Medicare/Coding/ICD10/2012-ICD-10-PCS.html	21-Nov-2013
40	DRG	List of Diagnosis Related Groups (DRGs), FY 2005, http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MedicareFeeforSvcPartsAB/downloads//DRGDesc05.pdf	30-Jul-2012
	MS-DRG	Acute Inpatient - Files for Download , https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Acute-Inpatient-Files-for-Download.html , Files for FY 20XX Final Rule and Correction Notice, Table 5	18-Feb-2011
41	MDC	MDC Description File, http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Acute-Inpatient-Files-for-Download-Items/CMS1247844.html	1-Jun-2013
42	APC	Hospital Outpatient PPS, Addendum A and Addendum B Updates, http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalOutpatientPPS/Addendum-A-and-Addendum-B-Updates.html	28-Sep-2011
46	NLM MeSH	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
47	NUCC	http://www.nucc.org/images/stories/CSV/nucc_taxonomy_131.csv , Version 13.1	7-Jan-2013
48	CMS Specialty	http://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/MedicareProviderSupEnroll/downloads/JSMTDL-08515MedicarProviderTypetoHCPTaxonomy.pdf	22-09-2008
50	FDA SPL	RxNorm distribution, http://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html	2-Dec-2013
		FDA NDC distribution, http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM070838.zip	26-Nov-2013

OMOP STANDARD VOCABULARY 4.4 2014-01-05

Vocabulary ID	Vocabulary Name	Source	Date used for Vocabulary
53	FDB Genseqno	FDB US distribution package	4-Nov-2013
55	OPCS	http://www.isb.nhs.uk/documents/isb-0084/amd-10-2013/0084102013summ.pdf	7-Nov-2013
56	Gemscript	http://www.resip.co.uk/gemscript	12-Dec-2013
57	HES Specialty		3-Jan-2014