Lab Results Data Dictionary

Overview:

The VDW Lab Data Structure consists of two tables.

- The main table, LAB RESULTS, has one record per lab result. If a test is not resulted for whatever reason (specimen not sufficient, patient did not show) then the test
- The second table, LAB NOTES, links to LAB RESULTS by the variable Row_ID. Not all rows in the LAB RESULTS data set will have corresponding records in LAB N⟨ more than 8 bytes character storage for results, normal low or normal high or a record that has notes associated with it will have one or more corresponding observati
- Please see the example after Table 2b (near the end of this document) for illustrations of this concept.

2a. Lab Results table data dictionary

Variable name	Type and Length	Format	Label	Definition	Valid Values
MRN	Char, (length is site specific)		Medical Record Number	An arbitrary identifier unique to an individual within a site.	Any, so long as they uniquely identify individual people.
Test_Type	Char (10)			HMORN's official name for the type of test	https://appliedresearch.cancer.gov/crnportal/data-resources/vdw/datasets/everloinc-records
LOINC	Char (10)			Codes in the form of nnnnn-n (left-justified)	Do not map LOINC codes. Only complete this field if the LOINC code is available in the data. Current codes are from 3-7 characters long. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. The hyphen, as well as all the numbers, is part of the LOINC and must be included.
Stat	Char(1)			Immediacy of test	'E'=Expedite 'S'=Stat 'R'=Routine 'U'=Unknown or missing
Pt_Loc	Char(1)			Location of the patient when the lab specimen was obtained	'I'=Inpatient 'O'=Outpatient 'E'='Emergency Department 'H'=Home 'U'=Unknown or missing
Result_Loc	Char(1)			Result location is used to identify whether the test was completed in a laboratory (e.g., hospital lab, outpatient clinic lab, emergency department lab) or if the test was completed using a point of care device at a location other than a certified laboratory (e.g., newborn nursery, finger stick in provider office, home).	'L'=Lab 'P'=Point of Care The variable will have the default value of 'L' unless the result is Point of Care. There are no missing values.
Specimen_ID	Char (length is site specific)			Specimen ID will be used to connect multiple records from the same	A possible way to create Specimen_ID is by concatenating lab accession number and date. At most sites lab accession number is reused, thus concatenating the date will make a unique specimen ID for that test_type. Since date is PHI consider applying a hash. md5(compress(AccessionNumber||SampleCollectionDt)) as Specimen_ID format=$hex32.

				blood sample, e.g. NTSEG_C and NTBAND_C to create a close estimate to the absolute neutrophil count
Specimen_Source	Char(6)		Sample or method that results in a specimen to be tested	BLOOD SERUM PLASMA SR_PLS = serum/plasma Some tests can be performed on either, and both values are listed in the specimen source field. PPP = Platelet Poor Plasma CSF= cerebral spinal fluid URINE STOOL NSWAB = nasal swab Any specimen source in your data listed as "nose" should be mapped to "nasal swab". NWASH = nasal wash NPH = nasopharyngeal swab NPWASH = nasopharyngeal wash THRT = throat, oropharyngeal swab SPUTUM BAL = bronchoalveolar lavage (BAL) specimen BALBX = BAL biopsy OTHER ' ' = missing value (assumptions can probably be made depending on test_type) NS = not specified
LOCAL_CD	Char, (length is site specific)		Non LOINC code related to lab test result	Unique to each HMO
BATTERY_CD	Char, (length is site specific)		Battery Code	Unique to each HMO
PX	Char, (length is site specific)		Related procedure code (most likely CPT or Local)	CPT-4 ##### HCPCS A#### ICD-9 ##.##or Local code
Codetype	Char(1)		Specifies what type of code PX is	C=CPT, I=ICD9,

Variable	Type			Description	Notes
					H=HCPCS,
					L=local home-grown,
					O=other
Order_ID	Char, (length is site specific)			This variable links orders to results.	Unique to each site.
Order_dt	Num(4)			Date test was ordered	SAS Date
Lab_dt	Num(4)			Date specimen was collected	SAS Date
Lab_tm	Num(4)			Time specimen was collected	SAS Time
Result_dt	Num(4)			Date of test result	SAS Date
Result_tm	Num(4)			Time of test result	SAS Time
Result_C	Char(8)			Test result as short character.	To accept digits, decimal, sign or text (e.g. POSITIVE, NEGATIVE, DETECTED) result values.The symbols >, <, >=, <= should be removed.If the input data contains a comment in the same field as the result value then Result_C should contain only the result value and the comment should be stored in the LAB NOTES table. The lab workgroup is working to standardize the acceptable values for qualitative test_types. Below is a list of test_types and acceptable values to date. <table><tr><th>Qualitative Test_Type</th><th>Acceptable Values</th></tr><tr><td>PG_B_QL, PG_B_QL_U, PG_QL, PG_QL_U</td><td>NEGATIVE, POSITIVE, BORDERLN</td></tr><tr><td>U_PROT_DIP</td><td>https://appliedresearch.cancer.gov/crnportal/data-resources/vdw/user-guidelines-and-tools/Guidelines_when_selecting_raw_data_records_to_include_in_U_PROT_DIP.doc/view</td></tr></table>
Modifier	Char(2)			Modifies the values in Result_C	'TX'=text 'EQ'=equal 'LT'=less than 'LE'=less than or equal to 'GT'=greater than 'GE'=greater than or equal to
Result_unit	Char (11)			Units in which result is reported.	possible values include %, IU/L, MEQ/L, MG/DL, THERAPEUTIC
Normal_low_C	Char(8)			Lowest value of normal range for this test as short character.	To accept digits, decimal, sign or text. The symbols >, <, >=, <= should be removed.
Modifier_low	Char(2)			Modifies the values in Normal_low_C	'TX'=text 'EQ'=equal (only used for a range when upper and lower boundaries are both supplied)

Field	Type			Description	Values
					'LT'=less than (this value will probably not be used)
					'LE'=less than or equal to (this value will probably not be used)
					'GT'=greater than
					'GE'=greater than or equal to
					null is appropriate when normal_low_c is null
Normal_high_C	Char(8)			Highest value of normal range for this test as short character.	To accept digits, decimal, sign or text. The symbols >, <, >=, <= should be removed.
Modifier_high	Char(2)			Modifies the values in Normal_high_C	'TX'=text 'EQ'=equal (only used for a range when upper and lower boundaries are both supplied) 'LT'=less than 'LE'=less than or equal to 'GT'=greater than (this value will probably not be used) 'GE'=greater than or equal to (this value will probably not be used) null is appropriate when normal_high_c is null
Abn_ind	Char(2)			Abnormal result indicator	'AB'=abnormal 'C' =critical 'AH'=abnormal high 'CH'=critical high 'AL'=abnormal low 'CL'=critical low 'IN'=Inconclusive 'UK'=unknown (null values should be set to UK) 'NL'=normal
Order_prov	Char, (length is site specific)			Identifier unique to an ordering provider.	Same coding scheme as rxmd in pharmacy or provider in utilization. Unique to each HMO.
Order_dept	Char(4)			Department of ordering provider as bucketed by scheme used in the utilization file.	Same coding scheme as department in utilization.
Facility_code	Char, (length is site specific)			Where order originated	Local facility code that identifies hospital or clinic. Same coding scheme as facility_code in utilization. Unique to each HMO.

Row_ID (changed from RowID to avoid Oracle keyword)	type and length are site specific			Facility-specific identifier that links a LAB RESULTS record to corresponding LAB NOTES record(s).	Missing value or a pointer

2b. Lab Notes table data dictionary

Variable name	Type and Length	Format	Label	Definition	Valid Values	General Notes
Row_ID (changed from RowID to avoid Oracle Keyword)	type and length are site specific			Facility-specific identifier that links a LAB NOTES record to the corresponding LAB RESULTS record.	A pointer	It is possible for more than one record in LAB N record in LAB RESULTS. Therefore it is possibl NOTES to have the same Row_ID value.
Result_Note	Char(80)			To accept text string results, comments or normal low or high range values that are longer than 8 bytes	This field could contain:result, lowest value of normal range, highest value of normal range, notes	
Note_Type (changed from Type to avoid Oracle keyword)	Char(1)			Describes what information is held in the Result_Note field.	'R' = result 'L' = low normal value 'H' = high normal value 'N' = note(other type values can be added as need arises)	
Line	Num(4)			Counter starting at 1 and incrementing by 1 as needed for each Row_ID by Type.	Integer.	As many rows as necessary can be associated w combination. For example, a Row_ID might ha it. The Note_Type value would = N and the line for the second note and so on through 5.

3. Notes:

Note 1: Charts illustrating a) the Row_ID link between LAB RESULTS and LAB NOTES and b) the relationship between the variables Note_Type and Line.

Selected variables and observations from LAB RESULTS table where ROWID not equal missing.

obs	Row_ID	Test_Type	Result_C	Result_unit
3596	471	K	4.0	MEQ/L
3709	472	K	5.9	MEQ/L
4843	473	K	3.6	MEQ/L
6045	474	K	3.4	MEQ/L
6786	475	K	5.4	MEQ/L

Selected observations from LAB NOTES table where ROWID corresponds to observations above.

obs	Row_ID	Result_Note	Note_Type	Line
512	471	SPECIMEN SLIGHTLY HEMOLYZED, MAY FALSELY INCREASE K	N	1
513	472	SPECIMEN MODERATELY HEMOLYZED, MODERATE HEMOLYSIS WILL FALSELY INCREASE SERUM	N	1
514	472	POTASSIUM	N	2
515	473	HEMOLYSIS, MAY INCREASE K.	N	1
516	474	SPECIMEN SLIGHTLY HEMOLYZED, MAY FALSELY INCREASE K	N	1
517	475	MOD HEMOLYSIS, FALSE INCREASES IN K.	N	1

Note 2: Changes that "break the structure" are defined as those changes that are not backwards compatible, that is, old programs will no longer work with the new data stru

3. Future plans:

A data structure for microbiology is being worked on.

4. Macros: The following macros are available for use with the lab file.

| Description	Macro
A macro to not break notes in the middle of a string that might be searched for (in other words in the middle of a word).	Split_Wrap_80 (Please see below for link to download macro)

5. Contact information:

Please contact the VDW Lab working group for questions and suggestions.

6. Prior Spec

This is what the prior version of the Lab Results datasets looked like

.

Related content
- VDW Lab QA Schedule
- Lab Metadata: Tests and Dates Available by Site
- VDW Lab Test Reference Document
- VDW Lab Priorities HAS BEEN COMBINDED WITH TEST_TYPE, LONGNAME AND LOINC DOCUMENT
- Split_Wrap_80 SAS Macro
- Lab Data: Collecting, Cleaning and Modeling
- Lab Values Discussion: Ideas for Producing High Quality Lab Data Sets
- Lab Workgroup Monthly QA Reviews Through 2012
- Test_Type, LongName, LOINC and Priority Spreadsheet
- V2 Lab Results Data Dictionaries
- Use and Limitations of the Three Date Variables in Labs
- Lab WorkGroup Multi-Site Compiled QA Reports Reviewed During Monthly Call Starting 2/2013

Comments (0)