PhUSE 2010

Paper CD03

From ACE to ZINC

Examples on the use of SDTM Controlled Terminology for lab tests

Paul Vervuren, Merck Sharp & Dohme, Oss, The Netherlands

ABSTRACT

The mapping of lab tests to the Laboratory Test Code controlled terminology in CDISC-SDTM[§] can be a challenge. One has to find candidates in the extensive controlled terminology list. Then there can be multiple lab tests that map to a single SDTM controlled term. This means additional variables must be used in order to produce a unique test definition (e.g. LBCAT, LBSPEC, LBMETHOD and/or LBELTM). Finally, it can occur that a controlled term is not available and a code needs to be defined in agreement with the rules for Lab tests. This paper describes my experience with the implementation of SDTM controlled terminology for lab tests during an SDTM conversion activity. In six clinical studies 124 lab tests were mapped to 101 SDTM controlled terms. The lab tests included routine lab parameters, coagulation parameters, hormones, glucose tolerance test and pregnancy test.

INTRODUCTION

This paper aims to give detailed examples of SDTM LB datasets that were created for six studies included in an FDA submission. Background information on the conversion project that formed the context of this work can be found in an earlier PhUSE contribution [1].

With the exception of part of the hormone data all laboratory data of these studies had been extracted from the Oracle Clinical™ NORMLAB2 system, which delivered complete and standardized lab data, i.e. standardized parameter (lab test) names, values, units and ranges. Subsequently, these NORMLAB2 extracts had been enriched with derived variables and records, following internal data standards and conventions, to form standardized analysis-ready datasets. These were the basis for conversion to SDTM LB datasets.

The combined source datasets of the six studies held 124 distinct lab tests, which were mapped to 101 distinct lab controlled terms. Controlled terminology for lab tests is part of the SDTM terminology, which is published on the NCI EVS[δ] website [2]. New lab test terms have been released for public review through a series of packages [3], starting in 2007. Since version 3.1.2. of the SDTM Implementation Guide [4], the use of SDTM controlled terminology for lab tests is assumed for LBTESTCD and LBTEST (codelists C65047 and C67154).

Table 1 provides an overview of the number of lab tests per study in the source data vs. the SDTM datasets (i.e. the number of LBTEST/LBTESTCD codes) and shows how these codes were distributed across different lab test categories. A set of 22 ‘routine safety parameters’ occurred in all four phase III studies (001-004), with 16 tests occurring in all six studies.

[§] Clinical Data Interchange Standards Consortium - Study Data Tabulation Model

[δ] National Cancer Institute - Enterprise Vocabulary Services

Table 1: The number of lab tests in six different studies and their distribution across lab test categories.

Study	Source	SDTM											
			LBTEST per category (SDTM Labtest package subdivision)										
	normlab2 tests	LBTEST terms	Chemistry - General	Chemistry - Enzymes	Chemistry - Proteins	Hema-tology	Chemistry - Hormones	Sediment Analysis	Coag-ulation	Immu-nology	General Obs.	Drug Screen	Vita-mins
001	22	22	6	5	2	9							
002	22	22	6	5	2	9							
003	33	33	6	5	2	9	10						1
004	68	58	14	5	6	9	12		12				
006	69	57	15	5	3	14	1	8	3	4	2	2	
011	62	55	15	6	3	14		5	2	6	2	2	

NOT ALL IS COMPLICATED

The mapping of the NORMLAB2 test names to lab controlled terminology was straightforward for many tests. It was performed by looking up the test name in the LBTEST codelist. Table 2 shows the mapping of twenty-two common chemistry and hematology safety parameters (these are all the tests performed in studies 001 and 002, see table 1). In most cases, there was an exact match between the names. In some cases the correctness of the mapping was confirmed by examining the test definition (or the synonym). For example, the LBTEST value 'Bilirubin' is defined as 'A measurement of the total bilirubin in a biological specimen', and it has a synonym of 'Total Bilirubin'. This verified that our NORMLAB2 test 'Bilirubin total' maps to LBTEST='Bilirubin'. Similarly, the mapping of NORMLAB2 test 'Uric acid' to LBTEST='Urate' was confirmed by the presence of synonym value 'Uric Acid'.

Table 2: Mapping of 22 common chemistry and hematology safety parameters.

Category	Lab test (source data)	LBTEST	LBTESTCD
Chemistry - Enzymes	Alanine aminotransferase (ALAT, SGPT)	Alanine Aminotransferase	ALT
	Alkaline phosphatase	Alkaline Phosphatase	ALP
	Aspartate aminotransferase (ASAT, SGOT)	Aspartate Aminotransferase	AST
	Gamma glutamyl transferase (GGT)	Gamma Glutamyl Transferase	GGT
	Lactate dehydrogenase (LDH)	Lactate Dehydrogenase	LDH
Chemistry - General	Bilirubin total	Bilirubin	BILI
	Creatinine	Creatinine	CREAT
	Potassium (K)	Potassium	K
	Sodium (Na)	Sodium	SODIUM
	Urea	Urea	UREA
	Uric acid	Urate	URATE
Chemistry - Proteins	Albumin	Albumin	ALB
	Protein total	Protein	PROT
Hematology	Basophils	Basophils	BASO
	Eosinophils	Eosinophils	EOS
	Hematocrit (Ht)	Hematocrit	HCT
	Hemoglobin (Hb)	Hemoglobin	HGB
	Leukocyte count (WBC count)	Leukocytes	WBC
	Lymphocytes	Lymphocytes	LYM
	Monocytes	Monocytes	MONO
	Neutrophils	Neutrophils	NEUT
	Platelet count (Thrombocyte count)	Platelets	PLAT

QUANTITATIVE VS. QUALITATIVE TESTS AND DIFFERENT SPECIMEN TYPES

One of the studies described here measured Erythrocytes, either quantitatively or qualitatively, in serum, urine or urine sediment. In the source data, there were four different lab tests. Figure 1 presents sample data from one the studies. These were all mapped to LBTEST='Erythrocytes', in line with the following 'Rules for Test Codes for Laboratory Test Names' (as defined in the Labtest package [2]):

1. *Test names do not contain the matrix of the specimen. This information is populated in the specimen type field (LBSPEC).*
2. *Lab tests where results can be expressed as qualitative, semi-quantitative or quantitative should all have the same variable name.*

Thus, the four Erythrocytes tests in this example are identified in the SDTM dataset using LBTEST, LBCAT and LBSPEC. For LBSPEC the 'Specimen Type' codelist (C78734) was applied. We extended the list with the term 'URINE SEDIMENT'. An alternative solution would have been to use LBSCAT or LBMETHOD to identify the assessments in urine sediment (e.g. 'SEDIMENT ANALYSIS'). The advantage would be that a standard code (i.e. 'URINE') could be used for LBSPEC; the drawback is that an additional variable is needed to identify a unique Erythrocyte test.

The categorization Hematology – Urine - Chemistry, present in source variable BLKF (fig. 1), was mapped to LBSCAT (where 'Urine' was mapped to 'URINALYSIS'). Following recommendations of the SDTM Implementation Guide (SDTMIG), controlled terminology was put in upper case text.

SAS System Viewer - lb.sas7bdat

File Edit View Window Help

saflab.sas7bdat

	PRFPARAM	BLKF	LVALUE	ORGUNIT	ORGNORLW	ORGNORUP	PANF
1	Erythrocyte count (RBC count)	Hematology	5	1/pL	3.8	5.3	-50
2	Erythrocyte count (RBC count)	Hematology	5	1/pL	3.8	5.3	-40
3	Erythrocyte count (RBC count)	Hematology	4.6	1/pL	3.8	5.3	10.001
4	Erythrocyte count (RBC count)	Hematology	4.5	1/pL	3.8	5.3	510
5	Erythrocytes (urine)	Urine	10 /uL		.	.	-50
6	Erythrocytes (urine)	Urine	25 /uL		.	.	-40
7	Erythrocytes (urine)	Urine	NEG		.	.	10.001
8	Erythrocytes (urine)	Urine	10 /uL		.	.	510
9	Erythrocytes (urine sediment)	Urine	0-2/HPF		.	.	510
10	Erythrocyte count (urine sediment RBC count)	Urine	8.4	1/uL	0	25	-50
11	Erythrocyte count (urine sediment RBC count)	Urine	6.5	1/uL	0	25	-40

lb.sas7bdat

	LBTEST	LBCAT	LBSPEC	LBSCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI	VISITNUM
1	Erythrocytes	QUANTITATIVE LAB TEST	BLOOD	HEMATOLOGY	5	1/pL	3.8	5.3	-50
2	Erythrocytes	QUANTITATIVE LAB TEST	BLOOD	HEMATOLOGY	5	1/pL	3.8	5.3	-40
3	Erythrocytes	QUANTITATIVE LAB TEST	BLOOD	HEMATOLOGY	4.6	1/pL	3.8	5.3	10.001
4	Erythrocytes	QUANTITATIVE LAB TEST	BLOOD	HEMATOLOGY	4.5	1/pL	3.8	5.3	510
5	Erythrocytes	QUALITATIVE LAB TESTS	URINE	URINALYSIS	10 /uL				-50
6	Erythrocytes	QUALITATIVE LAB TESTS	URINE	URINALYSIS	25 /uL				-40
7	Erythrocytes	QUALITATIVE LAB TESTS	URINE	URINALYSIS	NEG				10.001
8	Erythrocytes	QUALITATIVE LAB TESTS	URINE	URINALYSIS	10 /uL				510
9	Erythrocytes	QUALITATIVE LAB TESTS	URINE SEDIMENT	URINALYSIS	0-2/HPF				510
10	Erythrocytes	QUANTITATIVE LAB TEST	URINE SEDIMENT	URINALYSIS	8.4	1/uL	0	25	-50
11	Erythrocytes	QUANTITATIVE LAB TEST	URINE SEDIMENT	URINALYSIS	6.5	1/uL	0	25	-40

Figure 1: Mapping of Erythrocyte parameters to SDTM LB, requiring the use of LBCAT and LBSPEC. Upper panel shows a source data example (one subject); lower panel the SDTM dataset. Only a selection of variables is shown; column order adjusted to show mapped columns above each other.

GLUCOSE TOLERANCE TEST (USE OF TIME POINT VARIABLES)

Another example where multiple lab parameters were mapped to a single SDTM LBTEST value is the oral glucose tolerance test (OGTT). OGTT is performed to assess glucose metabolism, and involves glucose measurement (and in this case also insulin) in a time series after drinking a glucose solution. The first measurement is the reference time point, which is taken under fasting conditions, just before glucose is administered. Figure 2 presents sample data from one of the studies. In this example, blood samples were collected 30, 60, 90, 120 and 180 minutes after 'glucose loading'.

The source data has a separate parameter for each time point. These individual parameters are all mapped to LBTEST='Glucose' (LBTESTCD='GLUC'). The time point information is captured in a number of timing variables (see fig. 3): two variables for the reference time point (or 'anchor'), three for the collection time point and one for the elapsed time between the reference and collection time point. The source variable that holds the lab parameter name (PRFPARAM) contains information on the collection time point ('... min after glucose loading') and therefore has a mapping to LBTPT (Planned Time Point Name). LBDTC contains the date/time of specimen collection and was created by combining the source

variables for date and time (SMPD and SMPT). LBTPTNUM (Planned Time Point Number) is described in the SDTMIG as a 'numerical version of LBTPT to aid in sorting'. While LBTPTNUM is a permissible variable, it is expected when time point is used (see SDTMIG v.3.1.2 section 4.1.4.10). Derived variable SAMPLENR serves this purpose, and is used to populate LBTPTNUM. LBFAST (Fasting Status) is defined using variable PRFPARAM (the fasting parameter contains the text string 'fasting').

SAS System Viewer - saflab.sas7bdat

File Edit View Window Help

saflab.sas7bdat

	PRFPARAM	ORGVALUE	ORGUNIT	SMPD	SMPT	ADMINT	SAMPLENR
1	Glucose tolerance test (oral), fasting	4.27	mmol/L	27FEB2007	8:35:00	8:40	0
2	Glucose tolerance test (oral), 30 min after glucose load	5.88	mmol/L	27FEB2007	9:10:00	8:40	1
3	Glucose tolerance test (oral), 60 min after glucose load	5.66	mmol/L	27FEB2007	9:40:00	8:40	2
4	Glucose tolerance test (oral), 90 min after glucose load	4.83	mmol/L	27FEB2007	10:10:00	8:40	3
5	Glucose tolerance test (oral), 120 min after glucose loa	4.83	mmol/L	27FEB2007	10:40:00	8:40	4
6	Glucose tolerance test (oral), 180 min after glucose loa	4.83	mmol/L	27FEB2007	11:40:00	8:40	5

lb.sas7bdat

	LBTEST	LBTPT	LBFAST	LBTPTREF	LBORRES	LBORRESU	LBDTC	LBRFTDTC	LBTPTNUM	LBELTM
1	Glucose	BEFORE GLUCOSE LOADING	Y	GLUCOSE LOADING	4.27	mmol/L	2007-02-27T08:35	2007-02-27T08:40	0	-PT5M
2	Glucose	30 MINUTES AFTER GLUCOSE LO	N	GLUCOSE LOADING	5.88	mmol/L	2007-02-27T09:10	2007-02-27T08:40	1	PT30M
3	Glucose	60 MINUTES AFTER GLUCOSE LO	N	GLUCOSE LOADING	5.66	mmol/L	2007-02-27T09:40	2007-02-27T08:40	2	PT1H
4	Glucose	90 MINUTES AFTER GLUCOSE LO	N	GLUCOSE LOADING	4.83	mmol/L	2007-02-27T10:10	2007-02-27T08:40	3	PT1H30
5	Glucose	120 MINUTES AFTER GLUCOSE L	N	GLUCOSE LOADING	4.83	mmol/L	2007-02-27T10:40	2007-02-27T08:40	4	PT2H
6	Glucose	180 MINUTES AFTER GLUCOSE L	N	GLUCOSE LOADING	4.83	mmol/L	2007-02-27T11:40	2007-02-27T08:40	5	PT3H

Figure 2: Mapping of Glucose Tolerance Test data to SDTM LB, involving the use of LBTPT, LBFAST, LBRFTDTC, LBTPTREF, LBTPTNUM and LBELTM. Upper panel shows a source data example (one subject, one visit); lower panel the SDTM dataset. Only a selection of variables is shown; column order adjusted to show mapped columns above each other where possible.

Variable ADMINT included in the source data presented the time of the reference time point, i.e. glucose administration. This time was used in combination with SMPD to produce LBRFTDTC (Date/Time of Reference Time Point). The source data did not have descriptive information on the reference time point. Therefore, it was added to the LB datasets (LBTPTREF='GLUCOSE LOADING').

Variable LBELTM (Planned Elapsed Time from Time Point Ref), see fig. 3, was defined on the basis of the planned time intervals of 30, 60, 90, 120 and 180 minutes between glucose loading and sample collection. In ISO 8601, which is the expected format, this series can be represented as PT30M, PT1H, PT1H30M, PT2H, and PT3H (see SDTMIG v.3.1.2, section 4.1.4.3). Alternatively, 60 minutes could have been represented as PT60M; 90 minutes as PT90M, etc.. There was no precise information on the planned timing of the reference time point. The protocol indicated: 'Blood glucose and insulin levels will be determined as fasting values just before oral glucose intake ...' From the actual timing of administration across all subjects it was established that the reference point was, in general, 5 minutes before glucose intake. Based on this, we applied LBELTM='–PT5M' for the reference time point.

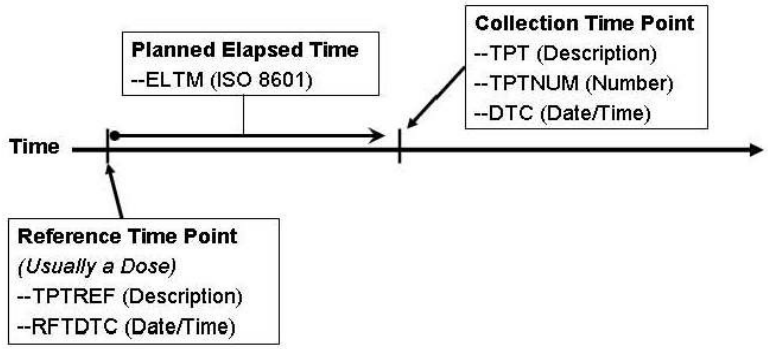


Figure 3: Interrelationship of time point variables (from SDTMIG v.3.1.2, section 4.1.4.10).

WHITE BLOOD CELL DIFFERENTIAL COUNT (NEUT VS. NEUTLE)

There are five basic types of leukocytes (white blood cells): neutrophils, lymphocytes, monocytes, eosinophils and basophils. The white blood cell differential count determines the number of each type of white blood cell present in the blood. It can be expressed as a percentage (relative numbers of each type of WBC in relationship to the total WBC) or as an absolute value, i.e. the number of cells per unit blood. The naming of the lab tests is different in these cases. See table 3, which displays the leukocyte parameters as determined in two of our studies.

Table 3: Mapping of relative and absolute WBC differential parameters to LB controlled terminology.

Source parameter		SDTM		
PRFPARAM	**ORGUNIT**	**LBTESTCD**	**LBTEST**	**LBSCAT***
Neutrophils (alternate unit)	%	NEUTLE	Neutrophils/Leukocytes	DIFFERENTIAL
Monocytes (alternate unit)	%	MONOLE	Monocytes/Leukocytes	DIFFERENTIAL
Lymphocytes (alternate unit)	%	LYMLE	Lymphocytes/Leukocytes	DIFFERENTIAL
Eosinophils (alternate unit)	%	EOSLE	Eosinophils/Leukocytes	DIFFERENTIAL
Basophils (alternate unit)	%	BASOLE	Basophils/Leukocytes	DIFFERENTIAL
Neutrophils	10^9/L	NEUT	Neutrophils	DIFFERENTIAL
Monocytes	10^9/L	MONO	Monocytes	DIFFERENTIAL
Lymphocytes	10^9/L	LYM	Lymphocytes	DIFFERENTIAL
Eosinophils	10^9/L	EOS	Eosinophils	DIFFERENTIAL
Basophils	10^9/L	BASO	Basophils	DIFFERENTIAL

* Value 'DIFFERENTIAL' suggested by SDTMIG v.3.1.2 (example 6.3.3.2).

The names for differentials are defined on the basis of the following CDISC rule for lab test names [3]:

"for all differential test names the absolute count is a short defined term (eg, EOS, BASO) and the ratio/percentage has the same short pneumonic with a second short pneumonic for the denominator eg, LE for leukocytes (EOSLE, Eosinophils/Leukocytes), LY for lymphocytes (LYMMCELY, Lymphoma Cells/Lymphocytes), RBC for erythrocytes (RETIRBC, Reticulocytes/Erythrocytes). LY will be used when there is a denominator. When lymphocyte(s) is part of the core test name use LYM."

This rule is also applicable to other relative measurements (ratios or percentages), for instance, CHOLHDL: Cholesterol to HDL-Cholesterol, ALBCREAT: Albumin/Creatinine.

USER-DEFINED TESTS AND CONTROLLED TERMINOLOGY UPDATES

The lab test controlled terminology is work in progress. CDISC releases new lab tests with every new production version of SDTM controlled terminology. The SDTM terminology production package of 28-MAR-2008 comprised 89 lab tests. This had increased to 581 lab tests in the release of 08-APR-2010, and the latest release (6-OCT-2010) includes 650 lab tests. For our data, new releases resulted in an additional 8 standard controlled terms that could be implemented (table 4), replacing terms that we had defined ourselves in 2008.

Table 4: Lab tests in studies 004 - 011 with newly released labtest controlled terminology (since Labtest package 3).

Source parameter		User-defined terminology *		SDTM controlled terminology	
PRFPARAM	**PRFUNIT**	**LBTESTCD**	**LBTEST**	**LBTESTCD**	**LBTEST**
Coagulation					
Antithrombin III (AT-III)	%	ATIII	Antithrombin III (AT-III)	ANTHRM	Antithrombin
Coagulation Factor VIII	%	FACTVIII	Coagulation Factor VIII	FACTVIII	Factor VIII
D-Dimer	mg/L FEU	DDIMER	D-Dimer	DDIMER	D-Dimer
Prothrombin fragments 1+2	nmol/L	PTHRFG12	Prothrombin Fragments 1+2	PTF1_2	Prothrombin Fragments 1 + 2
Lipoproteins					
Lipoprotein a	g/L	LIPOA	Lipoprotein (a)	LPA	Lipoprotein-a
High density lipoprotein 2 (HDL2) cholesterol	mmol/L	HDL2	HDL2 cholesterol	HDL2	HDL-Cholesterol Subclass 2
High density lipoprotein 3 (HDL3) cholesterol	mmol/L	HDL3	HDL3 cholesterol	HDL3	HDL-Cholesterol Subclass 3
Enzymes					
Tryptase	ug/L	TRYPTASE	Tryptase	TRYPTASE	Tryptase

* 5 LBTESTCD's and 3 LBTEST's appeared to be in line with SDTM terminology.

Nevertheless, for a number of tests no suitable candidate was found in the SDTM controlled terminology list (table 5). The 'Drug screen (urine)' was a qualitative multi-drug test that screened for benzodiazepines, opiates, (meth)amphetamines, PCP, methadone, tricyclic antidepressants, cocaine, cannabinoids, and barbiturates (for each of which separately an SDTM lab test name is present). The 'Other cells (differential count)' was part of the WBC differential count (excluded from table 3), and included cells that could not be classified to any of the five leukocyte types.

Table 5: User-defined terminology (LBTESTCD/LBTEST) for lab tests in studies 003 - 011.

PRFPARAM	PRFUNIT	LBTESTCD	LBTEST	LBCAT
Coagulation				
Activated protein C (APC) resistance ratio (APTT-based)	1	APCRAPT	APTT-based Act. Prot. C Resistance Ratio	QUANTITATIVE LAB TESTS
Coagulation Factor II	%	FACTII	Coagulation Factor II	QUANTITATIVE LAB TESTS
Coagulation Factor VIIa	U/L	FACTVIIA	Coagulation Factor VIIa	QUANTITATIVE LAB TESTS
Coagulation Factor VIIc	%	FACTVIIC	Coagulation Factor VIIc	QUANTITATIVE LAB TESTS
Activated protein C (APC) resistance ratio (ETP-based)	1	APCRETP	ETP-based Act. Prot. C Resistance Ratio	QUANTITATIVE LAB TESTS
Partial thromboplastin time (PTT)	s	PTT	Partial Thromboplastin Time	QUANTITATIVE LAB TESTS
Protein C	%	PROTC	Protein C	QUANTITATIVE LAB TESTS
Protein S free	%	PROTSFR	Protein S free	QUANTITATIVE LAB TESTS
Protein S total	%	PROTS	Protein S total	QUANTITATIVE LAB TESTS
Drug screen				
Drug screen (urine)		DRUGSCRN	Drug Screening	QUALITATIVE LAB TESTS
Endocrinology				
Dehydroepiandrosterone sulphate (DHEA-S)	umol/L	DHEAS	Dehydroepiandrosterone Sulphate	QUANTITATIVE LAB TESTS
Cortisol binding globulin (CBG)	nmol/L	CBG	Cortisol Binding Globulin	QUANTITATIVE LAB TESTS
Immunology				
Hepatitis B Core Antibody (Anti-HBc)		HBCAB	Hepatitis B Core Antibody	QUALITATIVE LAB TESTS
Hematology				
Other cells (differential count)	%	OTHCELLE	Other cellls/Leukocytes	QUANTITATIVE LAB TESTS

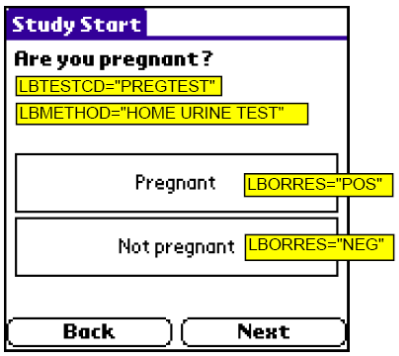
PREGNANCY TEST

The investigational product in the studies described in this paper was a contraceptive. Pregnancy testing was an essential assessment in these studies, and was planned at screening, at the end of treatment, and during follow-up. Furthermore, unscheduled pregnancy testing was performed whenever pregnancy was suspected. There were three types of pregnancy tests, as described in table 6.

Table 6: Representation of the three pregnancy test types in LB.

Type	Timing	Source	LBTEST	LBMETHOD	LBSPEC
Beta-hCG lab test	screening, end of treatment and follow-up	lab data	Pregnancy test	SERUM TEST (B-HCG)	SERUM
Home pregnancy test	before first IP intake	eDiary*	Pregnancy test	HOME URINE TEST	URINE
Pregnancy test at the clinic	suspected pregnancy	CRF **	Pregnancy test	URINE/BLOOD TEST	URINE/BLOOD
				ULTRASOUND	

* see figure 4; ** see figure 5.



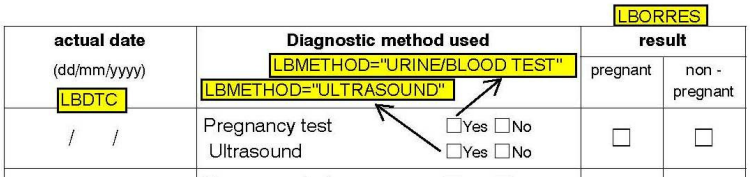


Figure 4 (left): eDiary, used to record the outcome from home pregnancy test.
Figure 5 (right): Pregnancy determination CRF, used when a pregnancy tests was performed in the clinic

The difficulty with mapping these data to SDTM was that each type was recorded differently, and, consequently, had different dataset structures. Also, only the serum Beta-hCG test was an actual laboratory test. The pregnancy test at the clinic could involve an ultrasound, which is not a laboratory test. Nevertheless, it was decided, to keep all pregnancy test data together in SDTM and load it into the LB dataset

Identical, user-defined labtest codes were used for all pregnancy test types (LBTEST='Pregnancy test' and LBTESTCD='PREGTEST'), including the B-hCG lab test, which could have been mapped to controlled term 'HCG'. The different test types were distinguished using variable LBMETHOD. At the clinic, two diagnostic methods could be used, which the investigator had to specify on the CRF (fig. 4). In the protocol, the 'Pregnancy test' method at the clinic was referred to as a 'urine or blood test'. Based on this it was decided to represent this option in the LB dataset using LBMETHOD ='URINE/BLOOD TEST'.

An example of pregnancy data is given in fig. 6, presenting data of two subjects. The source dataset (pregdet) was an analysis dataset that combined three source datasets. The Beta-hCG lab test data (source='BLOOD SAMPLE') originated from a lab dataset, where the NORMLAB2 test code was 'B_HCG_URINE'. The variables PRGTSTY and USSY represent the 'Pregnancy test' and 'Ultrasound' fields on the CRF (source='PREGDET'). .

SAS System Viewer - lb.sas7bdat

File Edit View Window Help

pregdet.sas7bdat

	USUBJID	SOURCE	PRGTSTY	USSY	PRGTSTD	PRGTSTDY	TUBEID	PRGRESF	PANF
1	001-MY_00	BLOOD SAMPLE	.	.	11AUG2006	-21	0608-22129	NEG	Screening
2	001-MY_00	DIARY	.	.	01SEP2006	1		NEG	Baseline
3	001-MY_00	BLOOD SAMPLE	.	.	07SEP2007	372	0709-22079	NEG	End of treatment
4	001-PL_14	BLOOD SAMPLE	.	.	26OCT2006	-12	0610-25770	NEG	Screening
5	001-PL_14	DIARY	.	.	07NOV2006	1		NEG	Baseline
6	001-PL_14	BLOOD SAMPLE	.	.	14NOV2007	373	0711-24010	NEG	End of treatment
7	001-PL_14	PREGDET	1	0	20NOV2007	379		NEG	Unscheduled
8	001-PL_14	PREGDET	0	1	10JAN2008	430		NEG	Unscheduled

lb.sas7bdat

	USUBJID	LBMETHOD	LBTEST	LBDTC	LBDY	LBREFID	LBORRES	VISIT	LBSPEC
1	001-MY_00	SERUM TEST (B-HCG)	Pregnancy test	2006-08-11	-21	0608-22129	NEG	SCREENING	SERUM
2	001-MY_00	HOME URINE TEST	Pregnancy test	2006-09-01	1		NEG	BASELINE	URINE
3	001-MY_00	SERUM TEST (B-HCG)	Pregnancy test	2007-09-07	372	0709-22079	NEG	END OF CYCLE	SERUM
4	001-PL_14	SERUM TEST (B-HCG)	Pregnancy test	2006-10-26	-12	0610-25770	NEG	SCREENING	SERUM
5	001-PL_14	HOME URINE TEST	Pregnancy test	2006-11-07	1		NEG	BASELINE	URINE
6	001-PL_14	SERUM TEST (B-HCG)	Pregnancy test	2007-11-14	373	0711-24010	NEG	END OF CYCLE	SERUM
7	001-PL_14	URINE/BLOOD TEST	Pregnancy test	2007-11-20	379		NEG	UNSCHEDULED	URINE/BLOOD
8	001-PL_14	ULTRASOUND	Pregnancy test	2008-01-10	430		NEG	UNSCHEDULED	

Figure 6: Mapping of pregnancy test data to LB. Upper panel shows a source data example (two subjects); lower panel the SDTM dataset. Only a selection of variables is shown; column order adjusted to show mapped columns above each other where possible.

CONCLUSION

This paper has shown various examples where lab data was mapped to the SDTM LB domain. Key to a successful application of the model is an understanding of the meaning and role of the different SDTM variables. There is one basic principle to keep in mind: the SDTM lab test (LBTESTCD/LBTEST) refers to what is measured only, not how or when it is measured or in which substance it is determined. This means a 'urine dipstick test' is not a valid test name in SDTM LB, but just refers to a method and specimen type. To map such a test, one needs to find out what the urine dipstick test actually measures (e.g. glucose) and find out whether SDTM provides a controlled term that describes it.

There are cases in which the mapping process involves the use of more than LBTEST and LBSPEC to define the test. In the Erythrocytes example, LBCAT was used to make the distinction between the quantitative and qualitative analysis in urine. In the example of pregnancy test data LBMETHOD was applied to distinguish between the diagnostic methods used. A glucose tolerance test involves the

measurement of glucose (LBTEST='Glucose') at different timepoints, requiring the use of special timing variables. For example LBELTM, which gave the elapsed time since glucose administration. This adds complexity to the mapping process, also because additional controlled terminology (e.g. ISO 8601) may be involved.

The number of SDTM controlled terms available for lab tests has greatly increased in the past years. The SDTM LB datasets created for the studies in this paper contained 101 distinct lab tests, 87 of which were from the SDTM controlled terminology list. For 14 lab tests (9 of which were coagulation parameters) no candidate was found in the list, and a new test was defined in line with the rules for controlled terminology. If the release of new controlled terms for lab tests continues at today's high pace, then the need for user-defined test names may be minimal in the near future.

REFERENCES

1. P. Vervuren & B. van Bakel, 2009. Experiences and lessons learned from a first SDTM submission project. PhUSE Conference 2009.
2. http://www.cancer.gov/cancertopics/terminologyresources/page6
3. http://www.cdisc.org/stuff/contentmgr/files/0/6bb810e0ab25d0b545742f104e9210d1/files/lab_controlled_terminology_for_v312_for_public_review_package___4.xls
4. SDTM Implementation Guide for Human Clinical Trials, version 3.1.2. Clinical Data Interchange Standards Consortium, Inc., 2008

ACKNOWLEDGMENTS

I would like to thank Monique van de Ven for providing her help and lab expertise.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author name :	Paul Vervuren
Company :	Merck Sharp & Dohme
Address :	PO Box 20
	5340 BH Oss
	The Netherlands
Work phone :	+31 (0) 412 663921
E-mail :	Paul.Vervuren[-at-]merck.com
Website :	www.merck.com