T-69

Challenges and Solutions for Mapping Pathology Data to SEND

Mike Wasko, Rich Buchanan, Fred Mura, and Laura Kaufman
PDS Life Sciences, Mt Arlington, New Jersey

SEND Express™ Your Virtual SEND Team

Abstract

Based on our experience creating SEND datasets for FDA submission, we can group SEND mapping challenges for pathology into several patterns: 1) inconsistent or incomplete data entry for gross and microscopic pathology findings; 2) inconsistent use of units for clinical pathology findings; 3) inconsistent use of terminology for clinical pathology (use of the same term to mean different things in different parts of a dataset); and 4) uncertainty of which domain to map biomarkers. For all of these disciplines, it is critical that data capture be consistent so that mapping can be automated, since manual solutions are impractical in a production environment. For example, microscopic findings as collected by the pathologist, including base pathological processes and modifiers, are placed into one SEND variable (MIORRES) in the Microscopic Domain (MI). The base pathological process and different modifiers get mapped again to different SEND variables, some of which require controlled terminology. Without a consistent way to record data across animals and studies at the time of data capture, it is difficult to automate this type of mapping. It is also critical that terminology be used in a consistent fashion for accurate mapping to controlled terminology. In conclusion, the use of SEND presents new challenges for data capture while at the same time creates opportunities for analyzing large quantities of data in ways that are not currently feasible.

Inconsistent Data Entry for Microscopic Pathology

Background: For the MA (macroscopic) and MI (microscopic) SEND domains, it is necessary to include each original finding as recorded (base pathological finding together with as shown in the table following.

SEND 3.0			SEND 3.1		
Variable	**Definition**	**CT Required?**	**Variable**	**Definition**	**CT Required?**
MIORRES	Original findings as recorded, including base pathological processes / findings & modifiers	No	MIORRES	Same as 3.0	No
MISTRESC	Base pathological process / finding without modifiers	Only for tumors	MISTRESC	Same as 3.0	Yes – for non neoplastic & neoplastic findings
MISEV	Severity modifier	Yes	MISEV	Same as 3.0	Yes
MILAT	Laterality	Yes	MILAT	Same as 3.0	Yes
MIDIR	Modifier	Yes	MIDIR	Same as 3.0	Yes
			MIDISTR	Distribution modifier	Yes
			MICHRON	Duration modifier	Yes

Inconsistent Data Entry for Microscopic Pathology continued

To automate SEND translation, it is important that raw data be consistently formatted across users, animals, and studies for a particular LIMS. For pathology, this means an ordered dataset with respect to primary findings and modifiers.

Examples of Actual Data that Present Mapping Challenges for Automation:

- Microscopic Pathology source data, same study, 2 different animals, 3 different findings, 3 different orders for modifiers:

  OPTIC NERVE: NEUROPATHY, MINIMAL, FOCAL
  TESTIS: INTERSTITIAL CELL HYPERPLASIA, DIFFUSE, SLIGHT
  KIDNEY: FOCAL NEPHROPATHY, MINIMAL

- Same data in a form that can be mapped automatically:

  OPTIC NERVE: NEUROPATHY, MINIMAL, FOCAL
  TESTES: HYPERPLASIA, INTERSTITIAL CELL, SLIGHT, DIFFUSE
  KIDNEY: NEPHROPATHY, MINIMAL, FOCAL

- Other Examples:
  - Inconsistent use of case
  - Inappropriate use of comments, such as inclusion of findings in comments

Inconsistent Use of Clinical Pathology Units and Abbreviations

Background: For the LB (Laboratory Test Results) SEND domain, the lab test or examination name (LBTEST) and units of the original test result (LBORRESU) variables must be mapped to controlled terminology. Accordingly, it is critical that data capture involve unique names and units that can be mapped to these variables and to controlled terminology.

Examples of Actual Data that Present Mapping Challenges for Automation:

Clinical Pathology source data, same study, same abbreviation for different lab tests:

Group/ SEX	Animal Number	Total WBC 1000/cm	N 100/cmm	L 100/cmm	M 100/cmm	E 100/cmm	B 100/cmm	LUC 100/cmm
Group/ SEX	**Animal Number**	**Total WBC 1000/cm**	**N %WBC**	**L %WBC**	**M %WBC**	**E %WBC**	**B %WBC**	**LUC %WBC**

How to Map Biomarkers

Background: According to the Biomarkers Definitions Working Group,[1] a biomarker is defined as a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, or biological responses to a therapeutic intervention. Biomarkers are highly varied and can fall into different categories, including anatomic, physiologic, biochemical, and genomic indicators. They be measured or imaged in different ways.

Proposed Solution: One approach to mapping biomarkers in SEND, under discussion by the FDA / PhUSE Working Group (Investigating Endpoint Modeling – Biomarkers, Antidrug Antibodies, and Immunophenotyping, M. Wasko, Co-Chair) supports presentation of biomarker data in the domain with which it is associated. This approach is consistent with the clinical CDISC teams. For example, anti-drug antibodies would be included in the LB (Laboratory Test Results) domain. QT intervals, if used as a biomarker, would be included in the EG (ECG Test Results) domain.

(1) Biomarkers Definitions Working Group, 2001, Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 69:89-95.

Conclusions

Based on our observations mapping SEND datasets and subsequent FDA submissions, we can make several conclusions:

- The industry is in a transition period with respect to SEND preparation and optimizing LIMS for SEND translation.
- Organized data recording fosters automated SEND translation.
- Consistent data collection processes across end users and up-to-date LIMS optimized for SEND are critical aspects of SEND readiness.
- Macroscopic and microscopic pathology can be viewed as hierarchical descriptive data systems that have the potential for accurate SEND translation, which will provide access to data in ways that are not currently feasible.
- Data collection systems need to exploit the underlying orderly and hierarchical nature of pathology to facilitate automated SEND mapping and avoid manual file manipulation.