Linking chemical and non-chemical data in Structured Product Labeling

Yulia Borodina, Phong Do, William Hess, Yiying Tsai
FDA/OC/OCS/Office of Health Informatics
Structured Product Labeling (SPL)

Structured Product Labeling is a Health Level Seven document markup standard accredited by the American National Standards Institute (ANSI) and adopted by the Food and Drug Administration (FDA) as a mechanism for exchanging product information electronically.
SPL Background

**Code of Federal Regulations:** On December 11, 2003, FDA published final regulations (the electronic labeling rule) requiring the submission of the content of labeling in electronic format for marketing applications (68 FR 69009).

**Food and Drug Administration Amendments Act of 2007 (Public Law 110-85) (FDAAA):** To facilitate the submission of drug establishment registration and drug listing information (including labeling as specified under 21 CFR 207.25), FDA adopted the use of Extensible Markup Language (XML) files in the Structured Product Labeling (SPL) format.

**FDA guidance 2008:** FDA's Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) will index the content of labeling for human drug and biological products using SPL.
SPL Indexing

- SPL indexing refers to the creation of one or more files that include machine-readable annotations to information that are linked to the original SPL file submitted by the manufacturer or distributor for their product.

- FDA creates and publishes (or intends to publish) SPL Index Files for:
  - Substances
  - Pharmacologic Class
  - Drug Interactions
  - Billing Units
  - Adverse Reactions
  - Indications
  - Product Concept
An SPL document is an XML document with inherent machine semantics capacity
Human Semantics vs. Machine Semantics

Concept definition: A vessel with a narrow neck designed to accept a specific closure.

Reference to terminology: NCI Thesaurus
Terminologies Referenced in SPL Documents

<table>
<thead>
<tr>
<th>Code System</th>
<th>Code System OID</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Cancer Institute Thesaurus</td>
<td>2.16.840.1.113883.3.26.1.1</td>
</tr>
<tr>
<td>Logical Observation Identifiers Names and Codes</td>
<td>2.16.840.1.113883.6.1</td>
</tr>
<tr>
<td>Food and Drug Administration Drug Registration and Listing System</td>
<td>2.16.840.1.113883.6.69</td>
</tr>
<tr>
<td>Food and Drug Administration Product Classification System</td>
<td>2.16.840.1.113883.6.303</td>
</tr>
<tr>
<td><strong>Food and Drug Administration Substance Registration System</strong></td>
<td><strong>2.16.840.1.113883.4.9</strong></td>
</tr>
<tr>
<td>Department of Veterans Affairs National Drug File Reference Terminology</td>
<td>2.16.840.1.113883.3.26.1.5</td>
</tr>
<tr>
<td>Systematized Nomenclature of Medicine</td>
<td>2.16.840.1.113883.6.96</td>
</tr>
<tr>
<td>Dun and Bradstreet D-U-N-S Number</td>
<td>1.3.6.1.4.1.519.1</td>
</tr>
<tr>
<td>Code of Federal Regulations</td>
<td>2.16.840.1.113883.3.149</td>
</tr>
<tr>
<td>FDA Submission Tracking System</td>
<td>2.16.840.1.113883.3.150</td>
</tr>
<tr>
<td>FDA Regulatory Compliance Service</td>
<td>2.16.840.1.113883.4.82</td>
</tr>
<tr>
<td>International Society of Blood Transfusion</td>
<td>2.16.840.1.113883.6.18</td>
</tr>
<tr>
<td>Biological Drug Substance Code</td>
<td>2.16.840.1.113883.3.6277</td>
</tr>
</tbody>
</table>
SPL Data Structure

- Data are structured as a set of XML documents of different types
- Each document has an **SPL header:**
  - Schema location
  - FDA stylesheet location
  - Document creation date
  - Author of document
  - Type of SPL document is identified by its *title* and *code*
  - **Document ID** – GUID for each version of SPL document
  - **setID** – GUID for series of versions of related SPL documents
  - **versionNumber**

- Each type of SPL documents has a specific **SPL Body:**
  - Body consists of sections and subsections
  - Each section has a unique ID, effective time and LOINC code
  - SPL elements specific for different types of SPL documents are described in the SPL guide:
SPL Product Resources

- Available

- Under development
SPL Product Data Exchange

Drug manufacturers

Product SPL file

Drug distributors

Product SPL file

SPL Substance index file
SPL Pharm Class index file
SPL Billing Unit index file
SPL Product Concept index file

Labels.fda.gov

DAILYMED

openFDA
Product SPL Files
Product Labels

- "INGREDIENTS AND APPEARANCE" section
  - Coded data elements
    - NDC
    - Active ingredients
    - Inactive ingredients
    - Dosage Form
    - Color, shape, size, imprint code
    - Route of administration
    - Packaging
    - Marketing information
    - ...
- DESCRIPTION SECTION
- CLINICAL PHARMACOLOGY SECTION
- PHARMACOKINETICS SECTION
- INDICATIONS & USAGE SECTION
- CONTRAINDICATIONS SECTION
- ...

12
# Product Labels

## TRIMETHOPRIM
trimethoprim tablet

### Product Information

<table>
<thead>
<tr>
<th>Product Type</th>
<th>HUMAN PRESCRIPTION DRUG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route of Administration</td>
<td>ORAL</td>
</tr>
<tr>
<td>Item Code (Source)</td>
<td>NDC:0093-2158</td>
</tr>
</tbody>
</table>

### DEA Schedule

<table>
<thead>
<tr>
<th>Color</th>
<th>Score</th>
<th>Size</th>
<th>Imprint Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHITE</td>
<td></td>
<td>10mm</td>
<td>9:3:2158</td>
</tr>
<tr>
<td>ROUND</td>
<td>2 pieces</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Active Ingredient/Active Moiety

- TRIMETHOPRIM (UNII: AN164J8Y0X) (TRIMETHOPRIM - UNII:AN164J8Y0X)

### Inactive Ingredients

- SILICON DIOXIDE (UNII: ETJ7Z6XBU4)
- CALCIUM PHOSPHATE, DIBASIC, DIHYDRATE (UNII: O7TSZ97GEP)
- MAGNESIUM STEARATE (UNII: 70097M6I30)
- CELLULOSE, MICROCRYSTALLINE (UNII: OP1R32D61U)
- STARCH, CORN (UNII: 08232NY3SJ)
- SODIUM STARCH GLYCOLATE TYPE A POTATO (UNII: 5856J3G2A2)

### Marketing Information

- Marketing Category: NDA
- Marketing Start Date: 09/30/1990
# Product Labels

## Stylesheet display:

<table>
<thead>
<tr>
<th>Product Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>WHITE</td>
</tr>
<tr>
<td>Shape</td>
<td>ROUND</td>
</tr>
<tr>
<td>Flavor</td>
<td></td>
</tr>
</tbody>
</table>

## Machine-readable representation:

```xml
<characteristic classCode="OBS" xmlns="urn:hl7-org:v3"><code code="SPLCOLOR" codeSystem="2.16.840.1.113883.1.11.19255" /><value xsi:type="CE" code="C48325" codeSystem="2.16.840.1.113883.3.26.1.1" displayName="WHITE" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" /></characteristic>
```
Product Labels

TRIMETHOPRIM TABLETS, USP
2158
2159
Rx only

To reduce the development of drug-resistant bacteria and maintain the effectiveness of trimethoprim tablets, USP and other antibacterial drugs, trimethoprim tablets, USP should be used only to treat or prevent infections that are proven or strongly suspected to be caused by bacteria.

DESCRIPTION
Trimethoprim is a synthetic antibacterial available in tablet form for oral administration. Each scored white tablet contains 100 mg trimethoprim or 200 mg trimethoprim.

Trimethoprim is 5-[(3,4,5-trimethoxyphenyl)methyl]-2,4-pyrimidinediamine. It is a white to light yellow, odorless, bitter compound with a molecular weight of 290.32 and the molecular formula C_{14}H_{13}N_{4}O_{3}. The structural formula is:

![Structural formula of trimethoprim]

CLINICAL PHARMACOLOGY
Trimethoprim is rapidly absorbed following oral administration. It exists in the blood as unbound, protein-bound, and metabolized forms. Ten to twenty percent of trimethoprim is metabolized, primarily in the liver; the remainder is excreted unchanged in the urine. The principal metabolites of trimethoprim are the 1- and 3-oxides and the 3'- and 4'-hydroxy derivatives. The free form is considered to be the therapeutically active form. Approximately 44% of trimethoprim is bound to plasma proteins.

Mean peak serum concentrations of approximately 1.0 mcg/mL occur 1 to 4 hours after oral administration of a single 100 mg dose. A single 200 mg dose will result in serum levels approximately twice as high. The half-life of trimethoprim ranges from 8 to 10 hours. However, patients with severely impaired renal function exhibit an increase in the half-life of trimethoprim, which requires either dosage regimen adjustment or not using the drug in such patients (see DOSAGE AND ADMINISTRATION). During a 13 week study of trimethoprim administered at a daily dosage of 200 mg (50 mg q.i.d.), the mean minimum steady-state concentration of the drug was 1.1 mcg/mL. Steady-state concentrations were achieved within 2 to 3 days of chronic administration and were maintained throughout the experimental period.

Excretion of trimethoprim is primarily by the kidneys through glomerular filtration and tubular secretion. Urine concentrations of trimethoprim are considerably higher than are the concentrations in the blood. A single oral dose of 100 mg, urine concentrations of trimethoprim ranged from 30 to 160 mcg/mL during the 0 to 4 hour period and declined to approximately 18 to 91 mcg/mL during the 8 to 24 hour period. A 200 mg single oral dose will result in trimethoprim urine levels approximately twice as high. After oral administration, 50% to 60% of trimethoprim is excreted in the urine within 24 hours, approximately 80% of this being unmetabolized trimethoprim.

Since normal vaginal and fecal flora are the source of most pathogens causing urinary tract infections, it is relevant to consider the distribution of trimethoprim into these sites. Concentrations of trimethoprim in vaginal secretions are consistently greater than those found simultaneously in the serum, being typically 1.6 times the concentrations of simultaneously obtained serum samples. Sufficient trimethoprim is excreted in the feces to markedly reduce or eliminate trimethoprim-susceptible organisms from the fecal flora.

Trimethoprim also passes the placental barrier and is excreted in human milk.

Microbiology
Trimethoprim blocks the production of tetrahydrofolic acid from dihydrofolic acid by binding to and reversibly inhibiting the required enzyme, dihydrofolate reductase. This binding is much stronger for the bacterial enzyme than for the corresponding mammalian enzyme. Thus, trimethoprim selectively interferes with bacterial biosynthesis of nucleic acids and proteins.

In vitro serial dilution tests have shown that the spectrum of antibacterial activity of trimethoprim includes the common urinary tract pathogens with the exception of Pseudomonas aeruginosa.

The dominant non-Enterobacteriaceae fecal organisms, Bacteroides spp. and Lactobacillus spp., are not susceptible to trimethoprim concentrations obtained with the recommended dosage.

Trimethoprim has been shown to be active against most strains of the following microorganisms, both in vitro and in clinical infections as described in the INDICATIONS AND USAGE section.
Trimethoprim is rapidly absorbed following oral administration. It exists in the blood as unbound, protein-bound, and metabolized forms. Ten to twenty percent of trimethoprim is metabolized, primarily in the liver; the remainder is excreted unchanged in the urine. The principal metabolites of trimethoprim are the 1- and 3-oxides and the 3'- and 4'-hydroxy derivatives. The free form is considered to be the therapeutically active form. Approximately 44% of trimethoprim is bound to plasma proteins.
SPL Substance Index Files
SPL Substance Index Files

Dictionary of substance definitions created and maintained by the FDA Substance Registration system
SPL Substance Index Files

- Each file has a UNII, defining characteristics, and a hash code computed based on the defining characteristics.

- Defining characteristics vary depending on the type of the substance:
  - Small molecules
  - Polymers
  - Biopolymers
  - Plant parts
  - Tissue parts
  - Vaccines, etc.

- Published in stages
  - Published for 50,000 chemical substances
  - Published for 3,000 biological substances (mostly plants)
§ Chemical substance
  § UNII: AN164J8Y0X

§ Chemical structure (MOLFILE)
  § InChI=1S/C14H18N4O3/c1-19-10-5-8(6-11(20-2)12(10)21-3)4-9-7-17-14(16)18-13(9)15/h5-7H,4H2,1-3H3,(H4,15,16,17,18)
  § IEDVJHCEMCRBQM-UHFFFAOYSA-N

§ Biological substance (plant)
  § UNII: 1KE45XD28S
  § Bibliographic reference: Cichorium intybus L.
Machine-readable representation:

```xml
<identifiedSubstance xmlns="urn:hl7-org:v3">
  <code code="AN164J8Y0X" codeSystem="2.16.840.1.113883.4.9" />
  <asEquivalentSubstance>
    <definingSubstance>
      <code code="829975fe-f6a4-2fcf-c445-9167b2df44ee" codeSystem="2.16.840.1.113883.3.2705" /></definingSubstance>
  </asEquivalentSubstance>
  <moiety>
    <quantity>
      <numerator value="1" unit="1" />
      <denominator value="1" unit="1" /></quantity>
    <partMoiety>
      <code code="AN164J8Y0X" codeSystem="2.16.840.1.113883.4.9" /></partMoiety>
  </moiety>
  <subjectOf>
    <characteristic>
      <code displayName="Chemical Structure" codeSystem="2.16.840.1.113883.3.26.1.1" code="C103240" />
      <value xmlns:p4="http://www.w3.org/2001/XMLSchema-instance" p4:type="ED" mediaType="application/x-mdl-molfile">![CDATA[...]]</value>
    </characteristic>
  </subjectOf>
  <subjectOf>
    <characteristic>
      <code displayName="Chemical Structure" codeSystem="2.16.840.1.113883.3.26.1.1" code="C103240" />
      <value xmlns:p4="http://www.w3.org/2001/XMLSchema-instance" p4:type="ED" mediaType="application/x-inchi">InChI=1S/C14H18N4O3/c1-19-10-5-8(6-11(20-2)12(10)21-3)4-9-7-17-14(16)18-13(9)15/h5-7H,4H2,1-3H3,(H4,15,16,17,18)</value>
    </characteristic>
  </subjectOf>
  <subjectOf>
    <characteristic>
      <code displayName="Chemical Structure" codeSystem="2.16.840.1.113883.3.26.1.1" code="C103240" />
      <value xmlns:p4="http://www.w3.org/2001/XMLSchema-instance" p4:type="ED" mediaType="application/x-inchi-key">IEDVJHECMCRBQM-UHFFFAOYSA-N</value>
    </characteristic>
  </subjectOf>
</identifiedSubstance>
```
SPL Pharm  Class Index Files

Pharmacologic class
SPL Pharm Class Index Files

- Pharm Class: a group of drugs that share scientifically documented properties
- Based upon any individual, or combination of 3, attribute(s) of the active moiety
  - Mechanism of Action (MoA)
  - Physiologic Effect (PE)
  - Chemical Structure (CS)
- FDA’s Established Pharmacologic Class (EPC)
  - Scientifically valid and clinically meaningful
  - Related to therapeutic effect(s) of approved indication(s)
SPL Pharm Class Index Files

• Source of standardized indexing concepts and codes: VA’s National Drug File Reference Terminology (NDF-RT)

• Non-therapeutic effects Pharm Class
  – May not be related to approved indications
  – Important for safe and effective use
    • Drug interactions and safety assessments
    • Example: CYP450 induction/inhibition
SPL Pharm Class Index Files

Stylesheet display:

<table>
<thead>
<tr>
<th>Pharmacologic Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dihydrofolate Reductase Inhibitor Antibacterial [EPC](dihydrofolate reductase inhibitor antibacterial), Dihydrofolate Reductase Inhibitors [MoA], Cytochrome P450 2C8 Inhibitors [MoA], and Organic Cation Transporter 2 Inhibitors [MoA]</td>
</tr>
</tbody>
</table>

Machine-readable representation:

```xml
<asSpecializedKind xmlns="urn:hl7-org:v3">
  <generalizedMaterialKind>
    <code code="N0000000191" codeSystem="2.16.840.1.113883.3.26.1.5" displayName="Dihydrofolate Reductase Inhibitors [MoA]" />
  </generalizedMaterialKind>
</asSpecializedKind>
```
SPL Product Concept Index Files
SPL Product Concept Index Files

• Product Concept Index File includes the product concepts found in a representative reference SPL and the drug applications associated with the product concepts

• Product concepts characterized by:
  – Dosage form
  – Active Ingredient (AI) / Active Moiety (AM)
    • Basis of Strength
  – Strength
SPL Product Concept Index Files

<manufacturedProduct xmlns="urn:hl7-org:v3">
  <code code="0242caa0-bcb1-b3bf-f430-5f9a7958a829" codeSystem="2.16.840.1.113883.3.3389" />
  <formCode code="C42998" displayName="TABLET" codeSystem="2.16.840.1.113883.3.26.1.1' />
  <ingredient classCode="ACTIB">
    <quantity>
      <numerator value="100" unit="mg" />
      <denominator value="1" unit="1" />
    </quantity>
    <ingredientSubstance>
      <code code="AN164J8Y0X" codeSystem="2.16.840.1.113883.4.9" />
      <name>TRIMETHOPRIM</name>
      <activeMoiety>
        <code code="AN164J8Y0X" codeSystem="2.16.840.1.113883.4.9" />
        <name>TRIMETHOPRIM</name>
      </activeMoiety>
    </ingredientSubstance>
  </ingredient>
</manufacturedProduct>
Linking Data within SPL and out of SPL

- **Versioning**
  - SPL files are versioned
  - Versions are archived
  - Versions are linked via setId

- **Linking index files to product SPLs**
  - Product SPLs reference Substance index files via UNIIIs
  - Product Concept index files reference products via application numbers
  - Pharm Class index files reference products via UNIIIs

- **Linking between index files**
  - Pharm Class index files reference Substance index files via UNIIIs
  - Product Concept index files reference Substance index files via UNIIIs

- **Linking to SPL data from outside**
  - Linking is supported by the use of identifiers and concept codes:
    - UNII
    - Product Concept ID
    - Abstract Product Concept ID
    - PharmClass terminology code
    - NDC, NDA, ANDA, BLA #
    - InChI
    - LOINC
Linking to Substances from Product SPL

Active Ingredient/Active Moiety

TRIMETHOPRIM (UNII: AN164J8Y0X) (TRIMETHOPRIM - UNII:AN164J8Y0X)

Inactive Ingredients

SILICON DIOXIDE (UNII: ETJ7Z6XBU4)
CALCIUM PHOSPHATE, DIBASIC, DIHYDRATE (UNII: O7TSZ97GEP)
MAGNESIUM STEARATE (UNII: 70097M6I30)
CELLULOSE, MICROCRYSTALLINE (UNII: OP1R32D61U)
STARCH, CORN (UNII: 08232NY3SJ)
SODIUM STARCH GLYCOLATE TYPE A POTATO (UNII: 58563J2A2)

UNII: AN164J8Y0X
UNII: ETJ7Z6XBU4
UNII: O7TSZ97GEP
UNII: 70097M6I30
UNII: OP1R32D61U
Linking to Pharm Classes from Substances

Active moiety UNII: AN164J8Y0X

<table>
<thead>
<tr>
<th>NDFRT Pharm Class code</th>
<th>Pharm Class description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N0000175489</td>
<td>Dihydrofolate Reductase Inhibitor Antibacterial [EPC]</td>
</tr>
<tr>
<td>N00000000191</td>
<td>Dihydrofolate Reductase Inhibitors [MoA]</td>
</tr>
<tr>
<td>N0000187062</td>
<td>Cytochrome P450 2C8 Inhibitors [MoA]</td>
</tr>
<tr>
<td>N0000187061</td>
<td>Organic Cation Transporter 2 Inhibitors [MoA]</td>
</tr>
</tbody>
</table>
Linking to Product SPLs from Substances

UNII:9J765S329G
## Linking Products to Product Concept

### Active Ingredient(s) | Strength | Dosage Form | Application Number
---|---|---|---
levothyroxine sodium | 0.05 mg levothyroxine sodium anhydrous | TABLET | NDA021210

### Active Ingredient(s) | Strength | Dosage Form | Application Number
---|---|---|---
levothyroxine sodium | 0.05 mg levothyroxine sodium anhydrous | TABLET | NDA021342

### Active Ingredient(s) | Strength | Dosage Form | Applications
---|---|---|---
levothyroxine sodium | 0.05 mg levothyroxine sodium anhydrous | TABLET | NDA021116; NDA021210; NDA021342; NDA021301; NDA021402; ANDA076187
Trimethoprim is a Dihydrofolate Reductase Inhibitor Antibacterial. The mechanism of action of trimethoprim is as a Dihydrofolate Reductase Inhibitor, Cytochrome P450 2C8 Inhibitor, and Organic Cation Transporter 2 Inhibitor.
SPL Resources

- **File download and searching**
  - [http://labels.fda.gov/](http://labels.fda.gov/)
  - [https://open.fda.gov/drug/label/](https://open.fda.gov/drug/label/)

- **Guidance, manuals, and info**
Querying Labels at openFDA

- API query example

- API result example

```json
{
  "meta": {
    "disclaimer": "openFDA is a beta research project and not for clinical use. While we make every effort to ensure that data is accurate, you should assume all results are unvalidated."
  },
  "results": []
}
```
Conclusions

- FDA provides to the public information about products, ingredients, substances that are used as ingredients, pharmacologic action of the substances, etc. in structured electronic format: SPL

- Specific information is covered by specific type of SPL documents

- Data provided by a particular type of SPL documents can be used independently, linked into a network of data, or integrated with other IT systems depending on user needs
Answer to the ultimate question

DATA STRUCTURIZATION

DATA STANDARDIZATION

DATA EXCHANGE MECHANISMS
Thank you