

Repurposing of drug label information to create actionable intelligence with BioWisdom's Metawise

Paul Bradley
12th August, 2011

Background

- * Toxicity is now the major cause of drug failure
- * Strong focus on developing predictive methods
 - * Fail early, avoid progressing toxic candidates
- * Improved prediction requires best view of current and *historic* data
 - * Current emphasis on repurposing legacy data
 - * Economic advantages, avoid costly repetition
- * Harnessing legacy data can be a big problem
 - * Different locations
 - * Lack of structure, i.e. free text
 - * Heterogeneity, lack of terminology standards
- * Impedes alignment and analysis
- * Difficult to answer simple questions
 - * Which drugs elevate ALT level?

Dealing with Unstructured Text

- * Impracticable to manually review and extract data from unstructured sources
 - * Automated methods needed to augment the process
 - * Employ costly curation judiciously
- * Conventional text analytics has limitations
 - * NLP and entity recognition leave us with a classification problem
 - * Tailoring to a specific domain, e.g. life sciences
 - * Dictionary look-ups require large quantities of “synonyms”
 - * Keeping current is a burden
- * Need something more flexible
 - * Tolerate linguistic variation
 - * Easier to update
 - * Facilitate adoption of standards
 - * Harmonise legacy data with current research

BioWisdom's Metawise

- * A terminology management system
 - * Applications in enterprise search, literature searching, document mark up, term translation/harmonisation
 - * More powerful than conventional, static dictionary-based methods
 - * Robust, can deal with the linguistic diversity that exists in free text
- * Template matching/extraction: recognition of important concepts based on term structure
 - * Utilises term *categories*, clusters of replaceable terms, to create *templates*
 - * hepatic necrosis - organ_location:process - pancreatic apoptosis
 - * Identifies terms that are similar to known terms
- * Term translation: convert free text references to any standard, e.g. SEND, SNOMED CT, ICD-10
 - * Exhaustive matching: word substitutions, rearrangements, stemming
 - * Translate/map new terms (e.g. from template matching) to a target ontology
 - * Map between ontologies and standards

Entity Recognition and Term Translation

INDICATIONS

For the treatment of anemia due to lack of iron and low folate as in menorrhagia, pregnancy, puberty, excessive blood loss, and advanced age. Also and vitamin C deficiency occur together, along with a poor B-complex vitamins in chronic and acute illness, as well as periods of extended recovery.

score	source	target
0.96	low folate	Decreased folic acid

CONTRAINDICATIONS

Centratex is contraindicated in patients with a known hypersensitivity to any of its ingredients. Iron compounds are contraindicated in patients with hemosiderosis, hemochromatosis, and hemolytic anemias. Folic acid may obscure signs and symptoms of pernicious anemia and is therefore a contraindication as well.

Warning

Accidental overdose of iron-containing products is a leading cause of fatal poisoning in children under 6. Keep this product out of reach of children. In case of accidental overdose, call a doctor or poison control center immediately. Iron is toxic and has been shown to cause severe reactions in children including death.

Candidate Term

low folate

Synonyms & Metadata

low : syn : decreased
 low : syn : reduced
 folate : syn : folic acid
 folate : syn : vitamin b9


Target Translation

Decreased folic acid

Background and Aims

- * DailyMed: online resource of drug package inserts
- * Metadata around section names, marketing information, etc.
 - * Query by Name, NDC Code, Drug Class
- * Data within each section is unstructured
- * No mark up of important concepts: indications, adverse events, etc.
- * Impossible to ask simple questions
 - * Which drugs treat asthma
 - * Do NSAIDs induce dyspnoea?
- * Repurpose labels to create actionable intelligence
- * Metawise used to mark up key language from SNOMED CT
 - * Drugs, AEs, boxed warnings, lab tests
- * Translation of free text descriptors to SNOMED CT
 - * Creates metadata index of SNOMED terms across the entire database
- * Manual curation of a subset to create *assertional* metadata

DailyMed



Daily Med
Current Medication Information

DailyMed provides high quality information about marketed drugs. Drug labeling on this Web site is the most recent submitted to the Food and Drug Administration (FDA) and currently in use; it may include, for example, strengthened warnings undergoing FDA review or minor editorial changes. These labels have been reformatted to make them easier to read.

At the present time this Web site does not contain a complete listing of labels for approved prescription drugs. Currently this Web site contains 27371 drugs.


Search :
Limits: Drug Name NDC Code Drug Class


About DailyMed

DailyMed provides high quality information about marketed drugs. This information includes FDA labels (package inserts). This Web site provides health information providers and the public with a standard, comprehensive, up-to-date, look-up and download resource of medication content and labeling as found in medication package inserts. The National Library of Medicine (NLM) provides this as a public service and does not accept advertisements.

Drug labeling and other information in the SPL is what has been most recently submitted by drug companies to the Food and Drug Administration (FDA) as drug listing information (See 21 CFR part 207). The drug labeling has been reformatted to make it easier to read but its content has not been altered or verified by FDA or National Library of Medicine. The drug labeling on this Web site may not be the labeling on currently distributed products or identical to the labeling that is approved. Drugs marked "OTC monograph final" or "OTC monograph not final" are not checked for conformance to the monograph. Drugs marked "unapproved" on this Web site have not been reviewed by FDA for safety and efficacy and their labeling has not been approved. For more information about unapproved drugs, [visit Enforcement Activities by FDA](#).

Other information about drugs may also be available. NLM regularly processes data files uploaded from FDA's system and provides and maintains this Web site for the public to use in accessing the information. Additional information about medicines is available on NLM's MedlinePlus Web site <http://www.nlm.nih.gov/medlineplus/medicines.html>.

 [Copyright](#), [Privacy](#), [Accessibility](#)
U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health, Health & Human Services

 We comply with the [HONcode](#) standard for health trustworthy information: [verify here](#).

DailyMed

DailyMed provides high quality information about marketed drugs.
Drug labeling on this Web site is the most recent submitted to the Food and Drug Administration (FDA) and currently in use; it may include, for example, strengthened warnings undergoing FDA review or minor editorial changes. These labels have been reformatted to make them easier to read.

Options

- Home
- E-mail Label Information
- Downloads
- Archives
- Notify of Updates
- Contact Us

Additional Resources

- Report Adverse Event
- SPL Format Previewer for Label Authors
- Product Identification System

At the present time this Web site does not contain a complete listing of labels for approved prescription drugs. Currently this Web site contains 27371 drugs.


Search : **GO**


Limits: Drug Name NDC Code Drug Class


Search results for

NSAID

No Drug Package Labels found.

 [Copyright, Privacy, Accessibility](#)
U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894
[National Institutes of Health, Health & Human Services](#)

 [Copyright, Privacy, Accessibility](#)
U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894
[National Institutes of Health, Health & Human Services](#)

 We comply with the [HONcode](#) standard for health trustworthy information: [verify here](#).

DailyMed

DailyMed provides high quality information about marketed drugs.

Daily Med
Current Medication Information

DailyMed provides high quality information about marketed drugs. Drug labeling on this Web site is the most recent submitted to the Food and Drug Administration (FDA) and currently in use; it may include, for example, strengthened warnings undergoing FDA review or minor editorial changes. These labels have been reformatted to make them easier to read.

At the present time this Web site does not contain a complete listing of labels for approved prescription drugs. Currently this Web site contains [27371](#) drugs.

Search:

Limits: Drug Name NDC Code Drug Class

Search results for
kinase inhibitor

Total Results Found: [Count:13]

- [AFINITOR \(everolimus\) tablet](#)
[Novartis Pharmaceuticals Corporation]
- [GLEEVEC \(imatinib mesylate\) tablet](#)
[Novartis Pharmaceuticals Corporation]
- [GLEEVEC \(imatinib mesylate\) tablet](#)
[Physicians Total Care, Inc.]
- [IRESSA \(gefitinib\) tablet, coated](#)
[AstraZeneca Pharmaceuticals LP]

Options

- Home
- E-mail Label Information
- Downloads
- Archives
- Notify of Updates
- Contact Us

Additional Resources

- Report Adverse Event
- SPL Format Previewer for Label Authors
- Product Identification System

NLM

DailyMed

DailyMed provides high quality information about marketed drugs.

Agitated states.
Patients with a history of drug abuse.
During or within 14 days following the administration of monoamine oxidase inhibitors (hypertensive crises may result).

WARNINGS

Serious Cardiovascular Events

Sudden Death in Patients with Pre-existing Structural Cardiac Abnormalities or Other Serious Heart Problems:

Children and Adolescents

Sudden death has been reported in association with CNS stimulant treatment at usual doses in children and adolescents with structural cardiac abnormalities or other serious heart problems. Although some serious heart problems alone carry an increased risk of sudden death, stimulant products generally should not be used in children or adolescents with known structural cardiac abnormalities, cardiomyopathy, serious heart rhythm abnormalities, or other serious cardiac problems that may place them at increased vulnerability to the sympathomimetic effects of a stimulant drug.

Adults

Sudden deaths, stroke, and myocardial infarction have been reported in adults taking stimulant drugs at usual doses for ADHD. Although the role of stimulants in these adult cases is also unknown, adults have a greater likelihood than children of having serious structural cardiac abnormalities, cardiomyopathy, serious heart rhythm abnormalities, coronary artery disease, or other serious cardiac problems. Adults with such abnormalities should also generally not be treated with stimulant drugs (see **CONTRAINDICATIONS**).

Hypertension and Other Cardiovascular Conditions

Stimulant medications cause a modest increase in average blood pressure (about 2 to 4 mmHg) and average heart rate (about 3 to 6 bpm), and individuals may have larger increases. While the mean changes alone would not be expected to have short-term consequences, all patients should be monitored for larger changes in heart rate and blood pressure. Caution is indicated in treating patients whose underlying medical conditions might be compromised by increases in blood pressure or heart rate, e.g., those with pre-existing hypertension, heart failure, recent myocardial infarction, or ventricular arrhythmia (see **CONTRAINDICATIONS**).

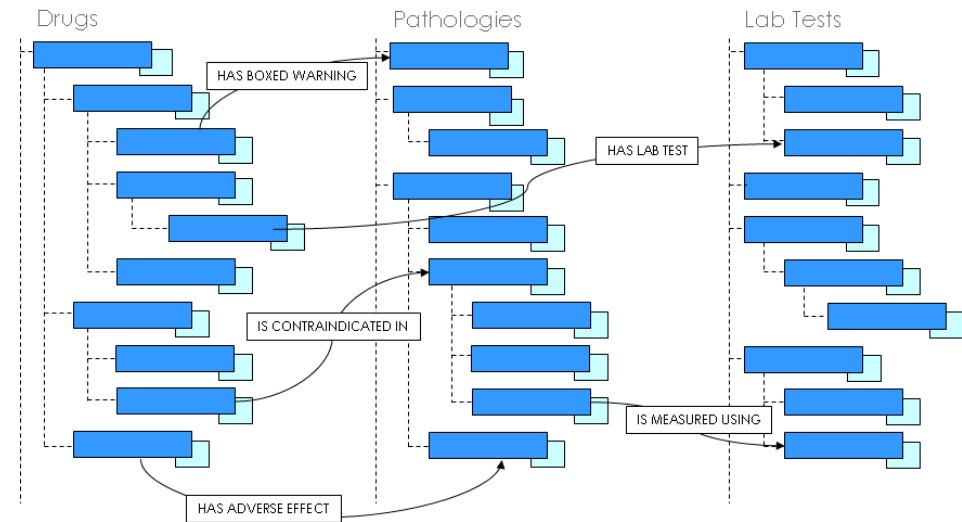
Assessing Cardiovascular Status in Patients Being Treated with Stimulant Medications

Children, adolescents, or adults who are being considered for treatment with stimulant medications should have a careful history (including assessment for a family history of sudden death or ventricular arrhythmia) and physical exam to assess for the presence of cardiac disease, and should receive further cardiac evaluation if findings suggest such disease (e.g., electrocardiogram and echocardiogram). Patients who develop symptoms such as exertional chest pain, unexplained syncope, or other symptoms suggestive of cardiac disease during stimulant treatment should undergo a prompt cardiac evaluation.

Assertional Metadata

- * Combines concepts as subject-relationship-object triplet, e.g.

Amiodarone – PROLONGS – QT Interval



- * Adds extra layer of detail over simple tagging of key terms
- * Creates a navigable information layer over documents
 - * Relationships add context
 - * Is myalgia an indication, adverse effect or boxed warning?
- * Allows more complex questions to be posed
 - * What adverse effects are observed in drugs that treat breast cancer?
 - * Which black box warnings occur in medicines that list haemolytic anaemia as an adverse event?
- * When combined with taxonomies, can ask broader, class related questions
 - * What testing should be carried out if prescribing a *corticosteroid*?
 - * Which *LFTs* should be performed with *statins*?

Mark-up and Curation of Drug Labels

- * Bespoke curation interface
- * Web based, allowing remote users to access concurrently

The screenshot displays the biowisdom curation interface. On the left is a 'Document List' table with columns for document name and 'Accepted / Total' counts. The main area shows a document with text and various mark-ups. A dropdown menu is open, showing 'bmo' and 'snomed' columns with corresponding terms like 'blood pressure' and 'Hypertension'. An 'Assertion Pane' at the bottom right shows a table of assertions with columns for State, Subject, Relation, Observation, and Action.

Document List

Doc	Accepted / Total
oxefiraxone (10)	121 / 472 (26%)
oxefiraxone (14)	121 / 472 (26%)
oxefiraxone (17)	121 / 472 (26%)
clomipramine	267 / 892 (30%)
clomipramine (1)	267 / 866 (30%)
clomipramine (2)	270 / 890 (30%)
clomipramine (3)	268 / 893 (30%)
clomipramine (4)	267 / 866 (30%)
clomipramine (5)	270 / 892 (30%)
clonidine	82 / 427 (19%)
clonidine (2)	82 / 427 (19%)
clonidine (4)	67 / 332 (20%)
dextroamphetamine (7)	65 / 326 (20%)
diclofenac (17)	0 / 18 (0%)
diclofenac, misoprostol	202 / 1024 (20%)
dihydroergotamine	52 / 326 (16%)
dihydroergotamine (1)	46 / 328 (14%)
dihydroergotamine (2)	125 / 582 (22%)
emtricitabine, tenofovir	127 / 893 (18%)
emtricitabine, tenofovir (1)	127 / 893 (18%)
eribovir	33 / 575 (6%)
	112 / 1840
	54897 / 278430

Text Editor with Mark-up:

blood pressure (bmo) Blood pressure (snomed) Blood pressure

increase in average blood pressure (bmo) Hypertension (snomed) Blood pressure

monoamine oxidase inhibitors (hypertensive crises may

Cardiac Abnormalities or Other Serious Heart

with CNS stimulation, treatment at usual doses in children or other serious heart problems. Although some risk of sudden death, blood pressure, stroke, and myocardial infarction have been reported in adults taking dextroamphetamine drugs at usual doses for ADHD. Although the risk of stroke in these adult cases is also unknown, adults have a greater likelihood than children of having serious structural cardiac abnormalities, cardiomyopathy, serious heart rhythm abnormalities, coronary artery disease, or other serious cardiac problems. Adults with such abnormalities should also generally not be treated with dextroamphetamine drugs (see CONTRAINDICATIONS).

Hypertension and Other Cardiovascular Conditions

Similar medications cause a modest increase in average blood pressure (about 2 to 4 mmHg) and average heart rate (about 3 to 6 bpm), and individuals may have larger increases. While the mean changes alone would not be expected to have short-term consequences, all patients should be monitored for larger changes in heart rate and blood pressure. Caution is indicated in treating patients whose underlying

Assertion Pane:

State	Subject	Relation	Observation	Action
[Only accepted]	[Hide rejected]	[Hide accepted]	[Hide flagged]	[Show advanced]
✓	dextroamphetamine	HAS ADVERSE EFFECT	Hypertension (SNOMED: Hypertension)	[sentence] [x] [i]
	dextroamphetamine [reject]	HAS LAB TEST	Blood pressure [reject]	[sentence] [x] [i]
	dextroamphetamine [reject]	HAS ADVERSE EFFECT	ventricular rate [reject]	[sentence] [x] [i]
	dextroamphetamine [reject]	HAS LAB TEST	Heart rate [reject]	[sentence] [x] [i]
✓	dextroamphetamine	HAS ADVERSE EFFECT	cardiac rate change (negative)	[sentence] [x] [i]

Entity Recognition from Drug Labels

SNOMED CT Concept	DailyMed Match
Anaphylactic-type reaction	anaphylactoid reaction, anaphylactic reactions, anaphylactic shock, anaphylactoid and allergic reactions, anaphylactic reaction
Hypertension	elevation of blood pressure, increase in average blood pressure, high blood pressure, increased blood pressure

- * Stemming
 - * Anaphylactic/anaphylactoid, reaction/reactions
- * Dictionary Look-ups
 - * Hypertension -> increased blood pressure
- * Stemming, word insertion
 - * increased blood pressure -> increased **in average** blood pressure

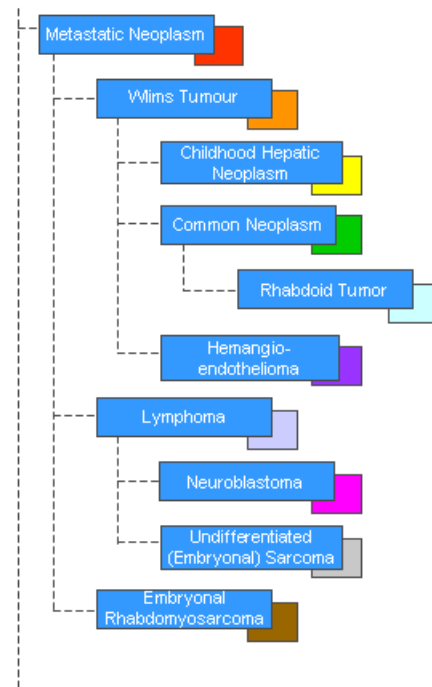
The Benefits of Organised Data

- * Document mark-up and translation creates metadata index of SNOMED terms across the entire database
 - * Enables semantic querying, document triage
 - * Adopt any target ontology structure for taxonomic querying

Target List, e.g. NCI Thesaurus Neoplasms

Document Source, e.g. Medline

Query at most granular level for individual concepts



Hepatic tumors in children.

Although they account for only 1% to 4% of solid tumors in children, hepatic tumors and pseudotumors offer a diagnostic challenge to the clinician seeing only an occasional case. Metastatic lesions such as neuroblastoma, Wilms' tumor, and lymphoma are the most common neoplasm seen in the liver, but 10 distinct primary tumors and pseudotumors of the liver occur with some regularity, and a few others may be seen rarely, including leiomyosarcoma, rhabdoid tumor, and endodermal sinus tumor. Five of these neoplasms--hepatoblastoma, infantile hemangio-endothelioma, mesenchymal hamartoma, undifferentiated embryonal sarcoma, and embryonal rhabdomyosarcoma of the biliary tree--occur only in children and are the major focus of the article.

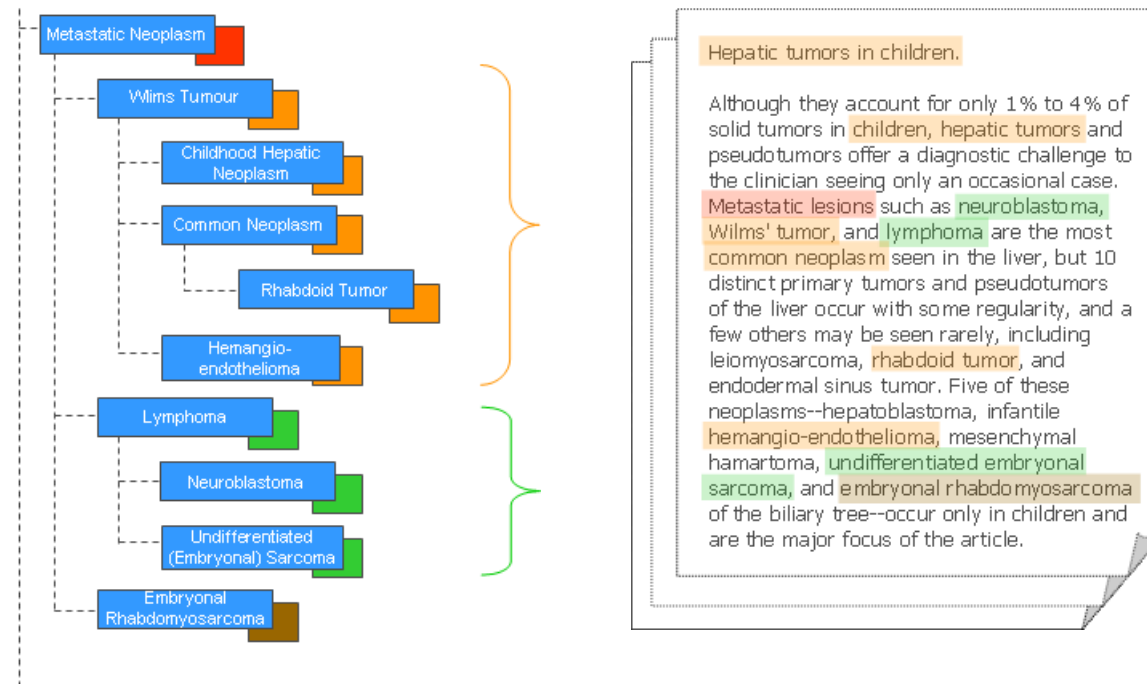
The Benefits of Organised Data

- * Document mark-up and translation creates metadata index of SNOMED terms across the entire database
 - * Enables semantic querying, document triage
 - * Adopt any target ontology structure for taxonomic querying

Target List, e.g. NCI Thesaurus Neoplasms

Document Source, e.g. Medline

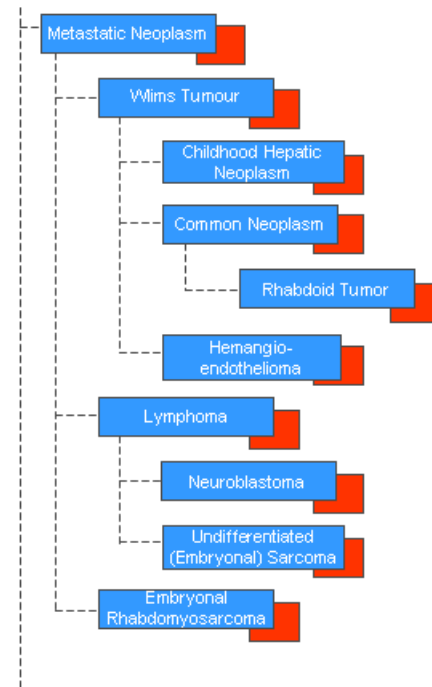
Query at selected parent level for groups of related concepts



The Benefits of Organised Data

- * Document mark-up and translation creates metadata index of SNOMED terms across the entire database
 - * Enables semantic querying, document triage
 - * Adopt any target ontology structure for taxonomic querying

Target List, e.g. NCI Thesaurus Neoplasms



Document Source, e.g. Medline

Hepatic tumors in children.

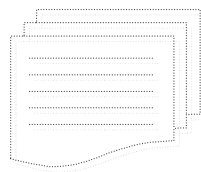
Although they account for only 1% to 4% of solid tumors in children, hepatic tumors and pseudotumors offer a diagnostic challenge to the clinician seeing only an occasional case. Metastatic lesions such as neuroblastoma, Wilms' tumor, and lymphoma are the most common neoplasm seen in the liver, but 10 distinct primary tumors and pseudotumors of the liver occur with some regularity, and a few others may be seen rarely, including leiomyosarcoma, rhabdoid tumor, and endodermal sinus tumor. Five of these neoplasms--hepatoblastoma, infantile hemangio-endothelioma, mesenchymal hamartoma, undifferentiated embryonal sarcoma, and embryonal rhabdomyosarcoma of the biliary tree--occur only in children and are the major focus of the article.

Query at selected top level i.e. Metastatic Neoplasm

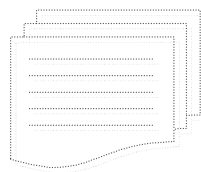
Interoperability with Other Resources

- * Several reports of “over-warning” in drug labelling
 - * Duke *et al*, “Archives of Internal Medicine”, July 2011
- * FDA took action in 2006 to prevent to discourage this act
- * Compare DailyMed boxed warnings with reported effects in AERS
- * Identify similarities / gaps
- * DailyMed and AERS expressed consistently in BioWisdom’s Safety Intelligence Program (SIP)

DailyMed (SNOMED CT)

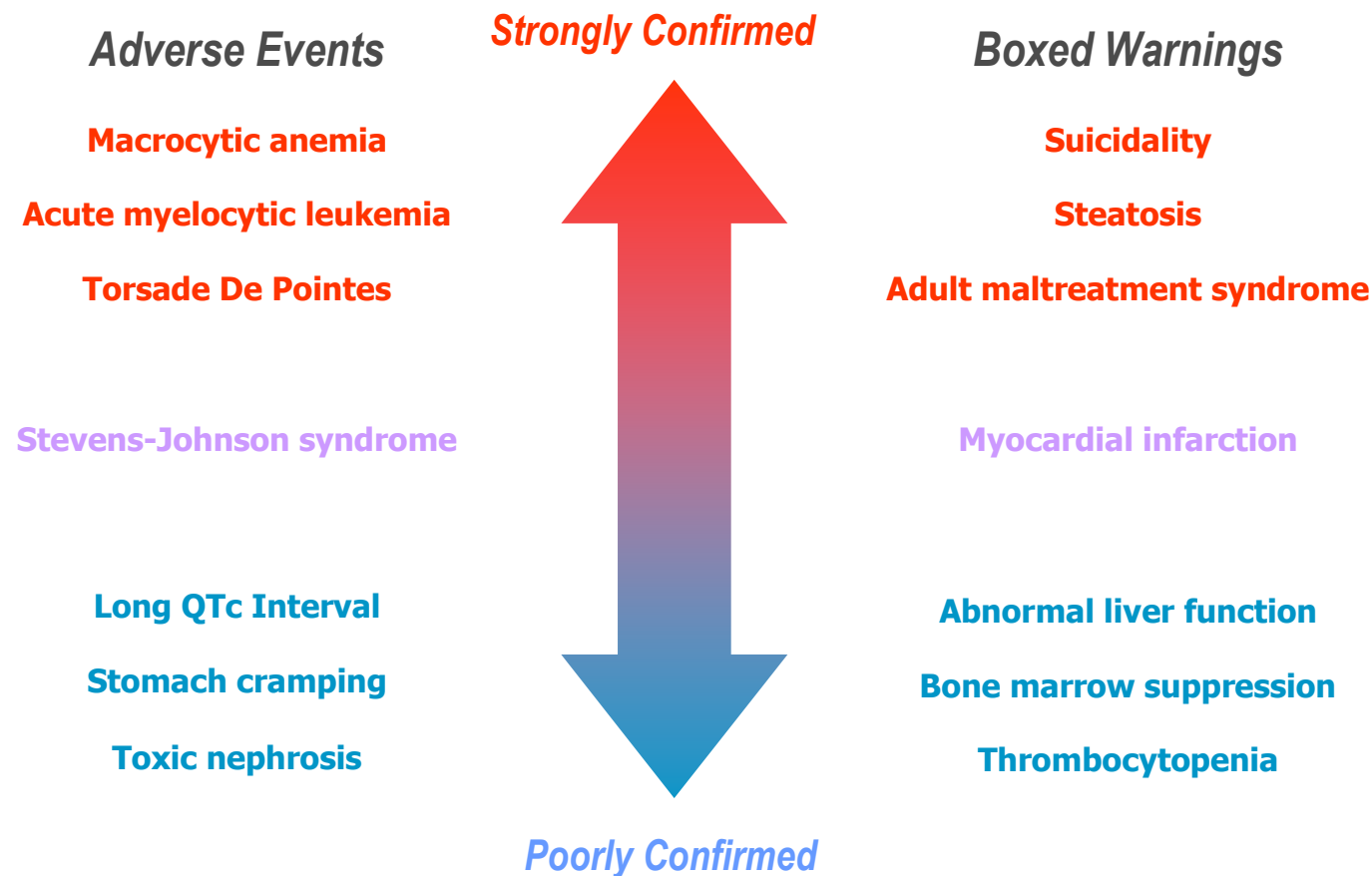


AERS (MedDRA)

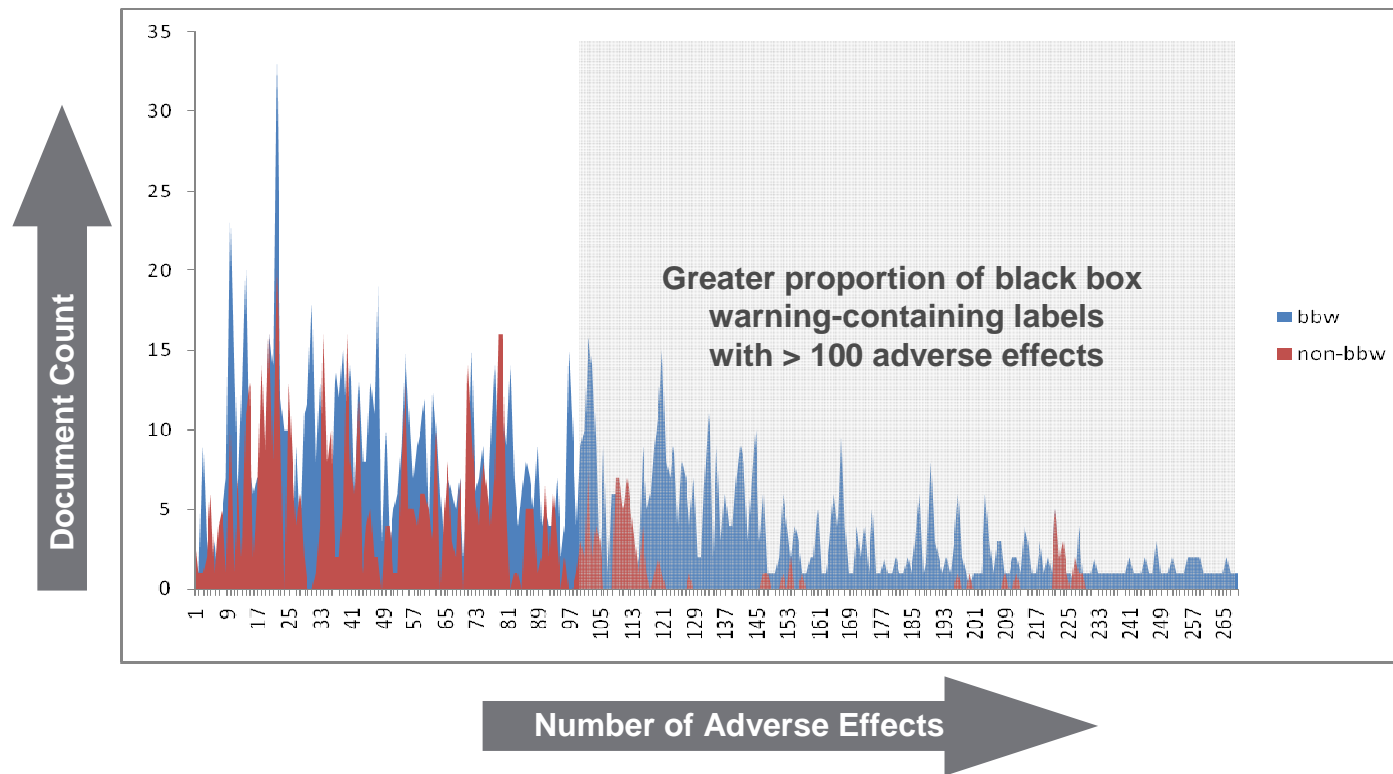


Drug	DailyMed Warning	In AERS?
Bleomycin	Confusion	✓
	Fever	✓
	Hypotension	✓
	Hypothermia	✓
	Pneumonitis	✓
	Pulmonary fibrosis	✓
	Pulmonary toxicity	✓
	Wheezing	✗
Carbamazepine	Agranulocytosis	✓
	Aplastic anemia	✓
	Bone marrow suppression	✓
	Fever	✓
	Leucopenia	✓
	Low platelet count	✗
	Stevens-johnson syndrome	✓
	Toxic epidermal necrolysis	✓

Confirmation of Labelled Adverse Effects/Boxed Warnings by AERS



Adverse Events per Label: Comparison of Black Box and Non-Black Box Drugs



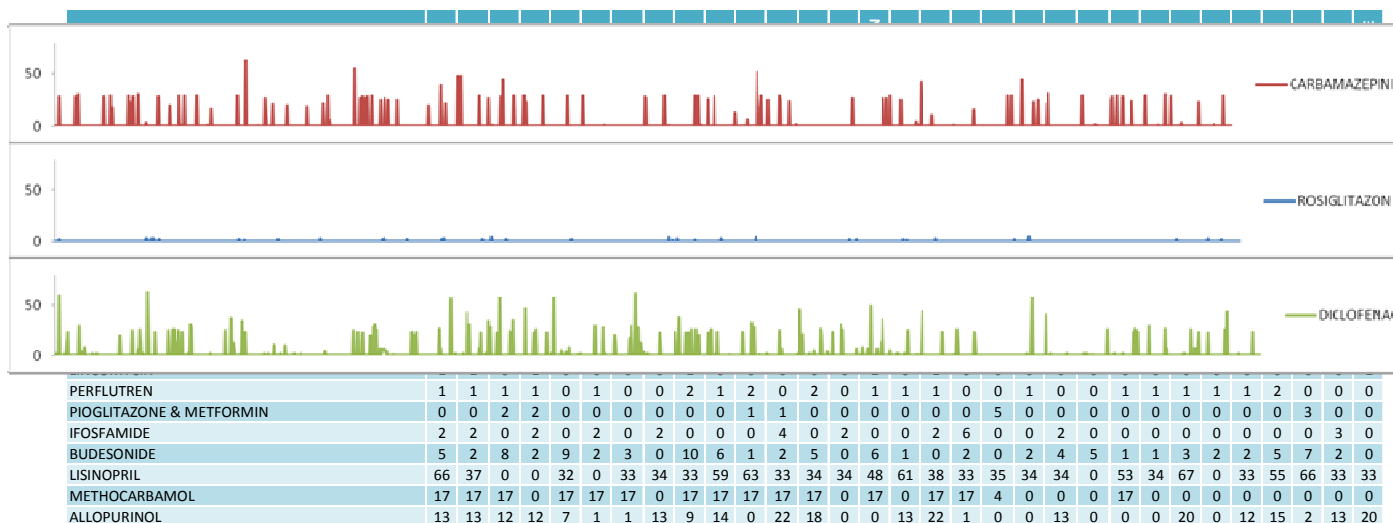
Consistent Language Facilitates Alignment

- * Align features and record frequency across all labels
- * Creates characteristic profile for each label, or each drug
- * Tractable to alignment algorithms
 - * Cluster drugs according to adverse event profile
 - * Could similar profiles be predictive for potential black box warnings in drugs not carrying a warning?
 - * Could some profiles act as early warning signals – how does my drug compare to withdrawn compounds?

Drug Name	NAUSEA	EMESIS	HEADACHE	DIARRHEA	RASH	DIZZINESS	INSOMNIA	THROMBOCYTOPENIA	URTICARIA	FEVER	HYPOTENSION	DROWSINESS	PRURITUS	CONSTIPATION	ANAPHYLACTIC-TYPE REACTION	ABDOMINAL PAIN	LEUCOPENIA	CONFUSION	EDEMA	DYSPNEA	ANOREXIA	ANXIETY	SYNCOPE	XEROSTOMIA	ARTHRALGIA	TACHYCARDIA	PARESTHESIA	DEATH	MYALGIA	ALOPECIA	STEVENS-JOHNSON SYNDROME	
COLISTIMETHATE	0	0	0	0	1	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
AMLODIPINE & TELMISARTAN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0
ERGOTAMINE & CAFFEINE	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	
OXYTOCIN	2	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
PERINDOPRIL	3	3	3	3	3	3	0	3	3	3	3	3	3	3	5	5	0	0	3	3	0	3	3	3	3	3	3	3	3	0	0	
NOREPINEPHRINE	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	
LINCOMYCIN	1	1	0	1	0	0	0	0	1	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
PERFLUTREN	1	1	1	1	0	1	0	0	2	1	2	0	2	0	1	1	1	0	0	1	0	0	1	1	1	1	1	1	2	0	0	
PIOGLITAZONE & METFORMIN	0	0	2	2	0	0	0	0	0	0	1	1	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	3	0	0
IFOSFAMIDE	2	2	0	2	0	2	0	2	0	0	4	0	2	0	2	6	0	2	0	0	2	0	0	0	0	0	0	0	0	0	3	0
BUDESONIDE	5	2	8	2	9	2	3	0	10	6	1	2	5	0	6	1	0	2	0	2	4	5	1	1	3	2	2	5	7	2	0	
LISINAPRIL	66	37	0	0	32	0	33	34	33	59	63	33	34	34	48	61	38	33	35	34	34	0	53	34	67	0	33	55	66	33	33	
METHOCARBAMOL	17	17	17	0	17	17	17	0	17	17	17	17	17	17	0	17	17	17	4	0	0	0	17	0	0	0	0	0	0	0	0	0
ALLOPURINOL	13	13	12	12	7	1	1	13	9	14	0	22	18	0	0	13	22	1	0	0	13	0	0	20	0	12	15	2	13	20		

Consistent Language Facilitates Alignment

- * Align features and record frequency across all labels
- * Creates characteristic profile for each label, or each drug
- * Tractable to alignment algorithms
 - * Cluster drugs according to adverse event profile
 - * Could similar profiles be predictive for potential black box warnings in drugs not carrying a warning?
 - * Could some profiles act as early warning signals – how does my drug compare to withdrawn compounds?

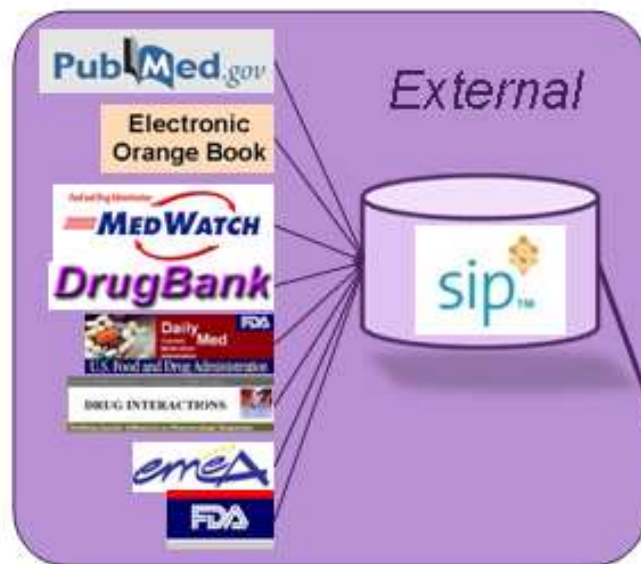


Safety Intelligence: Systems Toxicology



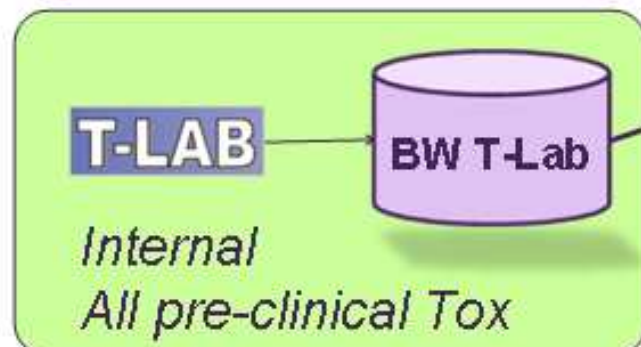
Vision

Augmentation of drug safety decision-making through integration of internal and external safety prior knowledge

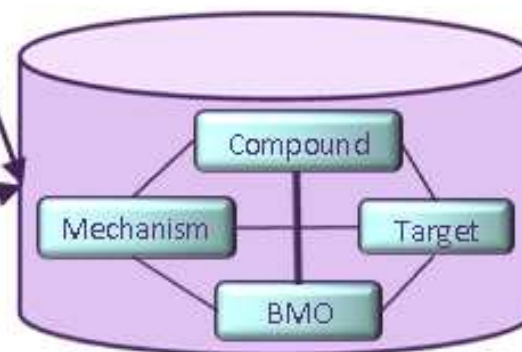


Typical Business Questions Addressed

- What compounds/targets have been tested preclinically and what are the findings?
- What compounds/targets show preclinical tox in 28 day rat studies?
- "Find all compounds causing glycogen accumulation in liver in man and rat" & "What are their common pharmacology and pathways?"



Partnership with **Biowisdom**



Integrated views in PharmaConnect



Conclusions

- * DailyMed is a rich source of drug intelligence
- * Unstructured, suffers all of the drawbacks typical of free text
- * Hinders its use in decision support
 - * Querying, analysis, integration with other applications
- * Concept mark-up and translation with Metawise
- * Curation of assertional metadata
 - * Actionable information layer
- * Metadata index provides powerful substrate for analysis
 - * Cluster drugs or individual labels
 - * Integrate with other data sources
- * Conversion of data to standards is of increasing importance
 - * SEND from C-DISC, interchange between organisations and regulators



Contact:

paul.bradley@biowisdom.com

paul.bradley@instem.com