



# **Drug-drug interaction extraction from Structured Product Labels**

**Johann Stan - [johann.stan@nih.gov](mailto:johann.stan@nih.gov)**

**Dina Demner-Fushman, Kin Wah Fung, Olivier Bodenreider**

**Lister Hill Center, National Library of Medicine, Bethesda, USA**

# Drug-drug Interactions

<http://alternativemedicine4you.com>



- A drug-drug interaction (DDI) is defined as a modification in the effect of a drug when administered with another drug

dofetilide and cimetidine



Grapefruit and grapefruit juice can react adversely with over 85 prescription medications:

- decreases the activity of the cytochrome P450 3A4 (CYP3A4) enzymes

# Food–drug interaction: grapefruit juice augments drug bioavailability—mechanism, extent and relevance (A Dahan and H Altman)



<i>Drug</i>	<i>Grapefruit juice influence</i>	<i>Potential risk</i>	<i>Recommendation</i>
<i>Calcium channel antagonists</i>			
Felodipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Nisoldipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Nicardipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Nitrendipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Pranidipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Nimoldipine	Increased bioavailability	Hypotension, tachicardia	Avoid combination
Nifedipine	No influence		None
Amlodipine	No influence		None
Verapamil	Increased bioavailability	Hypotension, thchicardia	Avoid combination
Diltiazem	No influence		None
<i>CNS modulators</i>			
Diazepam	Increased bioavailability	Increased CNS depression	Avoid combination
Triazolam	Increased bioavailability	Increased CNS depression	Avoid combination
Midazolam	Increased bioavailability	Increased CNS depression	Avoid combination
Alprazolam	No influence		None
Carbamazepine	Increased bioavailability	Increased adverse effects	Avoid combination
Buspirone	Increased bioavailability	Increased adverse effects	Avoid combination
Sertraline	Increased bioavailability	Increased adverse effects	Avoid combination
<i>HMG coA reductase inhibitors</i>			
Simvastatin	Increased bioavailability	Rhabdomyolysis, acute renal failure	Avoid combination
Lovastatin	Increased bioavailability	Rhabdomyolysis, acute renal failure	Avoid combination
Atorvastatin	Increased bioavailability	Rhabdomyolysis, acute renal failure	Avoid combination
Pravastatin	No influence		None
<i>Immunosuppressants</i>			
Cyclosporine	Increased bioavailability	Nephrotoxicity, hypertension, cerebral toxicity	Avoid combination
<i>HIV protease inhibitor</i>			
Saquinavir	Increased bioavailability	Increased adverse effects	Avoid combination
<i>Phosphodiesterase-5 inhibitor</i>			
sildenafil	Increased bioavailability	Increased adverse effects	Avoid combination
<i>Antihistamines</i>			
Terfenadine	Increased unmetabolised drug in plasma	QT prolongation, torsade de pointes	Avoid combination
<i>Prokinetics</i>			
cisapride	Increased bioavailability	QT prolongation, torsade de pointes	Avoid combination
<i>Antiarrhythmics</i>			
Amiodarone	Blockage of Metabolite formation	Arrhythmias	Avoid combination

# Are drug-drug interactions important to detect?

**60 cohort and case-control studies reported an elevated risk of hospitalization in patients who were exposed to DDIs.**

**(Hines & Murphy, 2011)**

**Clinically important events attributable to DDIs 5.3-14.3% of inpatients and are responsible for 0.02 – 0.17% of the nearly 130 million emergency dept. visits each year.**

**(“FASTSTATS” – Emergency Department Visits, Magro et al., 2013)**



The potential for **reducing medication errors** by **using** computerized medical records as well as **drug-interaction screening software** that **detects and alerts the physician** and/or **pharmacist** to potentially serious drug interactions has been recognized.

**(Committee on Quality of Health Care in America: Institute of Medicine. To err is human: building a safer health system. Washington, D.C.: National Academy Press,2000)**

# Public Sources of Drug-Drug Interactions (Boyce et al., 2014)

- 13 publicly available sources
- **5 sources for clinical application**
  - **CredibleMeds** ("Crediblemeds.org," 2013) - a list of DDIs thought to be clinically relevant and be supported by strong scientific evidence
  - **VA-NDF-RT** (Olvey, Clauschee, & Malone, 2010)
  - **ONC High Priority** (Phansalkar et al., 2012)
  - **ONC Non-interruptive** (Phansalkar et al., 2013)
  - **OSCAR** - a list of DDIs derived by expert consensus in the late 1990s (Crowther, Holbrook, Kenwright, & Kenwright, 1997)
- **3 sources to support Natural Language Processing (NLP)**
  - **DDI Corpus 2011** (Segura-Bedmar, Martinez, & Sánchez-Cisneros, 2011)
  - **DDI Corpus 2013** (Segura-Bedmar, Martinez, & Herrero-Zazo, 2013)
  - **PK DDI Corpus** (Boyce, Gardner, & Harkema, 2012)
- **5 other sources** were developed to support either pharmacovigilance or bioinformatics applications
  - **KEGG DDI, TWOSIDES, DrugBank, SemMedDB-SemRep, DIKB**

# Structured Product Labels published by DailyMed (NLM)

**LABEL: SIMVASTATIN- simvastatin tablet, orally disintegrating**

## SAFETY

[Report Adverse Events](#)

[FDA Safety Recalls](#)

[Presence in Breast Milk](#)

## RELATED RESOURCES

[Medline Plus](#)

[Clinical Trials](#)

## MORE INFO FOR THIS DRUG

[Get Label RSS Feed](#)

**NDC Code(s):** 63672-0001-1, 63672-0001-3, 63672-0001-5

**Packager:** Synthon Pharmaceuticals, Inc.

**Category:** HUMAN PRESCRIPTION DRUG LABEL

## DRUG LABEL INFORMATION

DOWNLOAD DRUG LABEL INFO: [PDE](#) | [XML](#) | [PDF](#) | [P](#)

[VIEW ALL SECTIONS](#)

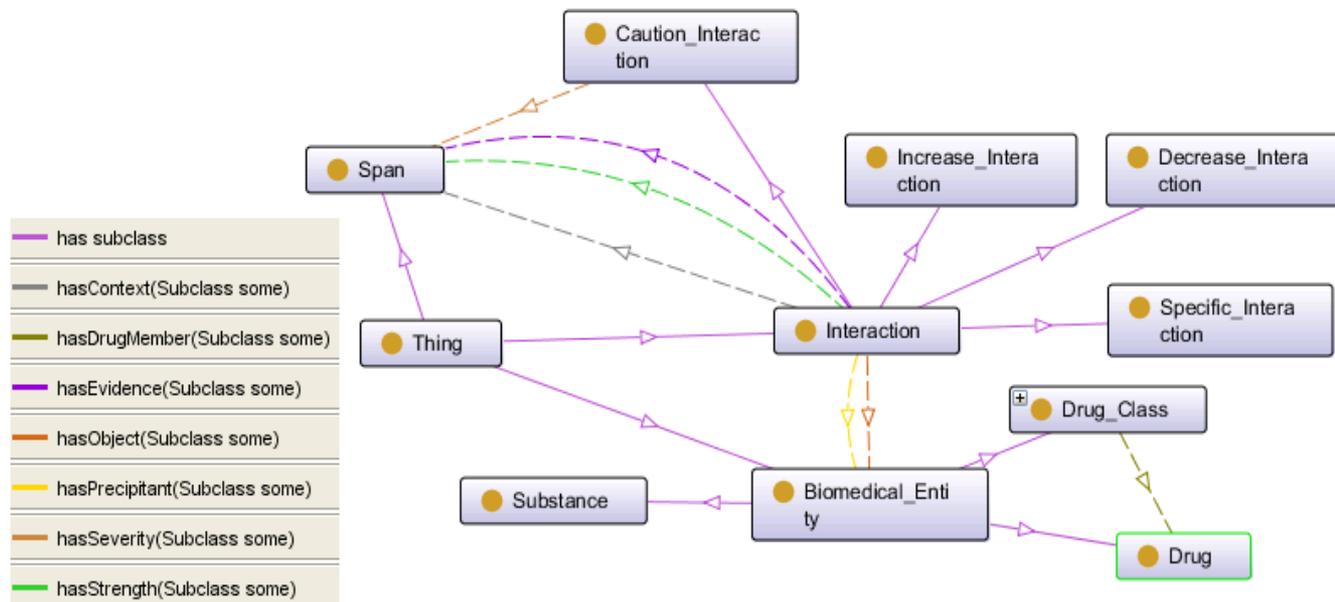
- [+ SPL UNCLASSIFIED SECTION](#)
- [+ DESCRIPTION](#)
- [+ CLINICAL PHARMACOLOGY](#)
- [+ INDICATIONS AND USAGE](#)
- [+ CONTRAINDICATIONS \(WHAT IS THIS?\)](#)
- [+ WARNINGS](#)
- [+ PRECAUTIONS](#)
- [+ ADVERSE REACTIONS](#)
- [+ OVERDOSAGE](#)
- [+ DOSAGE AND ADMINISTRATION](#)
- [+ HOW SUPPLIED](#)
- [+ SPL UNCLASSIFIED SECTION](#)
- [+ INGREDIENTS AND APPEARANCE](#)

[VIEW ALL SECTIONS](#)

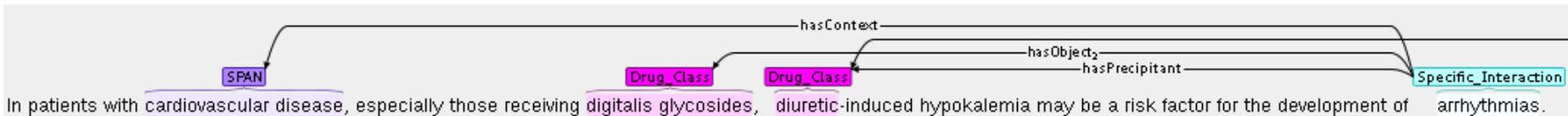
The risk of myopathy/rhabdomyolysis is increased by concomitant use of simvastatin with the following:

- Potent inhibitors of CYP3A4:** Simvastatin, like several other inhibitors of HMG-CoA reductase, is a substrate of cytochrome P450 3A4 (CYP3A4). When simvastatin is used with a potent inhibitor of CYP3A4, elevated plasma levels of HMG-CoA reductase inhibitory activity can increase the risk of myopathy and rhabdomyolysis, particularly with higher doses of simvastatin.
- The use of simvastatin concomitantly with the potent CYP3A4 inhibitors itraconazole, ketoconazole, erythromycin, clarithromycin, telithromycin, HIV protease inhibitors, nefazodone, or large quantities of grapefruit juice (>1 quart daily) should be avoided.** Concomitant use of other medicines labeled as having a potent inhibitory effect on CYP3A4 should be avoided unless the benefits of combined therapy outweigh the increased risk. If treatment with itraconazole, ketoconazole, erythromycin, clarithromycin or telithromycin is unavoidable, therapy with simvastatin should be suspended during the course of treatment.
- Gemfibrozil, particularly with higher doses of simvastatin: The dose of simvastatin should not exceed 10 mg daily in patients receiving concomitant medication with gemfibrozil. The combined use of simvastatin with gemfibrozil should be avoided, unless the benefits are likely to outweigh the increased risks of this drug combination.**
- Other lipid-lowering drugs (other fibrates or  $\geq 1$  g/day of niacin):** Caution should be used when prescribing other fibrates or lipid-lowering doses ( $\geq 1$  g/day) of niacin with simvastatin, as these agents can cause myopathy when given alone. **The benefit of further alterations in lipid levels by the combined use of simvastatin with other fibrates or niacin should be carefully weighed against the potential risks of these combinations.**
- Cyclosporine or danazol, with higher doses of simvastatin: The dose of simvastatin should not exceed 10 mg daily in patients receiving concomitant medication with cyclosporine or danazol.** The benefits of the use of simvastatin in patients receiving cyclosporine or danazol should be carefully weighed against the risks of these combinations.

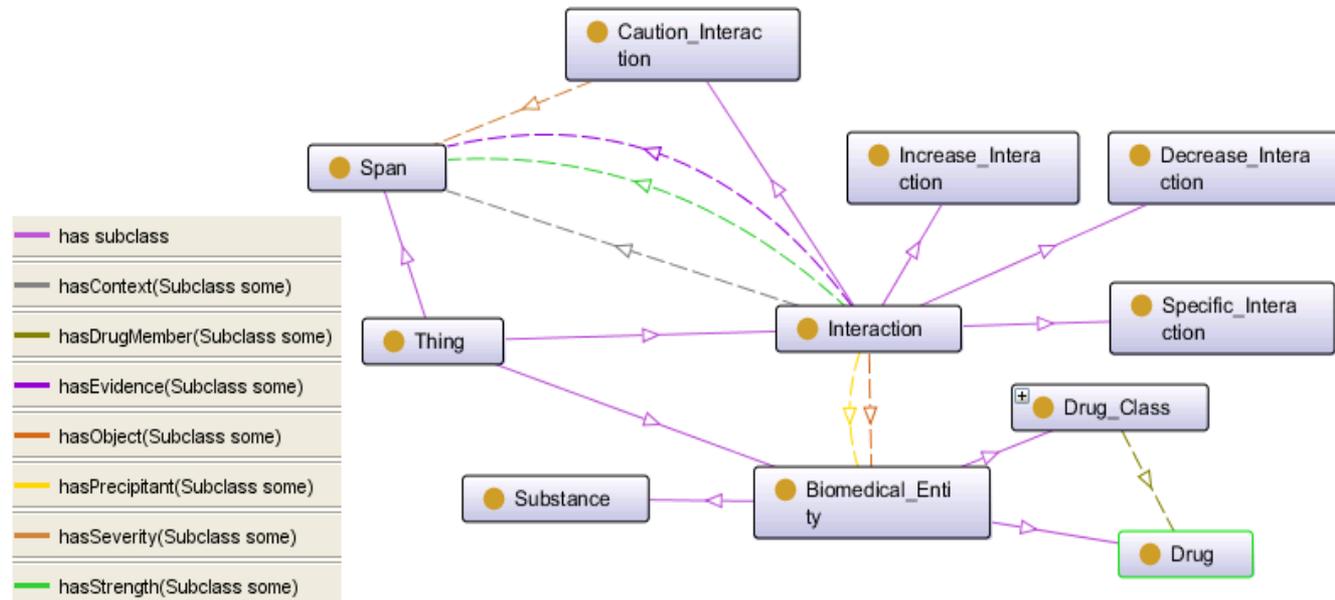
# Annotation Schema (Fragment)



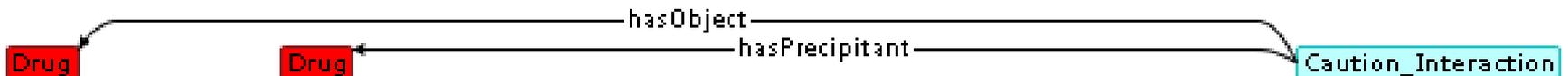
***Specific Interaction*** - specific effects resulting from the interaction



# Novel Annotation Schema (Fragment)

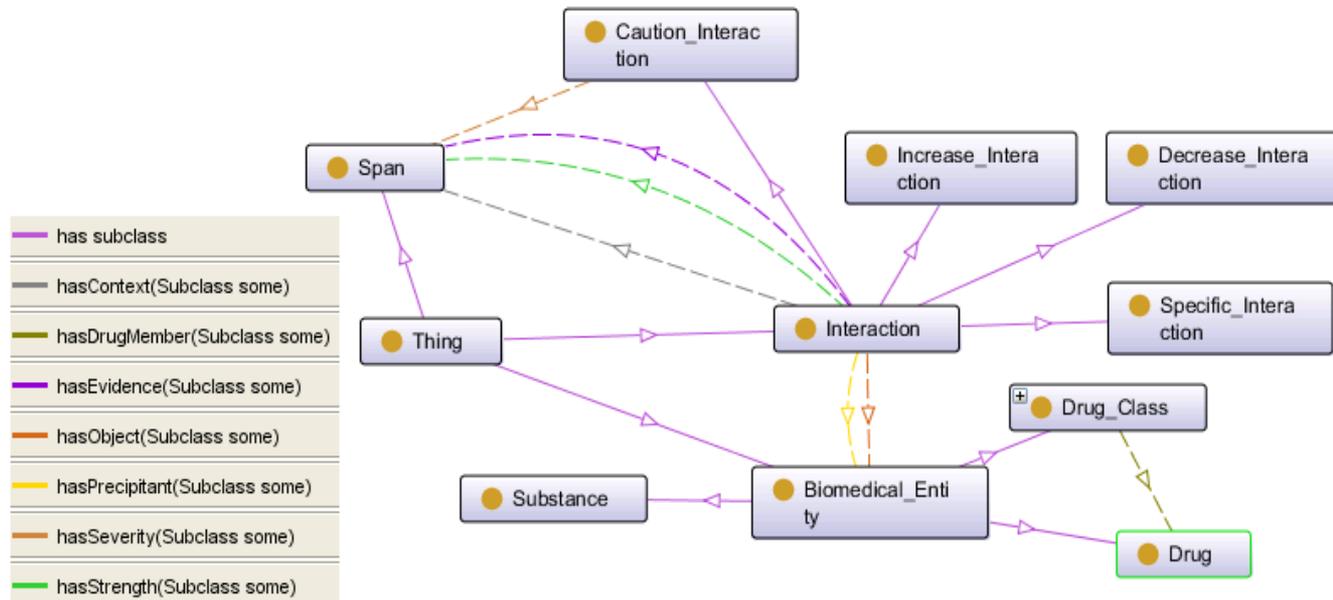


***Caution Interaction*** - precautions about the use of two entities together without specific mention of an effect.

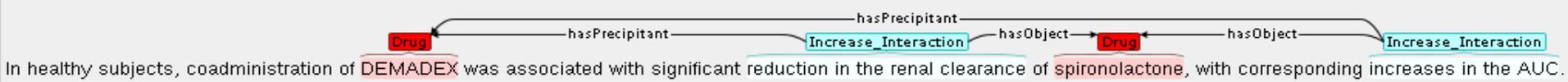


If **DEMADEX** and **cholestyramine** are used concomitantly, simultaneous administration is not recommended.

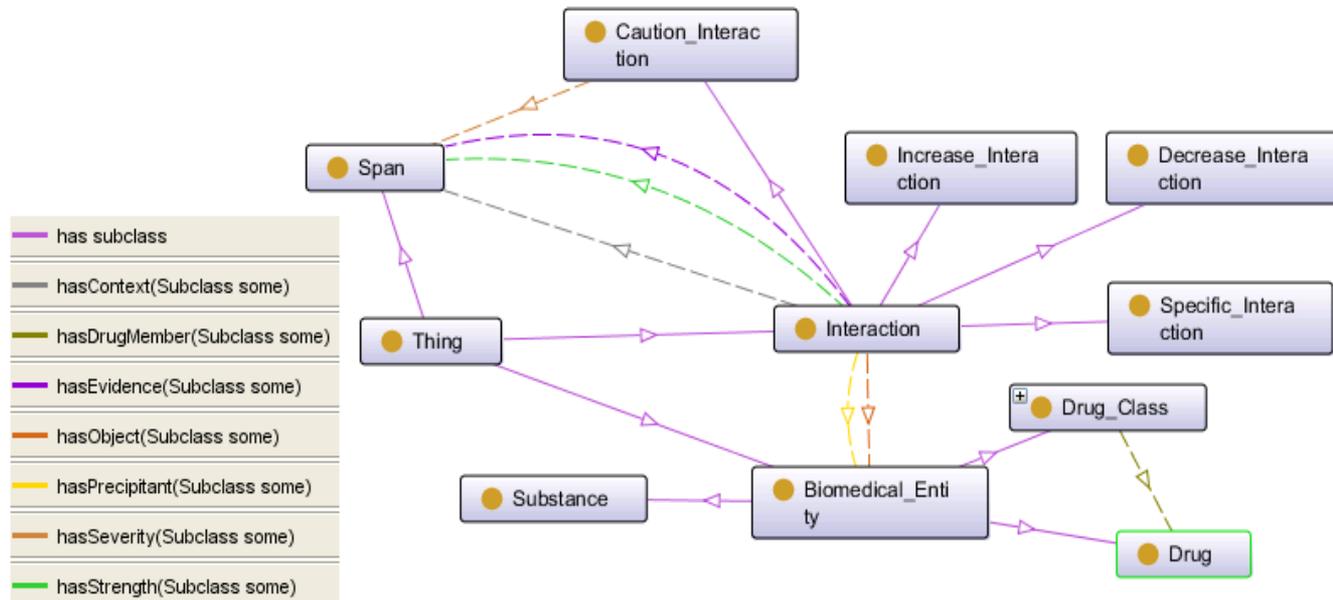
# Novel Annotation Schema (Fragment)



***Increase Interaction*** - indicate higher levels or increased effects of the object drug in the system as a result of the precipitant drug.



# Novel Annotation Schema (Fragment)



***Decrease Interaction*** - lower levels or decreased effects of the object drug in the system as a result of the precipitant drug.

Concomitant use of torsemide and cholestyramine has not been studied in humans but, in a study in animals, coadministration of **cholestyramine** decreased the absorption of orally administered torsemide.

# Annotation Process

[http://lhce-brat.nlm.nih.gov/#/GoldenStandard\\_Iteration\\_13\\_Reconciled/2e0570e7-d28f-4936-cba8-81ee0c4c3547](http://lhce-brat.nlm.nih.gov/#/GoldenStandard_Iteration_13_Reconciled/2e0570e7-d28f-4936-cba8-81ee0c4c3547)

7 LEVOPHED should not be given to patients who are hypotensive from blood volume deficits except as an emergency measure to maintain coronary and cerebral artery perfusion until blood volume replacement therapy can be complete.

8 If LEVOPHED is continuously administered to maintain blood pressure in the absence of blood volume replacement, the following may occur: severe peripheral and visceral vasoconstriction, decreased renal perfusion and urine output acidosis.

9 LEVOPHED should also not be given to patients with mesenteric or peripheral vascular thrombosis (because of the risk of increasing ischemia and extending the area of infarction) unless, in the opinion of the attending physician, the

10 Cyclopropane and halothane anesthetics increase cardiac autonomic irritability and therefore seem to sensitize the myocardium to the action of intravenously administered epinephrine or norepinephrine.

11 Hence, the use of LEVOPHED during cyclopropane and halothane anesthesia is generally considered contraindicated because of the risk of producing ventricular tachycardia or fibrillation.

12 The same type of cardiac arrhythmias may result from the use of LEVOPHED in patients with profound hypoxia or hypercarbia.

15 WARNINGS

17 LEVOPHED should be used with extreme caution in patients receiving monoamine oxidase inhibitors (MAOI) or antidepressants of the triptyline or imipramine types, because severe, prolonged hypertension may result.

18 LEVOPHED Bitartrate Injection contains sodium metabisulfite, a sulfite that may cause allergic-type reactions including anaphylactic symptoms and life-threatening or less severe asthmatic episodes in certain susceptible people.

19 The overall prevalence of sulfite sensitivity in the general population is unknown.

20 Sulfite sensitivity is seen more frequently in asthmatic than in nonasthmatic people.

# Annotated corpus from Product Labels (DailyMed)



## Use in Patients with Heart Failure

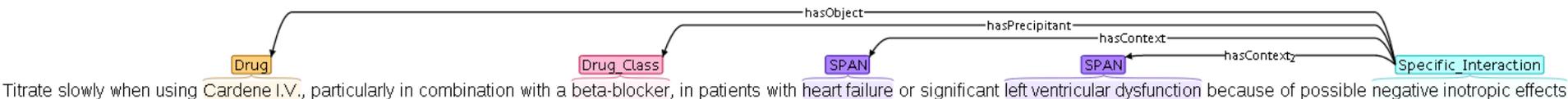
Titrate slowly when using Cardene® I.V., particularly in combination with a beta-blocker, in patients with heart failure or significant left ventricular dysfunction because of possible negative inotropic effects.

## Intravenous Infusion Site

To reduce the possibility of venous thrombosis, phlebitis, local irritation, swelling, extravasation, and the occurrence of vascular impairment, administer drug through large peripheral veins or central veins rather than arteries or small peripheral veins, such as those on the dorsum of the hand or wrist. To minimize the risk of peripheral venous irritation, change the site of the drug infusion every 12 hours.



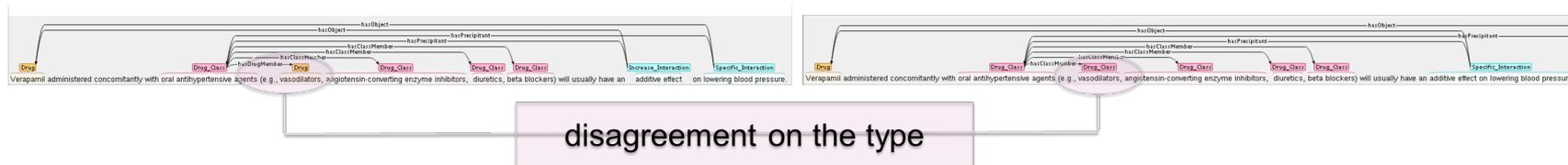
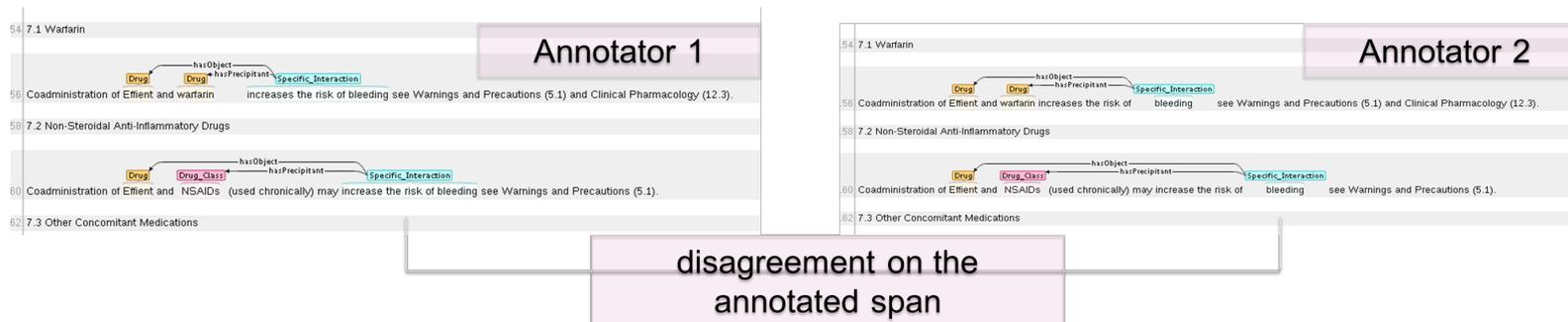
2 expert annotators using Brat



# Corpus Format

```
- <sentence type="regular" text="Because DEMADEX and salicylates compete for secretion by renal tubules, patients receiving high doses of salicylates may experience salicylate toxicity when DEMADEX is concomitantly administered. " id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67" lineNumber="64" biomedicalEntities="2">
- <entity type="Drug_Class" text="salicylate" id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.0" charOffset="132:142">
- <Normalization>
  <mmtx semType="phsu, orch" preferredWord="salicylate" cui="C0036075" phraseText="salicylate toxicity"/>
  <mmtx semType="phsu, orch" preferredWord="Salicylates" cui="C0036077" phraseText="salicylate toxicity"/>
  <RxNorm RxCui="9522"/>
</Normalization>
</entity>
<entity type="Increase_Interaction" text="toxicity" id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.1" charOffset="143:151"/>
- <entity type="Drug" text="DEMADEX" id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.3" charOffset="157:164">
- <Normalization>
  <mmtx semType="phsu, orch" preferredWord="Demadex" cui="C0243946" phraseText="DEMADEX"/>
  <RxNorm RxCui="71974"/>
</Normalization>
</entity>
- <drugInteraction id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.ddi.1">
- <interaction trigger="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.1">
- <relations>
- <relation type="hasObject">
  <entity id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.0"/>
</relation>
- <relation type="hasPrecipitant">
  <entity id="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.3"/>
</relation>
</relations>
</interaction>
</drugInteraction>
<pair type="Increase_Interaction" e2="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.3" e1="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.0" ddi="true" trigger="Dailymed.01b388a0-7dfb-11de-bc4e-0002a5d5c51b.s.67.e.1"/>
</sentence>
```

# Disagreement Analysis Tool (AMIA, 2014)



<http://lhce-brat.nlm.nih.gov/disagreementAnalyzer.htm>



## Facilitating Reconciliation of Inter-Annotator Disagreements

Johann Stan, PhD, Dina Demner-Fushman, MD, PhD, Kin Wah Fung, MD, MS, Olivier Bodenreider, MD, PhD

Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, National Institutes of Health, DHHS, Bethesda, MD



### Introduction

In the process of annotating a corpus of drug package inserts for drug-drug interactions, we were faced with a problem of reconciling differences in fairly complex annotations of interactions between drugs, drug classes and substances. Our goal was to annotate interactions for training supervised machine learning (ML) algorithms and evaluating the results. Annotated corpora are most useful for training ML tools if they are consistent.

To ensure consistency two experts annotated the interactions and two senior annotators adjudicated the disagreements. To facilitate annotations, we used Brat [1] that is fairly convenient for annotation, but does not provide mechanisms for reconciliation of disagreements. Therefore we have developed functions that allow reconciling disagreements and ensure consistency of annotation across similar interactions mentioned multiple times in the drug package inserts.

The tool was tested on a collection of 180 DailyMed product labels with the objective to create a corpus with sentences describing drug-drug interactions.

### Materials and methods

#### Materials

- 180 DailyMed Product Labels annotated with Brat [1] by two independent annotators
- Annotations consisted in identifying drugs, drug classes, food and drug interactions in the text within such entities

Example of annotated product label



#### Methods

- We divided the annotation process into iterations consisting of 10 product labels. After each iteration, we analysed the disagreements and developed functions to identify them automatically

### Results

#### Main Disagreement Reconciliation Functions

##### 1. Disagreement Reconciliation module

- Compares annotations by two annotators and uses line numbers in the files loaded to Brat for annotation to indicate the location of disagreements.

- Example of disagreement (increase the risk of bleeding - bleeding)



- Other Example of disagreement (spans annotated with different types: drug and drug class)



##### 2. The Sentence Validation module

- Identifies similar sentences and checks if they have been annotated consistently.
- Sentence similarity was computed using an implementation of the Jaccard similarity measure (threshold 0.75).
- Annotated biomedical entities as well as those identified by Metamap were replaced with standard names, e.g. "DRUG"

#### Example Output of the tool

- Results are written into a text file in the form of [Line number] - [Disagreement type]
- F-score computed for each entity type and global F-scores (the collection of one of the annotators is used as gold standard)

### Results

We tested the annotation reconciliation tools on 176 manually annotated package inserts (8081 biomedical entities and 4841 interactions). The tools identified 2584 discrepancies, of which 1200 were in entity annotation and 1384 in interaction annotations. 320 similar sentences were annotated inconsistently (e.g. different types attributed to same drugs or different interaction types).

### Conclusions

- We developed a Java toolkit for facilitating reconciliation of inter-annotator disagreements
- The tools helped us improve annotation guidelines and detect and reconcile discrepancies. To the best of our knowledge, this is the first publicly available extension for assisting with discrepancies and assuring consistency of annotations using Brat.
- The toolkit helped speed up the process of creating a gold standard
- The toolkit can be downloaded from the following link:

<http://goo.gl/8js841>

### Literature

Stenetorp P, Topic G, Ohta T, Ananiadou S, Tsujii J. Brat: Web-based Tool for NLP-Assisted Text Annotation. In Proceedings of the Demonstrations Session at EACL 2012, 2012.

### Acknowledgments

This work was supported by the Intramural Research Program of the NIH, National Library of Medicine (NLM).

We thank Larizza Rodriguez and Sonya Shooshan for annotating the 180 drug labels.



# NLM-DDI Corpus Statistics

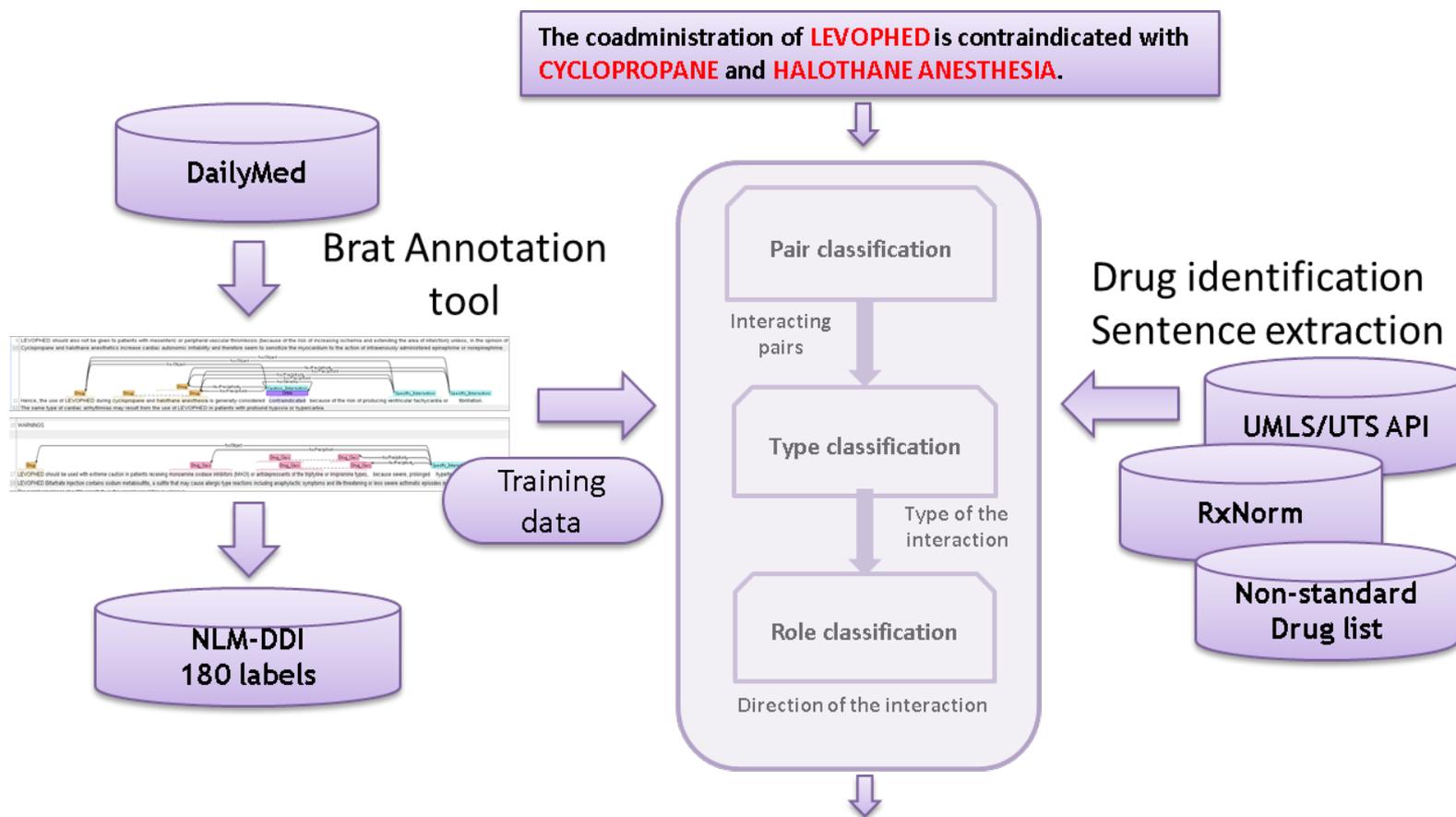
	Entities	Total number	%	Inter-Annotator Agreement
<b>Drug entities</b>	<b>Drug</b>	<b>4584 (592 distinct)</b>	<b>54.2%</b>	<b>0.81</b>
	<b>Drug Class</b>	<b>2816 (670 distinct)</b>	<b>33.3%</b>	<b>0.84</b>
	<b>Substance</b>	<b>221 (33 distinct)</b>	<b>2.7%</b>	<b>0.82</b>
	<b>Span</b>	<b>823 (290 distinct)</b>	<b>9.8%</b>	<b>0.56</b>
	<b>Total</b>	<b>8444</b>	<b>100%</b>	
<b>DDI roles</b>	<b>Specific Interaction</b>	<b>2595 (560 distinct triggers)</b>	<b>52.2%</b>	<b>0.79</b>
	<b>Caution Interaction</b>	<b>1308 (204 distinct triggers)</b>	<b>25.7%</b>	<b>0.72</b>
	<b>Increase Interaction</b>	<b>894 (289 distinct triggers)</b>	<b>17 %</b>	<b>0.84</b>
	<b>Decrease Interaction</b>	<b>262 (128 distinct triggers)</b>	<b>5%</b>	<b>0.9</b>
	<b>Total</b>	<b>5059</b>	<b>100%</b>	

- **2,963 DDIs, 14519 DDI candidate pairs**
  - **2705 positive pairs**
  - **11814 negative pairs**

# Public Sources of Drug-Drug Interactions

- 13 publicly available sources
- **5 sources for clinical application**
  - **CredibleMeds** ("Crediblemeds.org," 2013) - a list of DDIs thought to be clinically relevant and be supported by strong scientific evidence
  - **VA-NDF-RT** (Olvey, Clauschee, & Malone, 2010)
  - **ONC High Priority** (Phansalkar et al., 2012)
  - **ONC Non-interruptive** (Phansalkar et al., 2013)
  - **OSCAR** - a list of DDIs derived by expert consensus in the late 1990s (Crowther, Holbrook, Kenwright, & Kenwright, 1997)
- **3 sources to support Natural Language Processing (NLP)**
  - **DDI Corpus 2011** (Segura-Bedmar, Martinez, & Sánchez-Cisneros, 2011)
  - **DDI Corpus 2013** (Segura-Bedmar, Martinez, & Herrero-Zazo, 2013)
  - **PK DDI Corpus** (Boyce, Gardner, & Harkema, 2012)
  - **NLM CV Corpus (Stan, Demner-Fushman, Fung, Bodenreider)**
- **5 other sources** were developed to support either pharmacovigilance or bioinformatics applications (not discussed in this talk)
  - **KEGG DDI, TWOSIDES, DrugBank, SemMedDB-SemRep, DIKB**

# Extraction of DDIs: ML framework

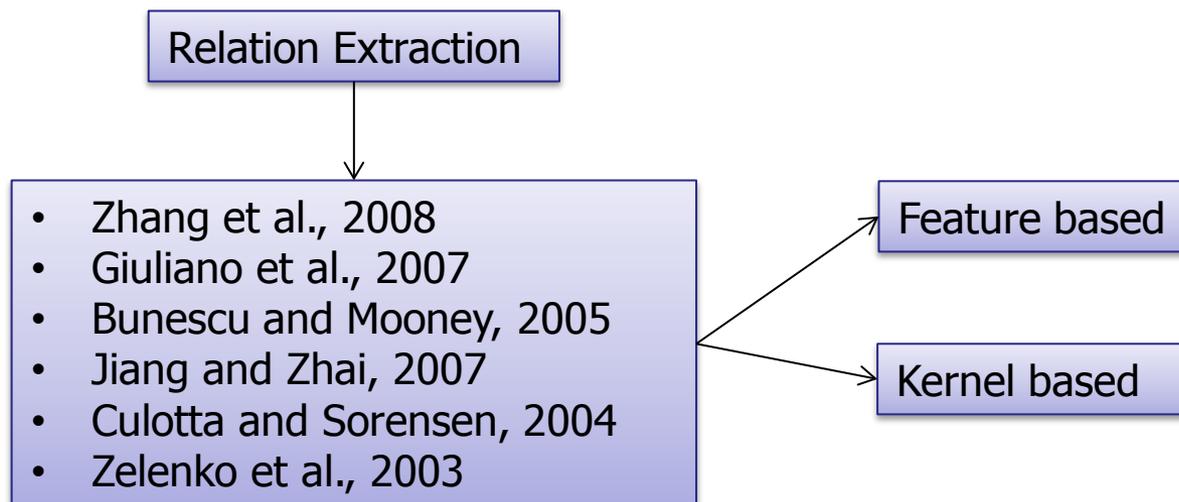


The coadministration of **LEVOPHED** is contraindicated with **CYCLOPROPANE** and **HALOTHANE ANESTHESIA**.

(LEVOPHED, CYCLOPROPANE) – DDI / CAUTION  
(LEVOPHED, HALOTHANE ANESTHESIA) – DDI / CAUTION  
(CYCLOPROPANE, HALOTHANE ANESTHESIA) – no DDI

recall of 25,7% and a precision of 48,7%  
lexical-syntactic patterns are not enough to detect  
all semantic relations occurring in text

# Machine Learning Approach for Relation Extraction

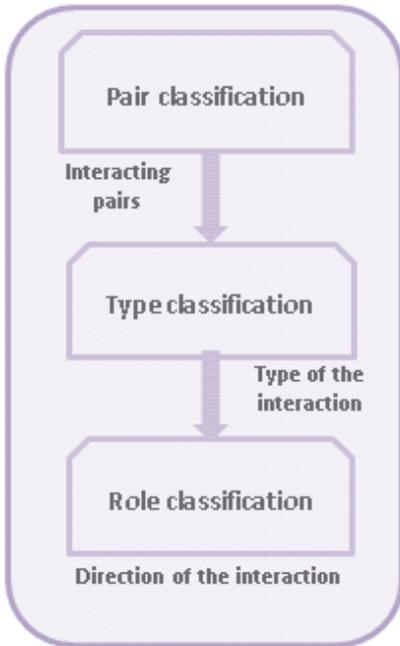


$$S = w_1 w_2 \dots DRUG_1 \dots DRUG_2 \dots w_n$$

$$F(T(S)) = \begin{cases} 1 & \text{if } DRUG_1 \text{ and } DRUG_2 \text{ interact} \\ 0 & \text{otherwise} \end{cases}$$

- tree kernels are relatively slow compared to feature classifiers and sequence kernels [Bunescu and Mooney, 2005, Li et al., 2008]
- integrated into real applications in which the processing time will be a priority

# Multi-Stage Classification Steps



DDI

**Inducers of CYP3A4** (e.g., rifampin) have caused a lowering of plasma levels of **verapamil**.

Decrease

**Inducers of CYP3A4** (e.g., rifampin) have caused a lowering of plasma levels of **verapamil**.

Precipitant

Object

**Inducers of CYP3A4** (e.g., rifampin) have caused a lowering of plasma levels of **verapamil**.

# Preprocessing

Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone

DDI (ASPIRIN, PROBENECID)

DDI (ASPIRIN, SULFINPYRAZONE)

DDI (ASPIRIN, PHENYLBUTAZONE)

- 1) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 1, because these drugs interact
- 2) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 1, because these drugs interact
- 3) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 1, because these drugs interact
- 4) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 0, because these drugs do not interact
- 5) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 0, because these drugs do not interact
- 6) Aspirin may decrease the effects of probenecid, sulfinpyrazone, and phenylbutazone  
=> label = 0, because these drugs do not interact

- 1) DRUG may decrease the effects of DRUG, OTHER, and OTHER => label = 1
- 2) DRUG may decrease the effects of OTHER, DRUG, and OTHER => label = 1
- 3) DRUG may decrease the effects of OTHER, OTHER, and DRUG => label = 1
- 4) OTHER may decrease the effects of DRUG, DRUG, and OTHER => label = 0
- 5) OTHER may decrease the effects of DRUG, OTHER, and DRUG => label = 0
- 6) OTHER may decrease the effects of OTHER, DRUG, and DRUG => label = 0

# Features for Support Vector Machines

## Linear Kernel/LIBSVM-Java Framework

### Feature types:

- Stems, n-grams of stems, POS tags, n-grams of POS tags
- Orthographic features
- Parse tree features
- Semantic features

### Global context features

Before-Between

Between

Between-After

Do not co-administer **DRUG** with **DRUG** in patients with diabetes.

### Local context features

Do not co-administer **DRUG** with **DRUG** in patients with diabetes.

# Examples of GC feature space

**DRUG1** may decrease the effects of **DRUG2**, DRUG3, and DRUG4.



Drug names replaced with DRUG or OTHER

**DRUG** may decrease the effects of **DRUG**, OTHER, and OTHER.



Word stems are used and POS tags  
N-grams with n=3  
Sparse POS n-grams

Fore-Between Features: **DRUG** may decrease the effects of **DRUG**, OTHER, and OTHER.

Between Features: **DRUG** may decrease the effects of **DRUG**, OTHER, and OTHER.

Between-After Features: **DRUG** may decrease the effects of **DRUG**, OTHER, and OTHER.

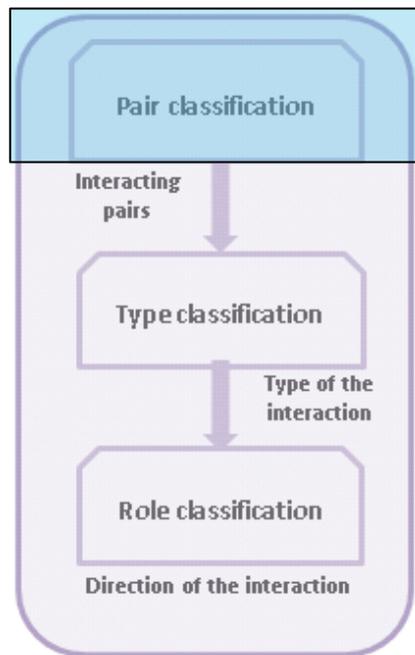
$K_{\text{Global context}}(\text{"**DRUG** may interact with **DRUG**"}, \text{"**DRUG** may interact with **DRUG**, OTHER, OTHER"})=2$

$K_{\text{Global context}}(\text{"**DRUG** may interact with **DRUG**"}, \text{"**DRUG** may decrease the effect of **DRUG**, OTHER, and OTHER"})=0$

$K_{\text{Global context}}(\text{"Coadministration of DRUG with DRUG may increase the risk of toxicity", "Coadministration of DRUG with DRUG may increase OTHER exposure"}, )=2$

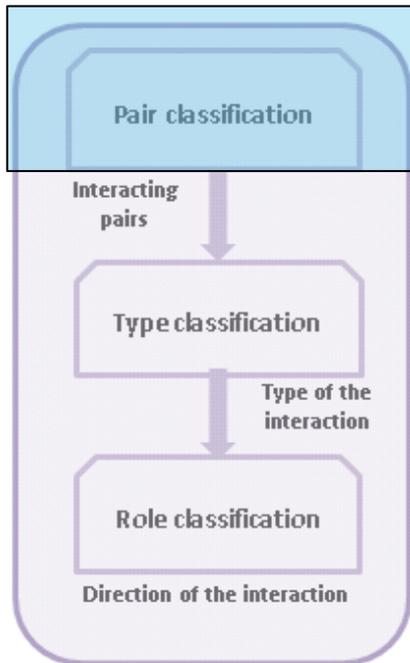


# Parameter selection for pair classification



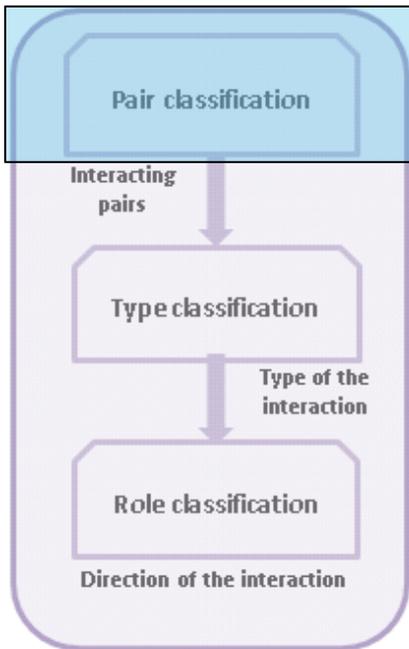
Feature set	Configuration	Precision	Recall	F-measure
Bag of words	N=1	0.573	0.902	0.701
	N=2	0.689	0.894	0.778
	N=3	0.689	0.888	0.776
Global Context (stems)	N=1	0.687	0.885	0.774
	N=2	0.738	0.861	0.795
	N=3	0.728	0.855	0.786
Global Context (stems, POS, sparse POS)	N=1	0.676	0.887	0.767
	N=2	0.744	0.853	0.795
	N=3	0.793	0.846	0.819
Shallow Linguistic (stems, POS, sparse POS)	N=1, W=3	0.797	0.872	0.833
	N=2, W=3	0.808	0.869	0.837
	N=3, W=2	0.814	0.870	0.841
Combined (stems, POS, sparse POS)	N=3, W=3 C=2	0.818	0.869	0.842

# Contribution of feature spaces for pair classification



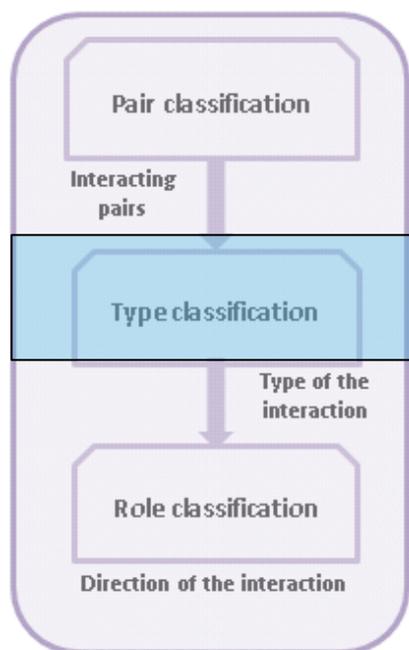
Feature set	Precision	Recall	F-measure
Local Context W = 3	0.74	0.826	0.781
Global Context Ngram = 3	0.778	0.854	0.814
<b>Combined features</b> <b>W = 3</b> <b>Ngram = 3</b>	<b>0.818</b>	<b>0.869</b>	<b>0.842</b>

# Contribution of GC feature types for pair classification



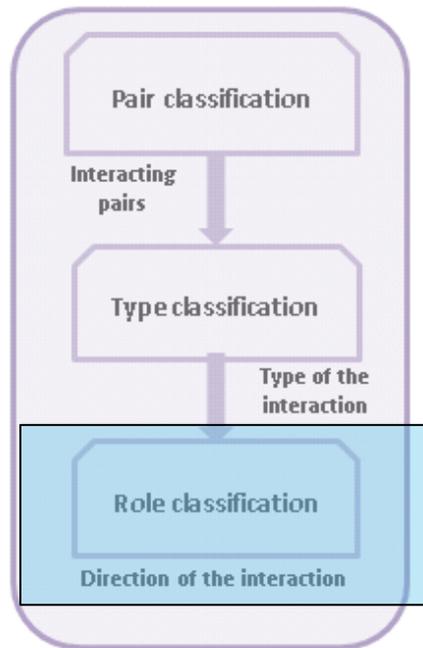
Feature set	Precision	Recall	F-measure
<b>Ngrams of stems</b>	<b>0.766</b>	<b>0.864</b>	<b>0.812</b>
<b>+ Ngrams of POS tags</b>	<b>0.816</b>	<b>0.869</b>	<b>0.841</b>
<b>+ Sparse stems</b>	<b>0.818</b>	<b>0.869</b>	<b>0.842</b>
<b>+ Sparse POS tags</b>	<b>0.797</b>	<b>0.861</b>	<b>0.828</b>
<b>All</b>	<b>0.805</b>	<b>0.854</b>	<b>0.829</b>

# Type Classification



Type of feature space	Type	Precision	Recall	F-measure
Local context only (LC) Window size 3	Specific	0.837	0.889	0.862
	Caution	0.865	0.851	0.858
	Increase	0.794	0.832	0.813
	Decrease	0.795	0.484	0.602
Global context only (GC) n-gram size 3	Specific	0.907	0.910	0.909
	Caution	0.898	0.913	0.905
	Increase	0.842	0.892	0.866
	Decrease	0.812	0.609	0.696
SL (LC + GC) Window, n-gram 3	Specific	0.893	0.926	0.909
	Caution	0.918	0.888	0.903
	Increase	0.843	0.897	0.869
	Decrease	0.812	0.625	0.714

# Role/direction Classification



Type of feature space	Type	Precision	Recall	F-measure
Local context only (LC)	1	0.769	0.938	0.845
	2	0.983	0.930	0.956
Global context only (GC)	1	1	1	1
	2	1	1	1
SL (LC + GC)	1	1	1	1
Window, n-gram 3	2	1	1	1

# Evaluation on the SemEval 2013 Dataset

Medline + DrugBank	Type	Tp	Fp	fn	Total	Precision	Recall	F-score	Best F-scores
Accuracy = 77.41585233441911% (713/921) (classification)	effect	175	49	43	921	0.781	0.803	0.792	0.662
	mechanism	229	41	58	921	0.848	0.798	0.822	0.679
	advice	268	113	54	921	0.703	0.832	0.762	0.692
	int	41	5	53	921	0.891	0.436	0.586	0.547

# DDI Extraction Tool

Drug-drug Interaction Extractor

**Choose LOINC section codes to be extracted**

- 34066-1 BOXED WARNING SECTION
- 34070-3 CONTRAINDICATIONS SECTION
- 34068-7 DOSAGE & ADMINISTRATION SECTION
- 34074-5 DRUG & OR LABORATORY TEST INTERACTIONS SECTION
- 34073-7 DRUG INTERACTIONS SECTION
- 42232-9 PRECAUTIONS SECTION
- 43685-7 WARNINGS AND PRECAUTIONS SECTION
- 34071-1 WARNINGS SECTION
- 34086-9 ABUSE SECTION
- 60555-0 ACCESSORIES
- 34084-4 ADVERSE REACTIONS SECTION
- 69761-5 ALARMS
- 34091-9 ANIMAL PHARMACOLOGY & OR TOXICOLOGY SECTION
- 60556-8 ASSEMBLY OR INSTALLATION INSTRUCTIONS
- 60557-6 CALIBRATION INSTRUCTIONS
- 34083-6 CARCINOGENESIS & MUTAGENESIS & IMPAIRMENT OF FERTILITY
- 34090-1 CLINICAL PHARMACOLOGY SECTION
- 60558-4 CLEANING, DISINFECTING, AND STERILIZATION INSTRUCTIONS
- 34092-7 CLINICAL STUDIES SECTION
- 69760-7 COMPATIBLE ACCESSORIES
- 60559-2 COMPONENTS
- 34085-1 CONTROLLED SUBSTANCE SECTION

**Metamap Configuration**

**User Name:** johannstan

**Password:** .....

**Mail:** johann.stan@nih.gov

Information about the extraction process will be displayed here

aapp-T116-Amino Acid, Peptide, or Protein  
arnas-T087-Amino Acid Sequence  
**antb-T195-Antibiotic**  
bacs-T123-Biologically Active Substance  
carb-T118-Carbohydrate  
chem-T103-Chemical  
**clnd-T200-Clinical Drug**  
enzy-T126-Enzyme  
**phsu-T121-Pharmacologic Substance**  
vita-T127-Vitamin

**Model file for binary classification**

C:\Users\stanj\workspace\DrugInteractionExtractor\Bin

**Model file for multiclass classification**

C:\Users\stanj\workspace\DrugInteractionExtractor\MClass

**Model file for direction classification**

C:\Users\stanj\workspace\DrugInteractionExtractor\Dir

**Input folder with DailyMed Product Labels**

C:\Users\stanj\workspace\DrugInteractionExtractor\DailyMedXML2

**Additional substance list**

C:\Users\stanj\workspace\DrugInteractionExtractor\NonStandardClasses

**Results folder**

C:\Users\stanj\workspace\DrugInteractionExtractor\Results

**Choose Functionality:**

Extract sentences with at least two substances

# Produced Output

A	B	C	D	E	F	G	H	I	J
SPL SET-ID	LABEL DRUG	GENERIC DRUG	SPL SECTION	SENTENCE	DRUG1	DRUG2	INTERACTION TYPE	OBJECT DRUG	PRECIPITANT DRUG
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	Human pharmacokinetic studies with an oral formulation of treprostinil ( treprostinil diolamine ) indicated that co-administration of the cytochrome P450 ( CYP ) 2C8 enzyme inhibitor gemfibrozil increases exposure ( both Cmax and AUC ) to treprostinil .	enzyme inhibitor	treprostinil	Increase Interaction	treprostinil	enzyme inhibitor
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	Human pharmacokinetic studies with an oral formulation of treprostinil ( treprostinil diolamine ) indicated that co-administration of the cytochrome P450 ( CYP ) 2C8 enzyme inhibitor gemfibrozil increases exposure ( both Cmax and AUC ) to treprostinil .	diolamine	enzyme inhibitor	No interaction		
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	Human pharmacokinetic studies with an oral formulation of treprostinil ( treprostinil diolamine ) indicated that co-administration of the cytochrome P450 ( CYP ) 2C8 enzyme inhibitor gemfibrozil increases exposure ( both Cmax and AUC ) to treprostinil .	enzyme inhibitor	gemfibrozil	Increase Interaction	enzyme inhibitor	gemfibrozil
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	7.2 Potassium Supplements and Potassium-Sparing Diuretics ( 250 mg/day ) and an oral formulation of treprostinil ( treprostinil diolamine ) , no pharmacokinetic interactions between treprostinil and bosentan were observed .	treprostinil	bosentan	No interaction		
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	7.2 Potassium Supplements and Potassium-Sparing Diuretics ( 250 mg/day ) and an oral formulation of treprostinil ( treprostinil diolamine ) , no pharmacokinetic interactions between treprostinil and bosentan were observed .	diolamine	bosentan	No interaction		
0fed2822-3a03-4b64-9857-c682fcd462bc	Aldactone	spironolactone	34073-7	7.2 Potassium Supplements and Potassium-Sparing Diuretics ( 250 mg/day ) and an oral formulation of treprostinil ( treprostinil diolamine ) , no pharmacokinetic interactions between treprostinil and bosentan were observed .	Potassium Supplement	bosentan	No interaction		

# FP Failure Analysis

Sentence in DailyMed	False Positive Pair	Possible Explanation
Hypokalemia may develop with LASIX, especially with brisk diuresis, inadequate oral electrolyte intake, when cirrhosis is present, or during concomitant use of corticosteroids, ACTH, licorice in large amounts, or prolonged use of laxatives.	corticosteroids- laxatives	Long sentence with several drugs.
In patients with an activated renin-angiotensin-aldosterone system, such as volume- or salt-depleted patients receiving high doses of diuretics, symptomatic hypotension may occur in patients receiving renin-angiotensin-aldosterone system (RAAS) blockers.	diuretics-renin- angiotensin- aldosterone	MetaMap wrongly identifies an entity renin-angiotensin-aldosterone is not a drug
As with other beta blockers, when discontinuation of TENORMIN is planned, the patients should be carefully observed and advised to limit physical activity to a minimum.	beta blockers- TENORMIN	Drug instance of a drug class, not drug-drug class interaction. Such cases can be captured using standard terminologies.

# FN Failure Analysis

Sentence in DailyMed	False Negative Pair	Possible Explanation
Phenytoin decreases serum amiodarone levels.	Phenytoin- amiodarone	Short sentences may lead to false negatives due to lack of context.
Use <b>amiodarone</b> with caution in patients receiving - receptor blocking agents (e.g., propranolol, a CYP3A inhibitor) or <b>calcium channel antagonists</b> (e.g., verapamil, a CYP3A substrate, and diltiazem, a CYP3A inhibitor) because of the possible potentiation of bradycardia, sinus arrest, and AV block, if necessary, amiodarone can continue to be used after insertion of a pacemaker in patients with severe bradycardia or sinus arrest.	amiodarone-calcium channel antagonists	Long sentences describing multiple interactions, composed of drug classes with examples of drug class members. Such sentences need additional processing, first decomposed and simplified. This example would then be transformed into several short sentences that are much simpler for the classifier.

# DDIs and drug classes

(Product labels cover only 9 ONC high priority interactions)

Substance level

In a few reported cases, co-administration of **verapamil** with **aspirin** has led to increased bleeding times greater than observed with aspirin alone.

**Grapefruit juice** may increase plasma levels of **verapamil**.

Drug-Drug Class Level Interaction

Clinically significant interactions have been reported with **inhibitors of CYP3A4** causing elevation of plasma levels of **verapamil** while **inducers of CYP3A4** have caused a lowering of plasma levels of **verapamil**.

Drug Class - Drug Class Interaction

Concomitant therapy with **beta-adrenergic blockers** and **calcium ion influx inhibitors** may result in additive negative effects on heart rate, atrioventricular conduction and/or cardiac contractility.

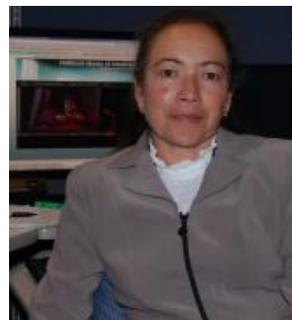
# Conclusion and Future Work

- A new corpus of DDIs extracted from DailyMed labels
- The most complete ML approach for DDI extraction (role, type, direction)
- Submitted to JBI Special Issue

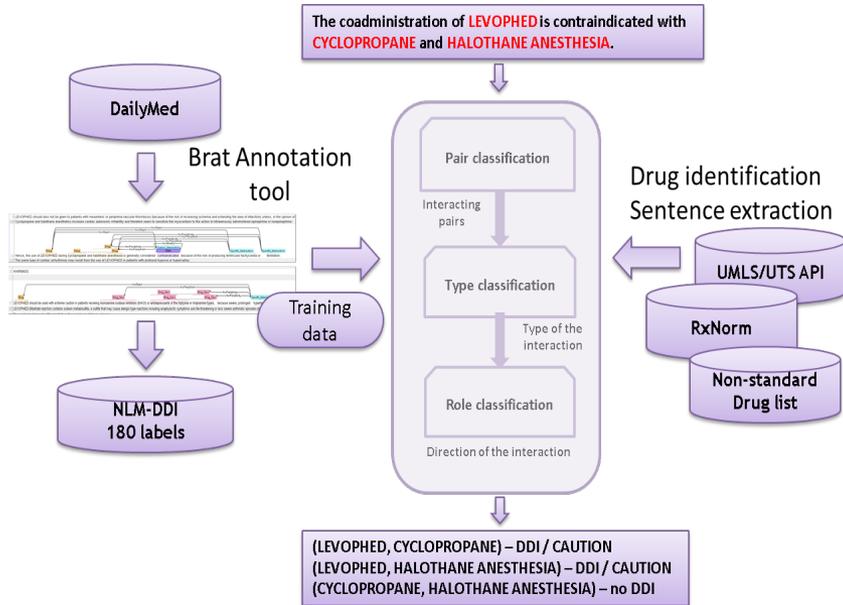
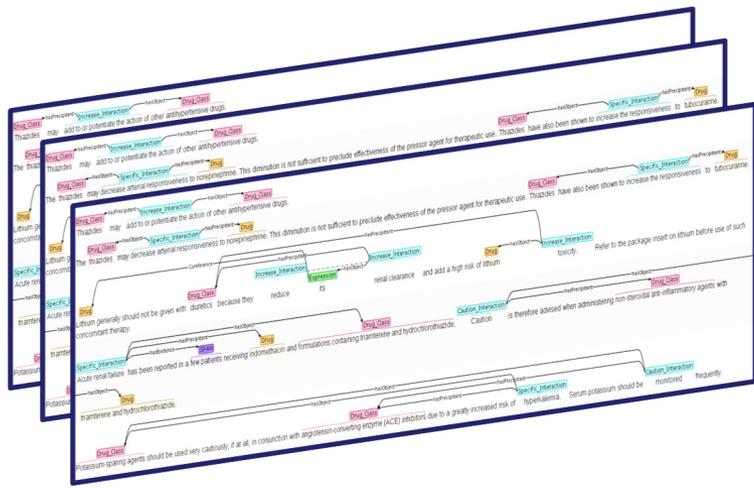
- Management of drug classes (standard and non-standard)
- Management of anaphoric constructions
- Improvements of the feature space with tree and semantic features
- Non-linear kernels
- Ensemble-learning

# Acknowledgment

## DDI Team



**Josephine O'Para**  
**Medical Informatics Postdoctoral Research Fellowships Program**  
**Dr. Paul Fontelo**  
**ORISE**  
**NLM Fellows**



johann.stan.phd@gmail.com

# Comparison of SemEval – NLM Annotation Schemas

- We do not consider brand names and active ingredient names as separate entities (all are annotated as DRUG)
- CAUTION – ADVICE, INT
- DECREASE/INCREASE – MECHANISM
- SPECIFIC – EFFECT

# Overlap Analysis using DrugBank mappings

- Subset of DDIs where both drugs involved in an interaction could be mapped

Clinically-oriented information sources			NLP Corpora			Bioinformatics/Pharmacovigilance		
Source	Description	mapped/ original	Source	Description	mapped/ original	Source	Description	mapped/ original
Crediblemeds.org	A list of clinically important drug-drug interactions	82/83	DDI Corpus 2011	Training corpus for the 2011 SemEval DDI NLP challenge	586/3,160	KEGG DDI	PDDIs extracted from the interaction tables of Japanese product labels	26,664/298,337
VA NDF-RT	PDDIs used by the Veteran's Administration health care system	2,606/5,265	DDI Corpus 2013	Training corpus for the 2013 SemEval DDI NLP challenge	1,287/5,021	TWOSIDES	Pharmacovigilance signals indicative of a possible drug-drug adverse event association	9,921/63,473
ONC High Priority	A consensus list of PDDIs that are high priority for inclusion in CDS alerting	1,150/1,150	PK DDI Corpus	Training corpus for NLP to extract pharmacokinetic PDDIs from drug product labels	166/298	DrugBank	Comprehensive drug information resource	12,113
ONC Non-interruptive	A consensus list of PDDIs for which alerts should be non-interruptive	2,101/2,101	NLM CV DDI Corpus			DIKB	An evidence-focused PDDI knowledge base	581/581
OSCAR	PDDIs used on an open source EHR system	7,969/7,969				SemMedDB	PDDIs extracted by NLP from the titles and abstracts in PubMed	3,952/190,219

# ONC High priority DDI examples

**Table 2** List of candidate drug–drug interactions (DDIs) discussed and the final pairs accepted by the expert panel as critical DDIs

#	Candidate drug–drug interaction pair (object–precipitant drug/class)	Status	Considerations suggested by the expert panel	Final DDI pair and suggested membership*	
				Object class	Precipitant class
3	Amphetamine and derivatives–MAO inhibitors	Accepted	Consider downgrading membership of selegiline due to its selective MAO-B inhibition; only at higher doses does it lose specificity and inhibit MAO-A	Amphetamine derivatives: Dexmethylphenidate Dextroamphetamine Methylphenidate Lisdexamfetamine Methamphetamine Phendimetrazine Pseudoephedrine Amphetamine Benzphetamine Diethylpropion Phentermine Atomoxetine	MAO inhibitors: Tranylcypromine Phenelzine Isocarboxazid Procarbazine Selegiline
4	Atazanavir–gastric pH alkalizing agents (proton pump inhibitors (PPIs) + H <sub>2</sub> blockers)	Accepted	1. Only include PPIs and remove H <sub>2</sub> blockers from precipitant class based on literature evidence 2. Add dexlansoprazole to precipitant class	Atazanavir	Proton pump inhibitors (PPIs): Omeprazole Lansoprazole Pantoprazole Rabeprazole Esmoprazole



# Overlap Analysis

- DrugBank and KEGG covered the most drug pairs across other sources (28.6% and 25.6% respectively).

	<b>SemMedDB</b>												
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>											
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>										
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>									
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>								
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>							
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>						
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>					
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>				
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>			
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>		
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>	
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)	

# Overlap Analysis

- SemEval 2013 -> SemEval 2011 (91.3%)

	<b>SemMedDB</b>												
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>											
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>										
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>									
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>								
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>							
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>						
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>					
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>				
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>			
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>		
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>	
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)	

# Overlap Analysis

- DrugBank -> CredibleMeds (69.5%)

	<b>SemMedDB</b>												
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>											
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>										
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>									
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>								
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>							
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>						
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>					
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>				
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>			
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>		
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>	
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)	

# Overlap Analysis

- no DDIs common to all 14 sources
- PK DDI, TWOSIDES, SemMedDB -> no overlap with CredibleMeds
- ONC non Interruptive -> no overlap with DIKB

	<b>SemMedDB</b>											
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>										
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>									
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>								
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>							
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>						
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>					
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>				
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>			
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>		
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>	
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)

# Overlap Analysis

- only 4 DDIs
  - Haloperidol/Clozapine,
  - Triazolam/Voriconazole,
  - Triazolam/Fluconazole,
  - Midazolam/Fluconazole

common to the Bioinformatics/Pharmacovigilance sources

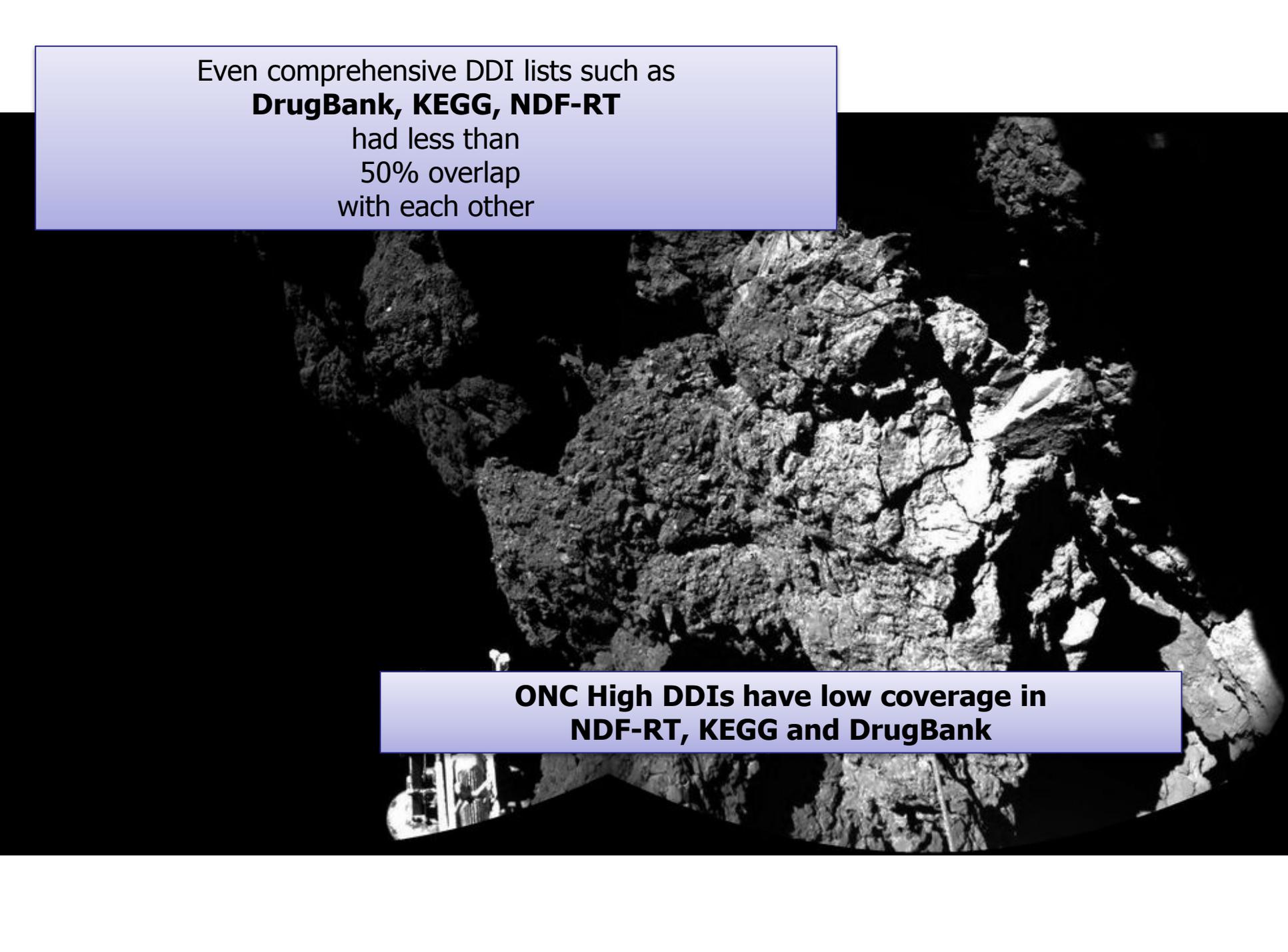
	<b>SemMedDB</b>											
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>										
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>									
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>								
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>							
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>						
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>					
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>				
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>			
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>		
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>	
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)



# Overlap Analysis

- CredibleMeds, NDF-RT, ONC Non-Interruptive no common DDI

	<b>SemMedDB</b>												
<b>Credible Meds</b>	0 (0.0%, 0.0%)	<b>Credible Meds</b>											
<b>NDF-RT</b>	69 (2.7%, 1.7%)	16 (0.6%, 19.5%)	<b>NDF-RT</b>										
<b>ONC High Priority</b>	12 (1.0%, 0.3%)	8 (0.7%, 9.8%)	225 (19.6%, 8.7%)	<b>ONC High Priority</b>									
<b>ONC Non-interruptive</b>	8 (0.4%, 0.2%)	4 (0.2%, 4.9%)	27 (1.3%, 1.0%)	2 (0.1%, 0.2%)	<b>ONC Non-interruptive</b>								
<b>OSCAR</b>	124 (1.6%, 3.1%)	23 (0.3%, 28.0%)	201 (2.5%, 7.7%)	44 (0.6%, 3.8%)	861 (10.8%, 41.0%)	<b>OSCAR</b>							
<b>DDI Corpus 2011</b>	68 (11.6%, 1.7%)	4 (0.7%, 4.9%)	162 (27.6%, 6.2%)	13 (2.2%, 1.1%)	4 (0.7%, 0.2%)	67 (11.4%, 0.8%)	<b>DDI Corpus 2011</b>						
<b>DDI Corpus 2013</b>	114 (8.9%, 2.9%)	5 (0.4%, 6.1%)	295 (22.9%, 11.4%)	23 (1.8%, 2.0%)	5 (0.4%, 0.2%)	112 (8.7%, 1.4%)	535 (41.6%, 21.3%)	<b>DDI Corpus 2013</b>					
<b>PK DDI Corpus</b>	12 (7.2%, 0.3%)	0 (0.0%, 0.0%)	50 (30.1%, 1.9%)	1 (0.6%, 0.1%)	1 (0.6%, 0.0%)	22 (13.3%, 0.3%)	28 (16.9%, 4.8%)	51 (30.7%, 4.0%)	<b>PK DDI Corpus</b>				
<b>KEGG</b>	403 (1.5%, 10.2%)	27 (0.1%, 32.9%)	777 (2.9%, 29.9%)	159 (0.6%, 13.8%)	511 (1.9%, 24.3%)	844 (3.2%, 10.6%)	218 (0.8%, 37.2%)	419 (1.6%, 32.6%)	77 (0.3%, 46.4%)	<b>KEGG</b>			
<b>TWOSIDES</b>	51 (0.5%, 1.3%)	0 (0.0%, 0.0%)	82 (0.8%, 3.2%)	25 (0.3%, 2.2%)	40 (0.4%, 1.9%)	101 (1.0%, 1.3%)	14 (0.1%, 2.4%)	25 (0.3%, 1.9%)	11 (0.1%, 6.6%)	724 (7.3%, 2.7%)	<b>TWOSIDES</b>		
<b>DRUGBANK</b>	150 (1.2%, 3.8%)	57 (0.5%, 69.5%)	1296 (10.7%, 49.9%)	319 (2.6%, 27.7%)	180 (1.5%, 8.6%)	490 (4.0%, 6.1%)	213 (1.8%, 36.3%)	448 (3.7%, 34.8%)	75 (0.6%, 45.2%)	2143 (17.7%, 8.0%)	289 (2.4%, 2.9%)	<b>DRUG BANK</b>	
<b>DIKB</b>	2 (0.4%, 0.1%)	21 (3.7%, 25.6%)	85 (15.2%, 3.3%)	33 (5.9%, 2.9%)	0 (0.0%, 0.0%)	7 (1.2%, 0.1%)	25 (4.5%, 4.3%)	36 (6.4%, 2.8%)	16 (2.9%, 9.6%)	152 (27.1%, 0.6%)	69 (12.3%, 0.7%)	189 (33.7%, 1.6%)	



Even comprehensive DDI lists such as  
**DrugBank, KEGG, NDF-RT**  
had less than  
50% overlap  
with each other

**ONC High DDIs have low coverage in  
NDF-RT, KEGG and DrugBank**