VANDERBILT UNIVERSITY School of Medicine

Medication resources for EHR research

Adverse Drug Effects

- **Adverse Drug Effects** (ADEs), also called Adverse Drug Reactions (ADRs), are defined by the World Health Organization as:

  "a response to a drug that is noxious and unintended and occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function."

Drug Knowledgebases

- Database of accurate drug-IND and drug-ADE relationships would benefit:
  - **Pharmacovigilance**
  - **Clinical Data Mining**
  - **Clinical Phenotyping**
  - **Decision Support Systems**
  - **Other applications**

Exisiting Work

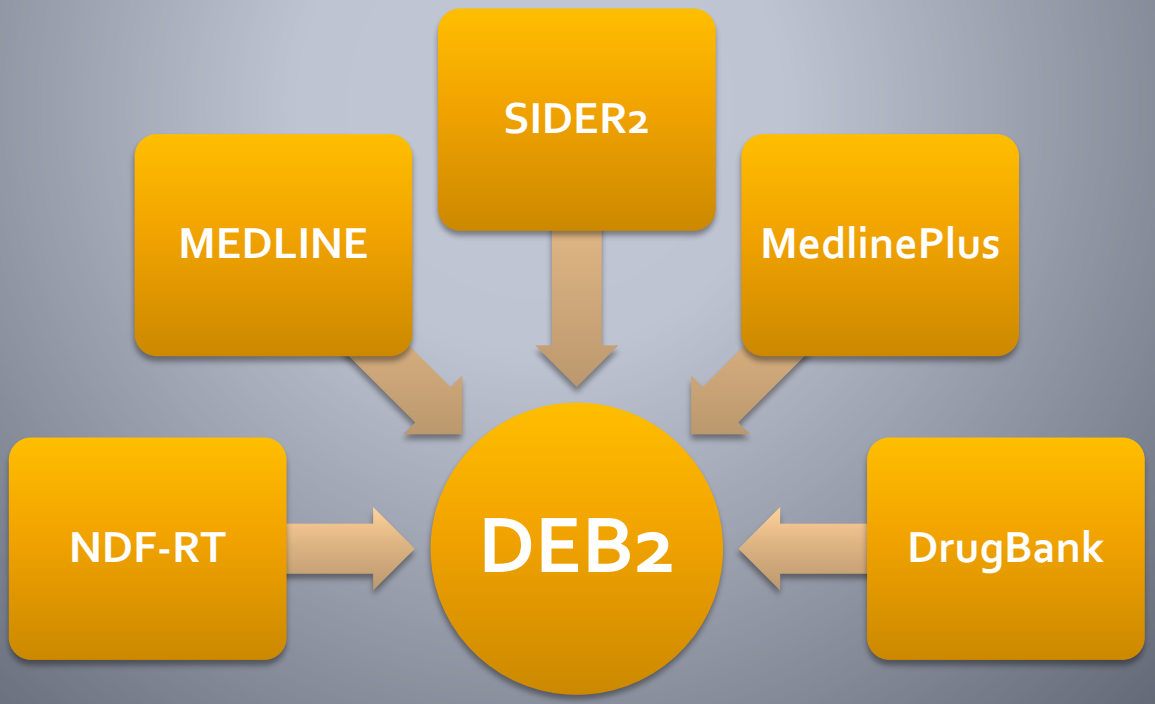
- **Commercial** repositories:
  - FDB, Micromedex, PDR, Epocrates, etc.
- **Public** data:
  - DailyMed, FAERS, RxNorm, NDF-RT, FDA, etc.
- **Academic** studies: (of many)
  - **2010, Wang, et al.**, combined data from AERS, SemMed, and NDF-RT to infer reasons for Rx
  - **2010, Kuhn, et al.**, SIDER database, extracted from FDA Structured Product Labels (SPLs)

Existing Work

- **Academic** studies: (continued)
  - **2011, Li, et al.**, combined data from FAERS, Micromedex, and NDF-RT to infer reasons for drug prescriptions
  - **2012, Kuhn, et al.**, updated SIDER2
  - **2012, Harpaz, et al.**, reviewed drug knowledge sources for pharmacovigilance.
  - **2013, Wei, et al.**, developed MEDI, combining indications from RxNorm, SIDER, MedlinePlus, and Wikipedia.
  - **2012,2013, Smith, et al.**, developed an early version of the Drug Evidence Base (DEB1) from MRCOC, NDF-RT, and FDA Structured Product Labels (SPLs).

The Drug Evidence Base (DEB2)

Slides from Josh Smith, PhD



The Original DEB1 Drug Evidence Base

- 2013, Smith, et al. – "Lessons Learned from Developing a Drug Evidence Base to Support Pharmacovigilance."
- DEB1 was only 61% accurate.
- Comparison of DEB1 to other knowledge-bases revealed:
  - Nomenclature mismatches impede comparison between drug information KBs
  - Different concepts used across sources and KBs

Drug Evidence Base (DEB2)

- **Objective** – Create an accurate, machine-processable drug knowledge base mined from reliable public sources.
- Concepts
  - **Drugs** – Single-ingredient medications
  - **Clinical Manifestations** (CMs) –Diseases, Syndromes, Symptoms, Findings, etc.
- Relationships (Drug-CM pairs)
  - **ADEs** – Drug causes exacerbates CM
  - **Indications** (INDs) – Drug treats or prevents CM
- Required relationships to be found in **at least 2 sources**

Constructing DEB2: CMs

- Using **UMLS2013ab**, CMs restricted to
  1. Concepts in SNOMED CT
  2. Specified UMLS semantic types

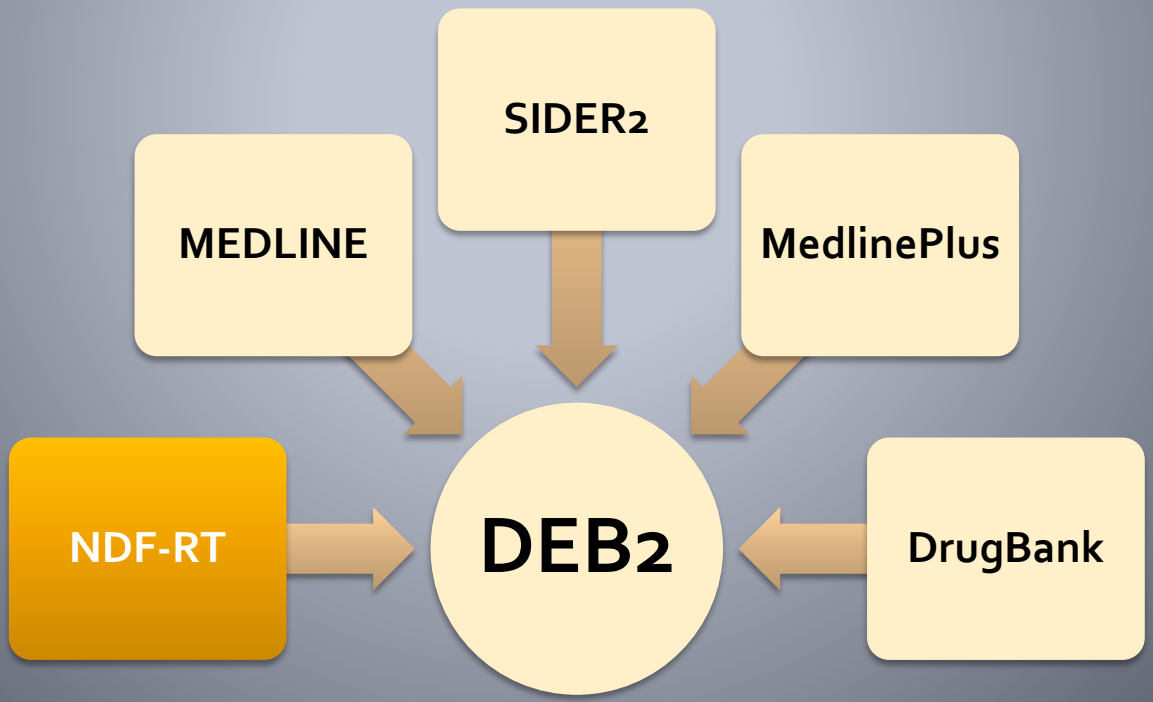
     Anatomical Abnormality, Injury or Poisoning, Congenital Abnormality, Finding, Sign or Symptom, Acquired Abnormality, Clinical Attribute, Disease or Syndrome, Mental or Behavioral Dysfunction, Neoplastic Process, Pathologic Function

Constructing DEB2: Medications

- Eliminated vague drug concepts by limiting DEB2 to only "clinical drug" concepts in the *RxNorm prescribable subset*
  - Extracted **76,212** "clinical drugs"
  - Normalized to **3059** single ingredients
  - Removed "drugs" with unwanted semantic types and unwanted terms
- Result: **1844** single-ingredient drugs (RxCUIs)

Constructing DEB2

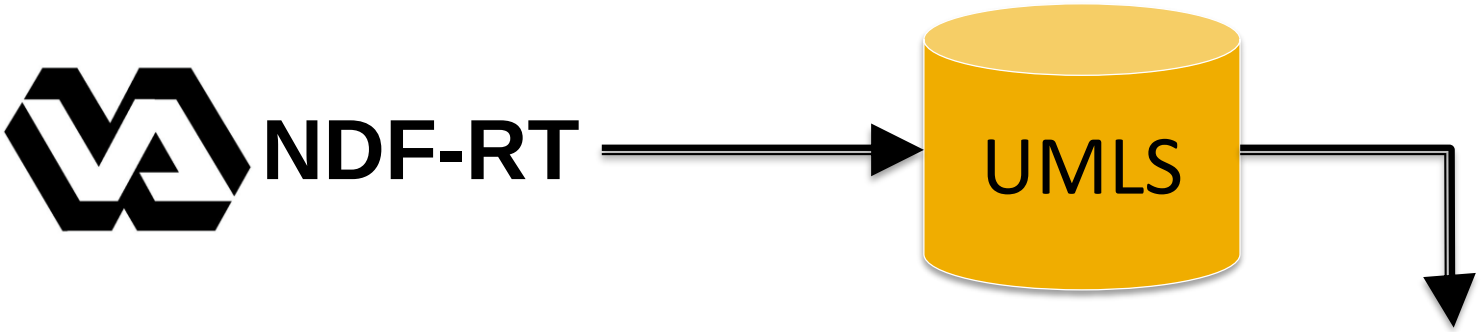
National Drug File – Reference Terminology (NDF-RT)



DEB2: Extraction from NDF-RT

- **NDF-RT is a formal drug representation**
  - Includes ingredients, dose forms, physiologic effects, mechanisms of action, and **25** distinct relationships
- **DEB2 extracts all drug-CM pairs with one of the following NDF-RT relationships:**
  - "induces" (ADE)
  - "may prevent" (IND)
  - "may treat" (IND)

DEB2: Extraction from NDF-RT



Drug Concept	Relationship	CM Concept
Lisinopril	INDICATION	Congestive Heart Failure
Lisinopril	INDICATION	Hypertension
Lisinopril	INDICATION	Left Ventricular Hypertrophy
Lisinopril	ADE	Cough

DEB2: Extraction from NDF-RT

NDF-RT	# Rows
induces	722
may_treat	48922
may_prevent	6114
[illegible]	[illegible]

Distinct Concepts	
Drugs	9596
CMs	1030

Modified Drugs	
Normalized Drugs	4133
In RXN subset	1153

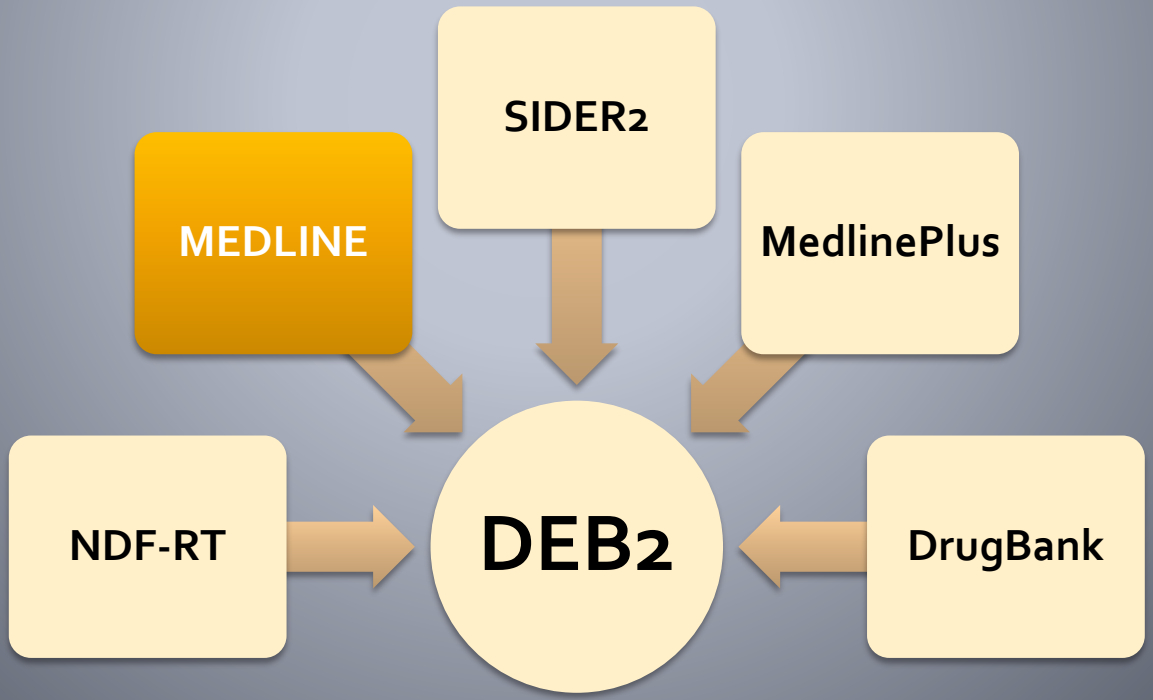
Modified CMs	
CMs also in SNOMEDCT	958 (-72)
CMs of correct semantic type	831

Pairs Extracted	
IND	4055
ADEs	78
Total	4133

NDF-RT

Constructing DEB2

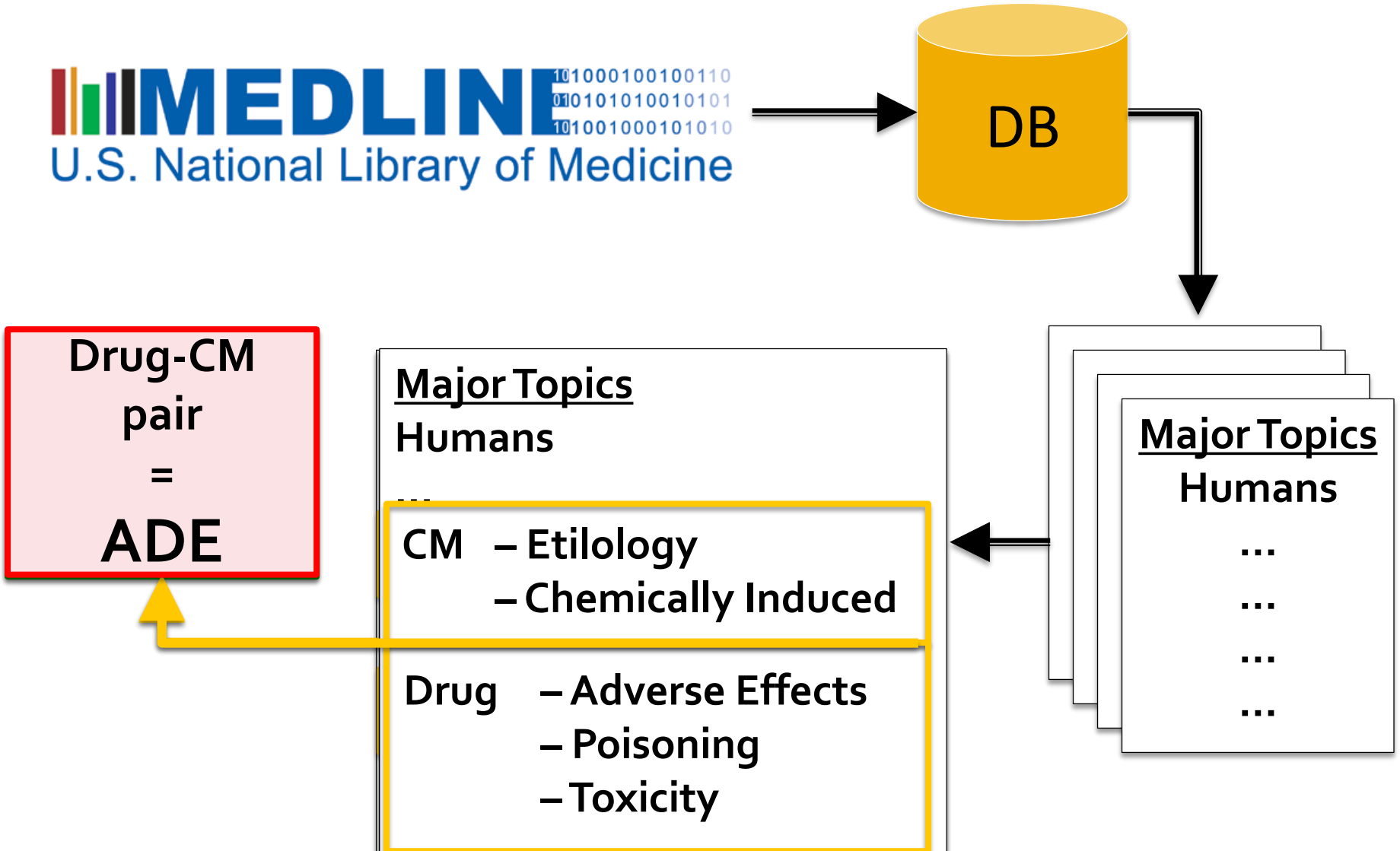
MEDLINE



Previous Work

- Drug knowledge using Medline:
  - **Zeng & Cimino, 1998** – extracted drug-disease relationships from MRCOC
  - **Shetty & Dalal, 2011** – disproportionality analysis of articles with predefined MeSH terms to discover unrecognized ADEs.
  - **Xu & Wang, 2013** – extracted drug-disease treatment relations using pattern-learning on MEDLINE abstracts.
  - **Avillach, et al., 2013** – extracted ADRs from MEDLINE using MeSH; minimum 3 articles.

DEB2: Extraction from MEDLINE



DEB2: Extraction from MEDLINE

MEDLINE2014	
Articles	~22M
"Humans"	~13M
MeSH	~50M
• ET	2.9M
• TU	2.6M
• DT	2.4M
• AD	1.5M
• AE	927K
• CI	671K
• TX	201K
• PO	24K

Relevant Articles	
Articles	~600,000
Pairs	~1,300,000

IND/ADE Relationships	
IND	32287
ADE	11420

Modified CMs	
Removed	290
Remaining	2522

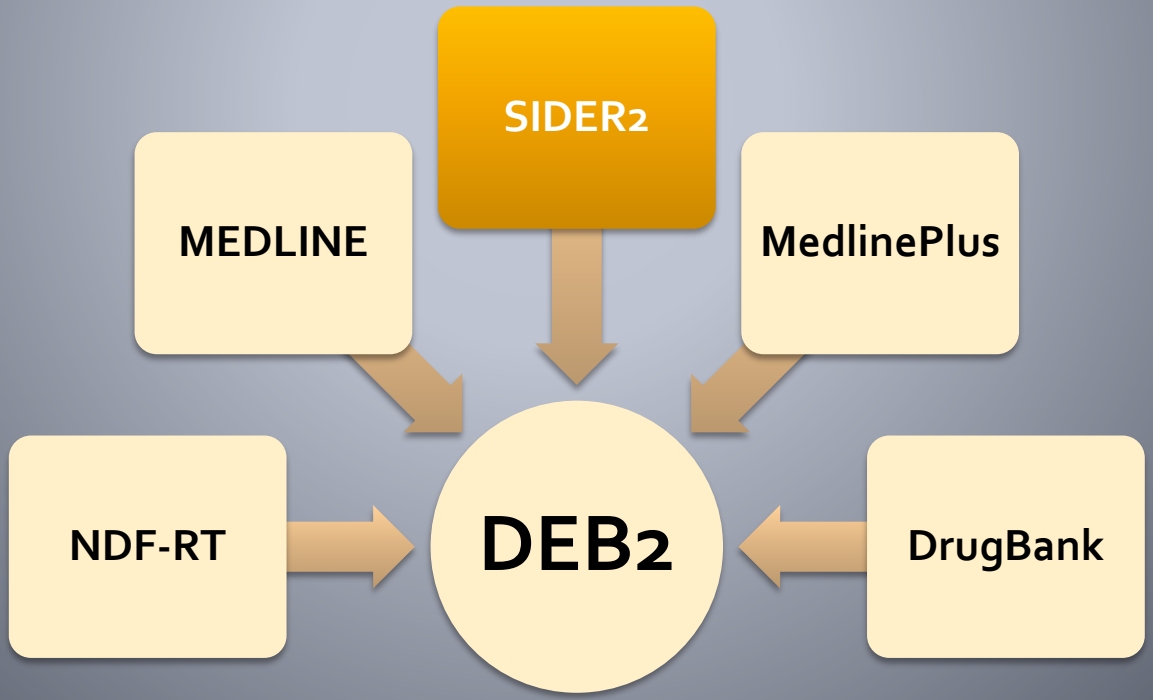
Combined Pairs	
INDs	31200
ADEs	9650
Total	40850

New Combined Pairs	
INDs	22732
ADEs	6331
Total	29063

- After manual review, we used an adjusted the article threshold and used each article's abstract to refine our inclusion criteria

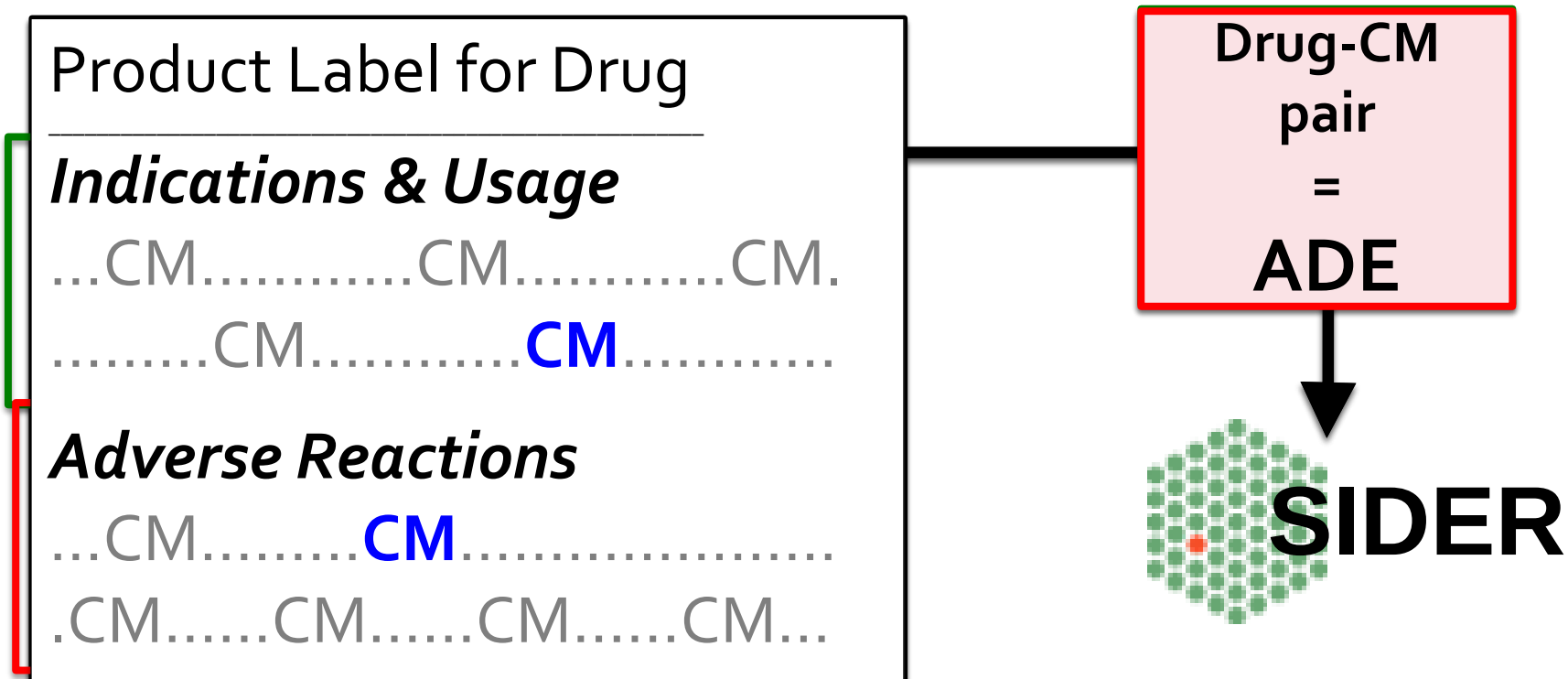
Constructing DEB2

SIDER2 Side Effects Resource



DEB: Extraction from SIDER2

- **SIDER2** is a database of indications and ADEs extracted from FDA Structured Product Labels (2012, Kuhn, et al.)



DEB: Extraction from SIDER2

Mapping SIDER2 Raw Data	
Label IDs	32140
Clinical Drugs	20507
Drugs from RXN Subset	931

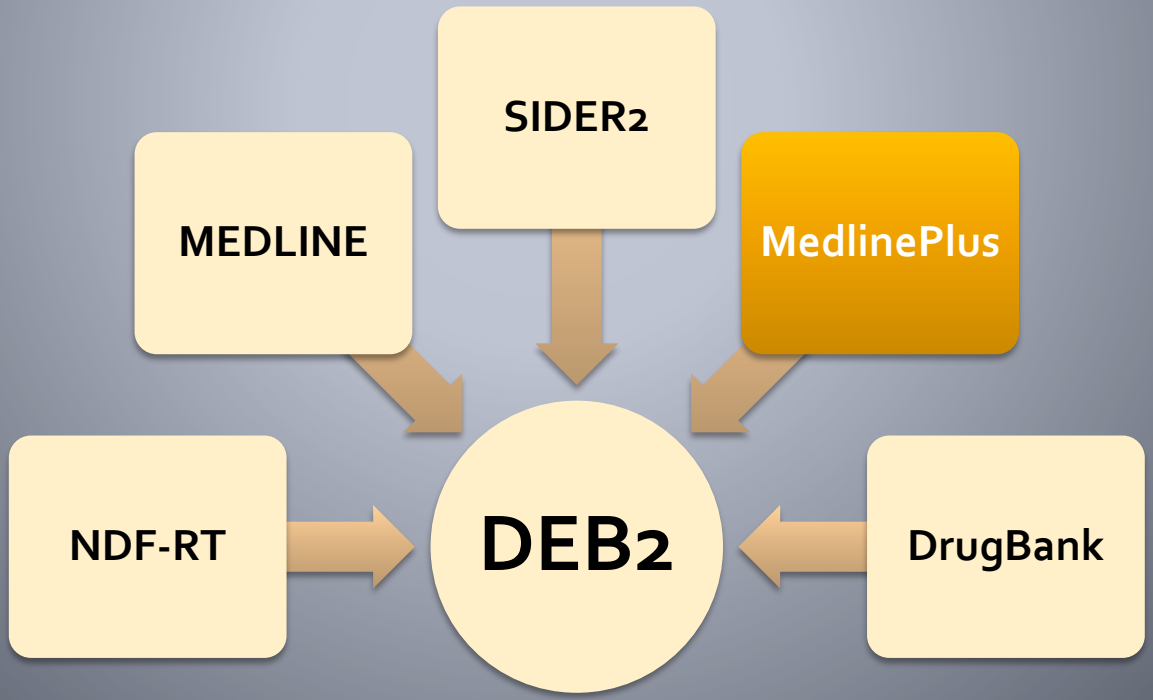
SIDER2 Concepts	
CMs in SNOMED	3815
CMs not in SNOMED	1351

CM Concepts by Section	
Indications & Usage	9646
Adverse Reactions	87126

Total Pairs	
IND	9646
ADEs	83956
Total	93602

Constructing DEB2

MedlinePlus

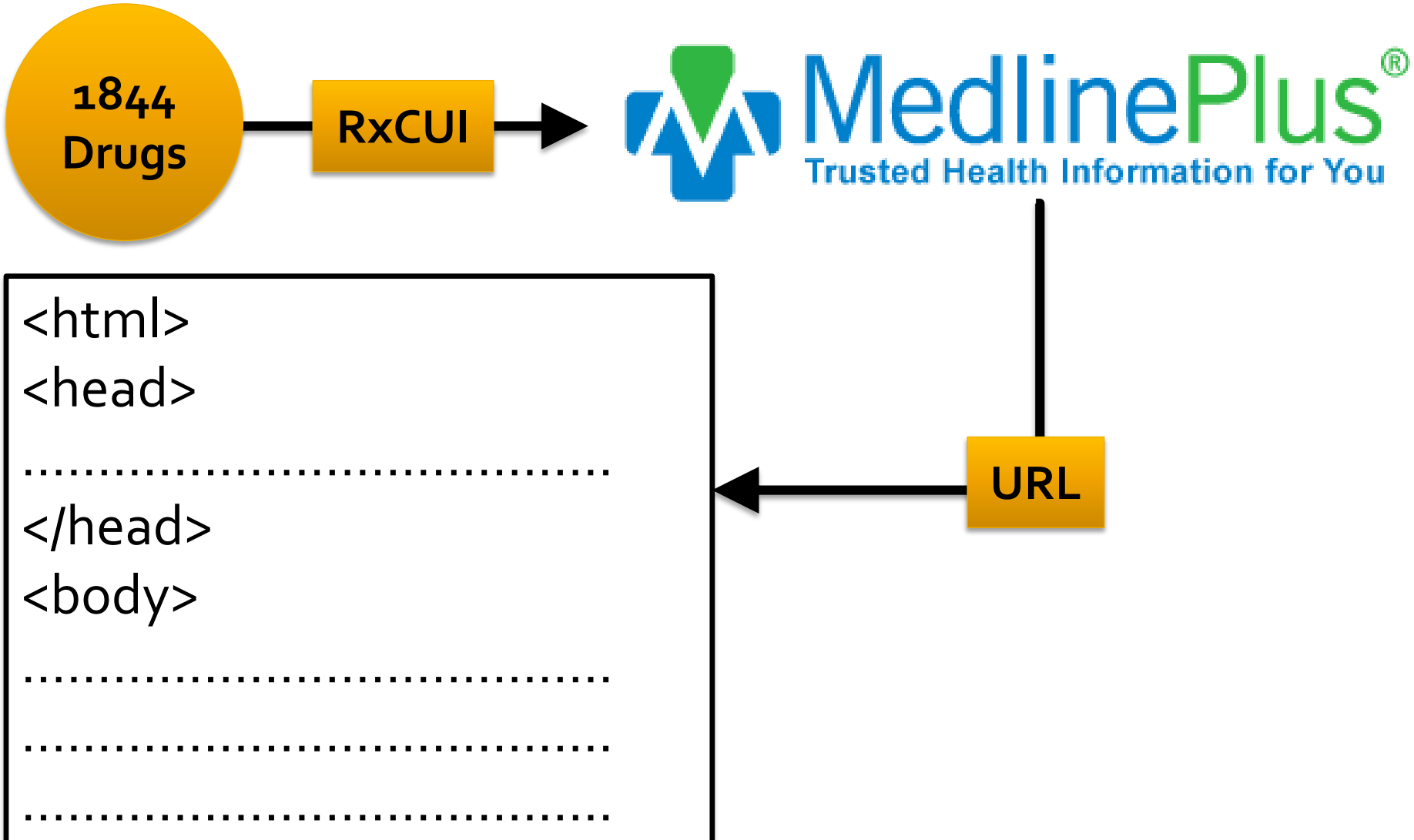


DEB2: From MedlinePlus

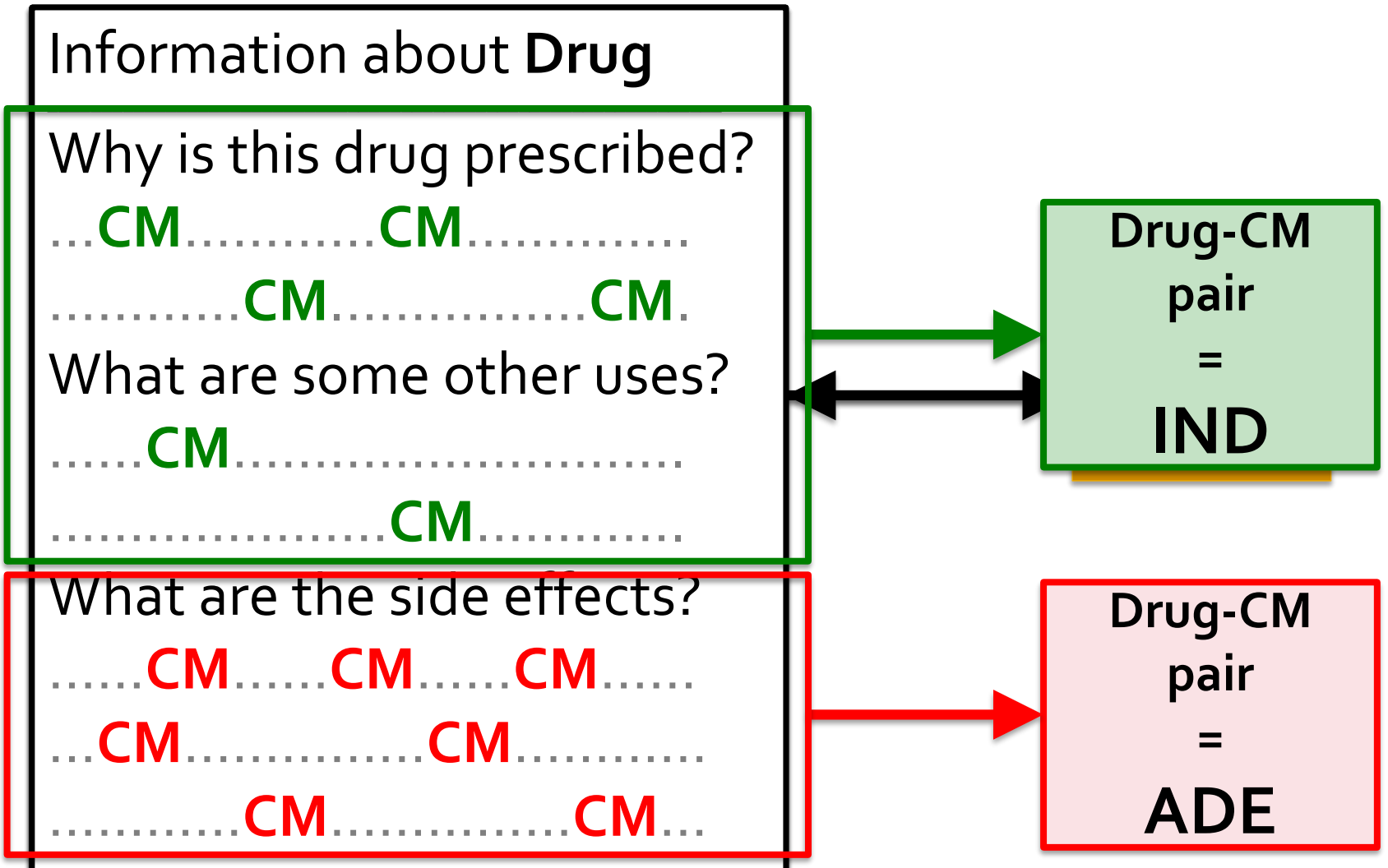
- **MedlinePlus** is a consumer health website from NLM and NIH.
- Among other items, the site contains drug monographs answering such questions as:
  - Why is this medication prescribed?
  - What are other uses of the medicine?
  - What side effects might this medication cause?

MedlinePlus®
Trusted Health Information for You

DEB2: From MedlinePlus



DEB2: From MedlinePlus



DEB2: From MedlinePlus

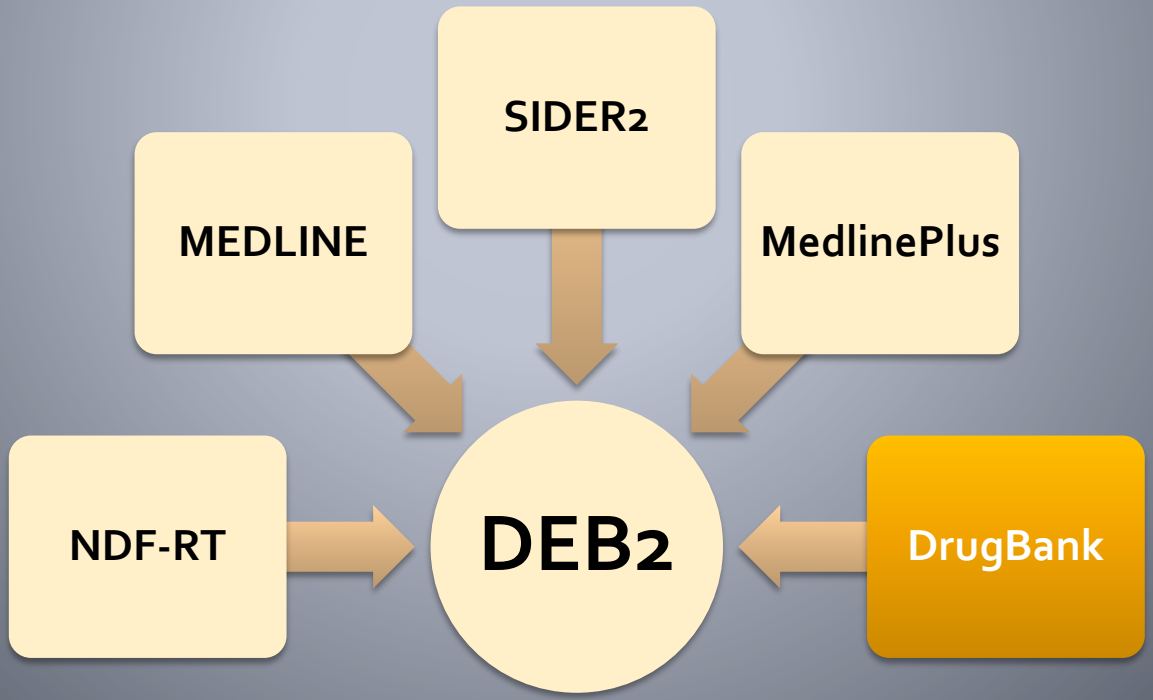
RxNorm Subset	
Drugs	1844
MedlinePlus	
Found	955

CM Concepts by Section	
Indications	5325
Side Effects	18699
• Serious	9640
• Common	8734
Overdose	3190
Boxed Warning	5004

Total Pairs	
IND	5325
ADEs	23444
Total	28769

Constructing DEB2

DrugBank

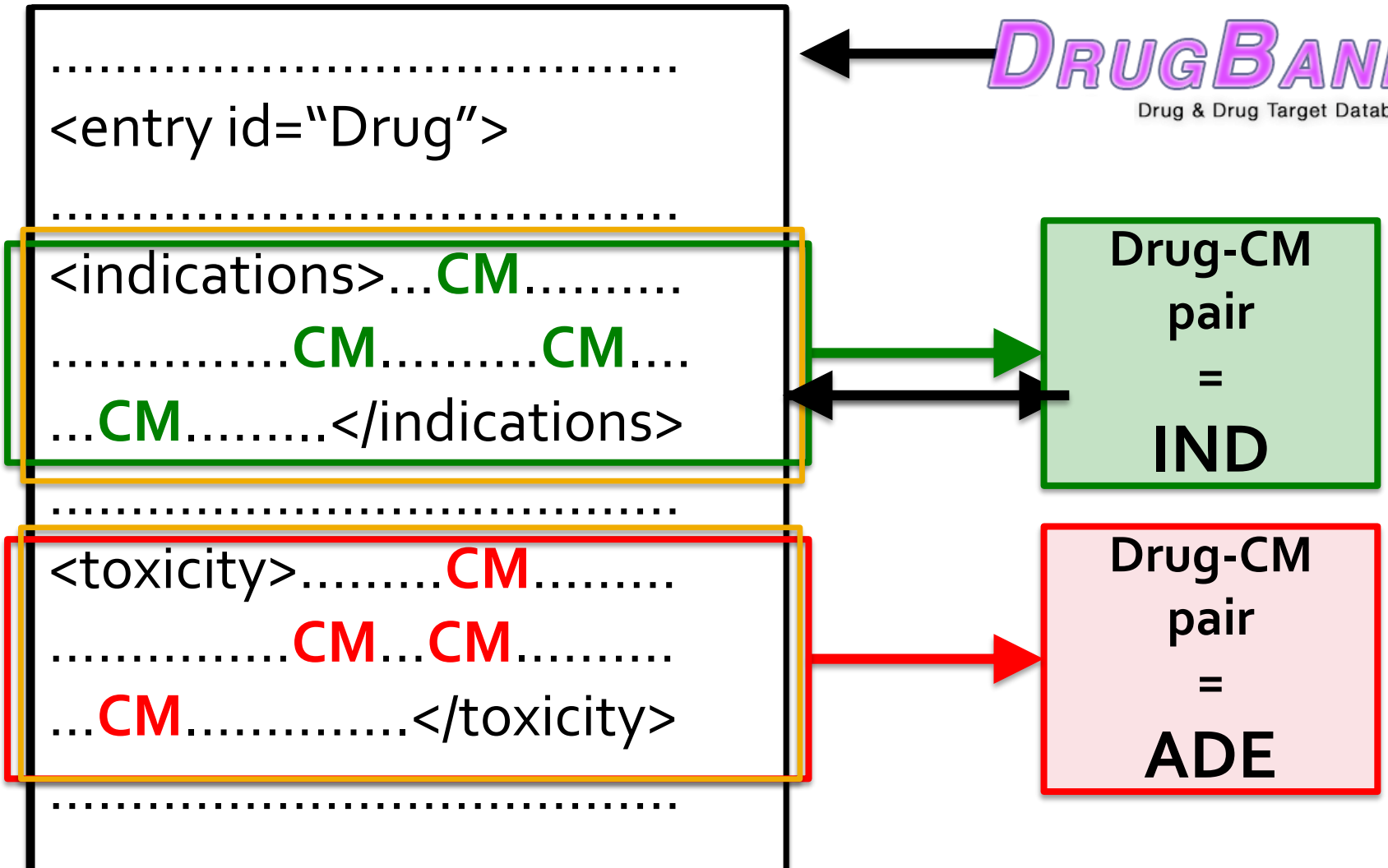


DEB2: From DrugBank

- **DrugBank.ca** is a manually curated database combining pharmacological and pharmaceutical and chemical data with drug target information.
- It includes **Indication** data manually curated from FDA, PubMed, KEGG, TTD, etc.
- It includes **ADE** and toxicity data manually curated from FDA, ToxNet, ASHP, etc.

DRUGBANK
Drug & Drug Target Database

DEB2: From DrugBank



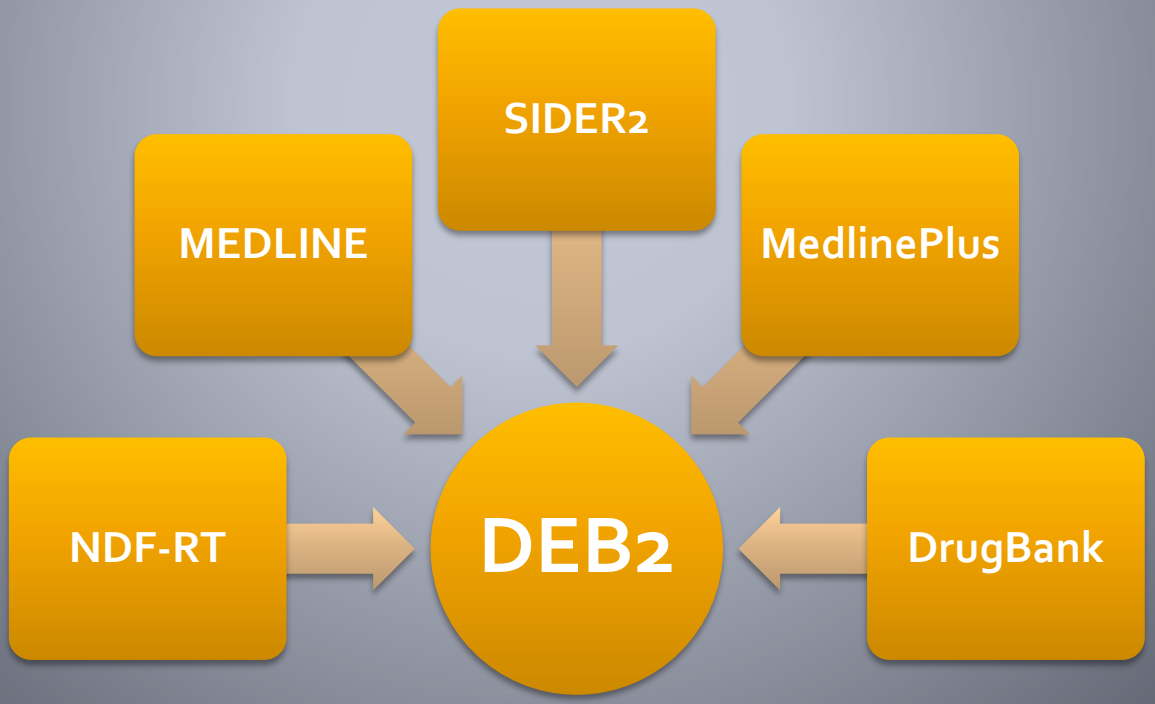
DEB2: From DrugBank

RxNorm Subset	
Drugs	1844
DrugBank	
Found	972

CM Concepts	
Indications	930
Side Effects	594

Total Pairs	
IND	3369
ADEs	4788
Total	8157

DEB2 Results



DEB2: Results

- DEB drug-CM pairs extracted from the 5 sources
- **Full Results**
  - Unique pairs: 138,418
  - Indications: 33,232
  - ADEs: 103,259
  - Unique drugs: 1556
  - Unique CMs: 5721
- **In at least 2 sources**
  - Unique pairs: 18068
  - Indications: 6451
  - ADEs: 11617
  - Unique drugs: 1163
  - Unique CMs: 1606

DEB2: Results

- Full results (from all sources)

	NDF-RT	MEDLINE	MedlinePlus	DrugBank	SIDER2	Unique Total
IND	4055	22732	5325	3369	9646	33232
ADE	78	6331	23444	4788	83956	103259
Total	4133	29063	28769	8157	93602	**138418**

(1927 ties)

- DEB2+ (appearing in at least 2 sources)

	NDF-RT	MEDLINE	MedlinePlus	DrugBank	SIDER2	Unique Total
IND	2684	5221	2293	1984	4217	6451
ADE	20	2846	8910	2042	11190	11617
Total	2704	8067	11203	4026	15407	**18068**

DEB2: Results

- DEB drug-CM pairs in a given source present in at another source:
  - **SIDER2 18%**
  - **MEDLINE 32%**
  - **MedlinePlus 41%**
  - **DrugBank 51%**
  - **NDF-RT 67%**
- Percentage of 18,068 DEB drug-CM pairs from multiple sources
  - **5 sources 1.2%**
  - **4 sources 4.3%**
  - **3 sources 17%**
  - **2 sources 78%**

DEB2: Preliminary Results

- Percentage of DEB drug-CM pairs in a given source agreeing with the consensus (IND/ADE) of the other sources (when present):

- Full Results *(ties included)*
  - MEDLINE 84.7%
  - SIDER2 88.3%
  - MedlinePlus 93.0%
  - DrugBank 95.8%
  - NDF-RT 96.8%

- DEB2+ *(ties excluded)*
  - MEDLINE 97.8%
  - SIDER2 98.5%
  - MedlinePlus 98.8%
  - NDF-RT 99.5%
  - DrugBank 99.6%

Evaluation

- Six physicians reviewed a random sample from DEB2 to estimate DEB2 validity.
  - 600 total pairs reviewed (half IND, half ADE)
  - Each reviewer reviewed 200 pairs
  - Each pair reviewed by two different reviewers
  - Disagreements decided by adjudication

Evaluation Results

- Based on the review, DEB2 is **86%** accurate overall, with indications slightly more accurate and ADEs slightly less accurate.

Overall	Percent	95% Confidence Interval
True	86%	(83%, 89%)
Indications	**Percent**	**95% Confidence Interval**
True	88%	(84%, 92%)
ADEs	**Percent**	**95% Confidence Interval**
True	84%	(81%, 87%)

Evaluation Results
(stratified by number of sources)

INDICATIONS by number of sources				
Sources	Count	TRUE	Percent TRUE	95% CI
2	140	110	**79%**	(72%, 86%)
3	60	55	**92%**	(85%, 99%)
4	50	48	**96%**	(89%, 100%)
5	50	49	**98%**	(87%, 100%)

ADEs by number of sources				
Sources	Count	TRUE	Percent TRUE	95% CI
2	180	151	**84%**	(79%, 89%)
3	70	58	**83%**	(74%, 92%)
4	50	47	**94%**	(79%, 100%)