



Sivaram Arabandi, MD  
National Library of Medicine

May 2011



**A KNOWLEDGE INTENSIVE APPROACH  
FOR IDENTIFYING ADVERSE DRUG  
EVENTS IN CLINICAL TEXT**

# motivation

- In any given week
  - 4 out of 5 people will use medications
  - 1 in 3 will take 4 or more different medications
- At least 1.5 million preventable ADEs per year  
[IOM 2006 report estimate]
- Each ADE added about \$8,750 (in 2006 dollars) to the cost of the hospital stay
- Estimated at \$3.5 billion per year
- Only 10 to 20 percent of errors are reported  
[IHI 2008 report]

ADE = Adverse Drug Event  
IOM = Institute of Medicine  
IHI = Institute for Healthcare Improvement



# how can ADEs be prevented?

- Acknowledge
  - Recognize
  - Understand
  - Address
- 



# recognize

"Because health care data are often narrative, natural language processing (NLP) is another important technique for mining data for quality improvement and patient safety purposes"

- Institute of Medicine Report on Patient Safety:  
*Achieving a New Standard of Care, 2003*
- 

# NLP in medicine

- 1980 – 90s:
  - Linguistic String Project-Medical Language Processor (LSP-MLP)
  - MedLEE
    - initially developed for Radiology reports
    - still in production (Columbia)
  - Metamap
  - and many more ....

# NLP for ADE detection

- Melton & Hripcsak (2005):
  - MedLEE to identify 45 NYPORTS event types using discharge summaries
  - Sensitivity = 0.28
  - Specificity = 0.985
- Wang X, Hripcsak G, Markatou M, Friedman C (2009)
  - Pharmacovigilance using MedLEE
  - Recall = 0.75
  - Precision = 0.31

NYPORTS = New York Patient Occurrence Reporting & Tracking System

# background

- “current applications, however, are rarely applied outside of the laboratories they have been developed in, mostly because of scalability and generalizability issues”
  - Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. *Extracting information from textual documents in the electronic health record: a review of recent research*, MedInfo 2008
- “with few exceptions, we do not know which of the reviewed NLP-CDS systems are actually implemented or deployed, and what makes these systems worthwhile.”
  - Demner-Fushman D, Chapman WW, McDonald CJ. *What can natural language processing do for clinical decision support?*, J Biomed Inform. 2009



goal

“What will it take to **extract information** from narrative clinical text  
– **reliably** and **generally**”





What is missing?  
What is holding us back?





# investigate

- Can we
    - **Exploit** capabilities of existing systems?
    - **Augment** them with additional knowledge?
- 

# use case

- Adverse Drug Events [ADEs]

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

-- World Health Organization



# identify

- Concepts & relations
    - Drugs, findings
  - Normal vs. Abnormal
  - Timeline (Events)
  - ADEs “events of harm”
- 

# approach

- **Exploit** existing resources (UMLS, MedLEE, MetaMap)
  - Concepts & relations (drugs and finding)
- **Enhance** with rules and knowledge
  - Normal vs. Abnormal
  - Timeline (Events)
  - Identify “events of harm (ADEs)”



# materials

- Data
  - Admission notes: ER note, H&P report, ....
  - Progress notes
  - Procedure reports: Lab, Radiology, ....
  - Discharge reports
- Sources
  - TREC reports
  - Clinical Center reports



TREC = Text REtrieval Conference



Concepts  
Manifestations  
Concept modification



# concepts

- Identification of Terms
  - observables, diagnosis, lab tests, procedures
  - medications
- 20+ years of research ....

A 76-year-old **male** with a 12-year history of **idiopathic Parkinson's disease (PD)** had been taking **levodopa** (L-dopa; 500 mg/day) and **pramipexole** (4 mg/day) for the past several years with relatively good control of his **motor symptoms**, primarily **tremor** and **rigidity**.



# concepts: challenges

- Expressiveness of language
    - Ambiguity
    - Synonymy
    - Negation
    - Abbreviations
- 

# concepts: ambiguity

- **One term** that can mean several things
  - **Cold** = feeling cold, cold infection, chronic obstructive lung disease
  - **MS** = Multiple sclerosis, Mitral stenosis

“No recent **catheterizations**”

Search Results (15)

C0007430 Catheterization  
C0004704 Balloon Dilatation  
C0006259 Catheterization, Bronchial  
C0007435 Catheterization, Central Venous  
C0007436 Catheterization, Peripheral  
C0018795 Cardiac Catheterization Procedures  
C0041953 Catheterization, Ureteral  
C0042019 Urinary Catheterization  
.....

“Patient's last **functional evaluation** approximately one year ago, with no reversibility / ischemia seen”

Search Results (5)

C0278372 Functional assessment  
C2732531 Functional capacity evaluation scale  
C0200299 Special audiologic evaluation for functional hearing loss  
C0203889 Physical medicine evaluation of functional activities, initial 30 minutes  
C1971412 Comprehensive preoperative assessment performed for cataract surgery with intraocular lens

# concepts: synonymy

- **Two or more terms** that mean the same thing
  - Myocardial infarction = MI, Heart Attack
  - Parkinson disease = Parkinson's disease, PD,  
Paralysis agitans, ...  
BUT not Parkinsonism

# negation

- Negation expressed in **many ways**

- **No**

- *no history of cough*
- *no evidence of tenderness*

- Patient **denies** ...

- *denies* any headache.
- *denies* any upper abdominal pain. He states he is very hungry right now. He has only lower abdominal cramps.

MetaMap with Negex

```
NEGATIONS:  
Negation Type: nega  
Negation Trigger: no  
Negation PosInfo: 0/2  
Negated Concept: C1961131:Cough  
Concept PosInfo: 14/5
```

```
NEGATIONS:  
Negation Type: nega  
Negation Trigger: denies  
Negation PosInfo: 0/6  
Negated Concept: C2096315:headache  
Concept PosInfo: 11/8
```

# concepts: abbreviations

- **Acronyms**

- Denies fevers/chills, **URI** sx
- 76 yoF with hx **MI** '96, s/p **CVA**, hx **GERD**

- **Shortcuts**

- 76 **yoF** with **hx** MI '96, **s/p** CVA, **hx** GERD
- 56 yo F w/restless leg **synd adm** for **eval**. No eating **probl/GI** **sx** **ID** 'ed. Wt 72.6kg, BMI=31.4. labs as yet **u/a**

- **??**

- **Lymphs**                      11 %

# findings

- Understanding at the **data level**
- **No interpretation** of normal or abnormal
- **Continuous** data → **Numeric values** and **Units**
  - e.g.: Temp. = **98.6° F**
  - Fahrenheit ↔ Centigrade
  - Meter ↔ Centimeter
  - Meter ↔ Feet

- Fizman et. al. *A Knowledge Intensive Approach to Mapping Clinical Narrative to LOINC*

- **Categorical** data  
e.g.: swollen knee



## use case

A 76-year-old **male** with a 12-year history of **idiopathic Parkinson's disease (PD)** had been taking **levodopa** (L-dopa; 500 mg/day) and **pramipexole** (4 mg/day) for the past several years with relatively good control of his **motor symptoms**, primarily **tremor** and **rigidity**.

# MedLEE output

A 76-year-old male with a 12-year history of

```
finding:demo
age>> [76,[idref,6],year,[idref,8]]
parsemode>> mode2
sectname>> report past history item
sex>> male
  idref>> 12
sid>> 1
```

```
finding:demo
age>> [12,[idref,18],year,[idref,20]]
parsemode>> mode5
sectname>> report past history item
sid>> 1
```

## idiopathic Parkinson's disease

```
problem:Parkinson's disease
certainty>> high certainty
  idref>> 42
descriptor>> idiopathic
  idref>> 26
  idref>> 28
parsemode>> mode4
```

```
sectname>> report past history item
sid>> 1
status>> history
  idref>> 22
code>> UMLS:C0030567_Parkinson
Disease
  idref>> [28]
```

# MetaMap output

A 76-year-old male with a 12-year history of

660 year [Temporal Concept]

827 History of previous events [Finding]

idiopathic Parkinson's disease



# next: enhanced processing

- Identify
    - Abnormal findings
    - Sequence of events
    - Identify ADE
  - Processing approach
    - Knowledge intensive
    - Rule based
- 



Abnormal manifestation

# *abnormal* manifestation

- Bodily feature
  - **Physiological** vs. **Pathological**
- Vitals: temperature
  - temp = 98.6° F → **NORMAL**
  - temp = 102.6° F → **ABNORMAL**
- Sign:
  - red, swollen knee → **ABNORMAL**

# *abnormal* manifestation

- Weight gain
  - Hypothyroidism → **ABNORMAL**
  - Pregnancy → **NORMAL**
- Pulse rate = 110/min.
  - In Adult *resting* → **ABNORMAL**
  - In Adult *exercising* → **NORMAL**
  - In *Child* resting → **NORMAL**

# knowledge sources

- UMLS

- Metathesaurus

- Concepts classified into Semantic Types

- Additional models

- Reference ranges

- LOINC?

- Wikipedia? [<http://en.wikipedia.org/wiki/Pulse>]

Parkinson Disease [Disease or Syndrome]  
Cough (Coughing) [Sign or Symptom]

Pulse rate

newborn	1 – 12 months	1 – 2 years	2 – 6 years	6 – 12 years	12 years - adults	adult athletes
120 - 160	80 - 140	80 - 130	75 - 120	75 - 110	60 - 100	40 - 70



# challenges

- Prospects for success? Feasibility?
  - Knowledge is available
  - Needs processing
- Approach
  - Develop rules
  - Limit in some principled way
    - Adults only or a few drugs, findings, disease



# Temporality

# temporal relations

- Basic ordering of events

“did the **symptom** precede the **medication** or follow it?”

Med	Dose	Status
Insulin	75/25 forty-five units in the morning and 30 units in the evening	Current
Cozaar	50 mg daily	Current
Lipitor	10 mg nightly	Current

CPOE

H & P report  
Progress note  
Nursing note  
Discharge summary

Date: 10 Aug 2008

50-year-old Caucasian male who presenting with severe nausea which started about 2 days ago. Since then he has been having about 20 episodes of vomiting. The patient also complains of some chills and fever and complains of postnasal drip. The patient has not moved his bowels in the last 2 days but he moved bowels today.

PAST MEDICAL HISTORY: Significant for

1. Diabetes mellitus diagnosed at the age of 40.
2. History of DKA requiring admission to ICU in July 2006.
3. Dyslipidemia.
4. Hypertension.

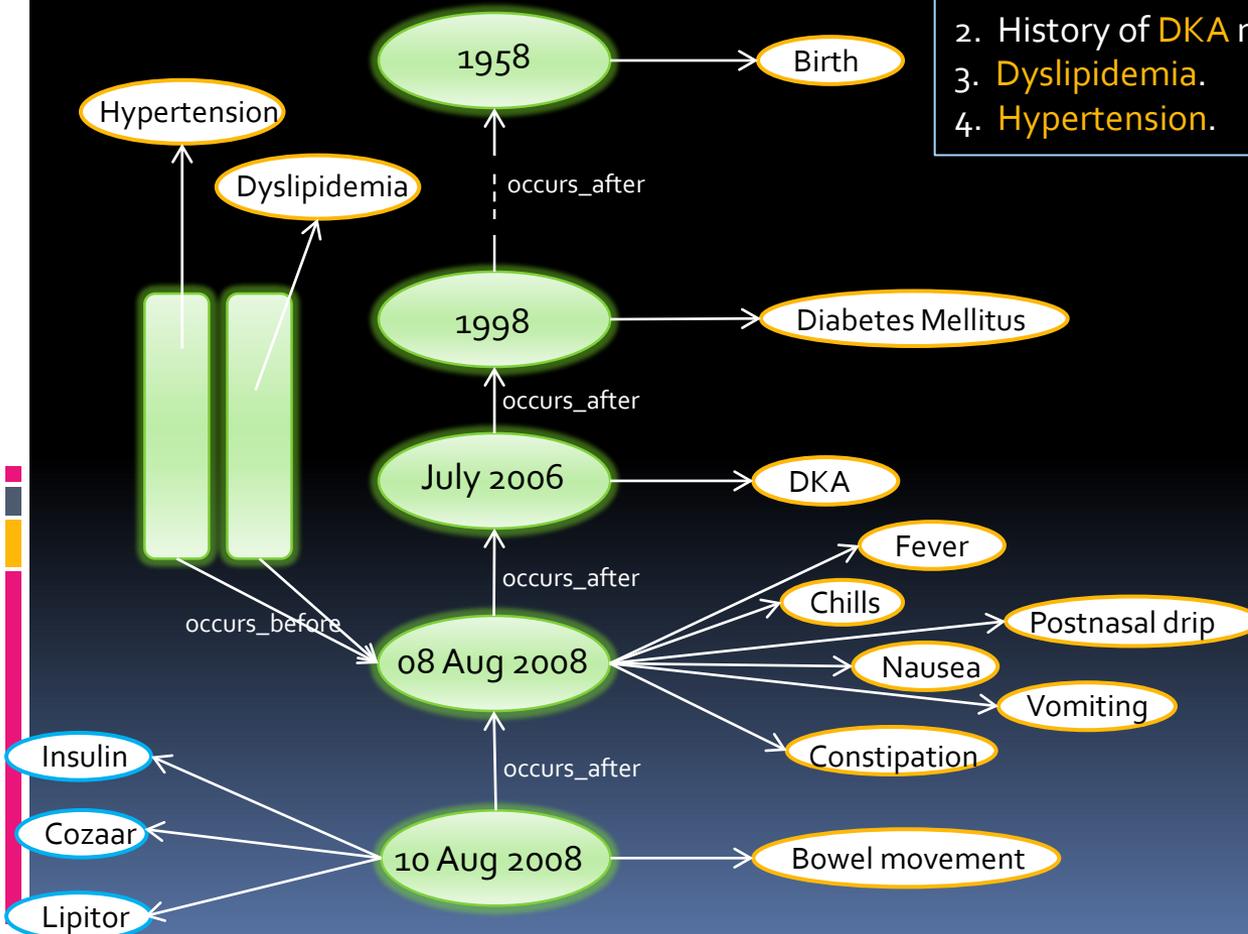
Med	Dose	Status
Insulin	75/25 forty-five units in the morning and 30 units in the evening	Current
Cozaar	50 mg daily	Current
Lipitor	10 mg nightly	Current

Date: 10 Aug 2008

50-year-old Caucasian male who presenting with severe **nausea** which started about 2 days ago. Since then he has been having about 20 episodes of **vomiting**. The patient also complains of some **chills** and **fever** and complains of **postnasal drip**. The patient has **not moved his bowels** in the last 2 days but he **moved bowels** today.

PAST MEDICAL HISTORY: Significant for

1. **Diabetes mellitus** diagnosed at the age of 40.
2. History of **DKA** requiring admission to ICU in July 2006.
3. **Dyslipidemia**.
4. **Hypertension**.



# knowledge sources

- UMLS
- Additional models of the world
- Reference ranges
- **Time models**
  - Time Ontology in OWL  
<http://www.w3.org/TR/owl-time/>
  - Clinical Narrative Temporal Relation Ontology [CNTRO]  
- Tao C, Wei WQ, Solbrig HR, Savova G, Chute CG. CNTRO: A Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives, 2010

# processing events

- Events

- Patient was admitted on July 24, 2007 10:27 am
- Lymphoma was diagnosed in August 2003
- Patient almost drowned in the summer of 2001
- Patient broke his right arm in the mid-seventies

- Relationships between events

- Patient started coughing 2 days after fever started
- Patient has a headache for the last 4 days

# processing events

- CNTRO
  - **Terms:** Event, Time, Duration, Granularity, Precision and TemporalRelationStatement.
  - **Relations:** equal, before, after, meet, overlap, contain, during, start, and finish

Lymphoma was diagnosed in August 2003

```
<event1> rdf:type Event;  
  rdfs:label "Lymphoma was diagnosed";  
  hasTimeStamp <tInst1>;  
<tInst1> rdf:type TimeInstant;  
  hasOrigTime "August 2003";  
  hasGranularity "month";
```



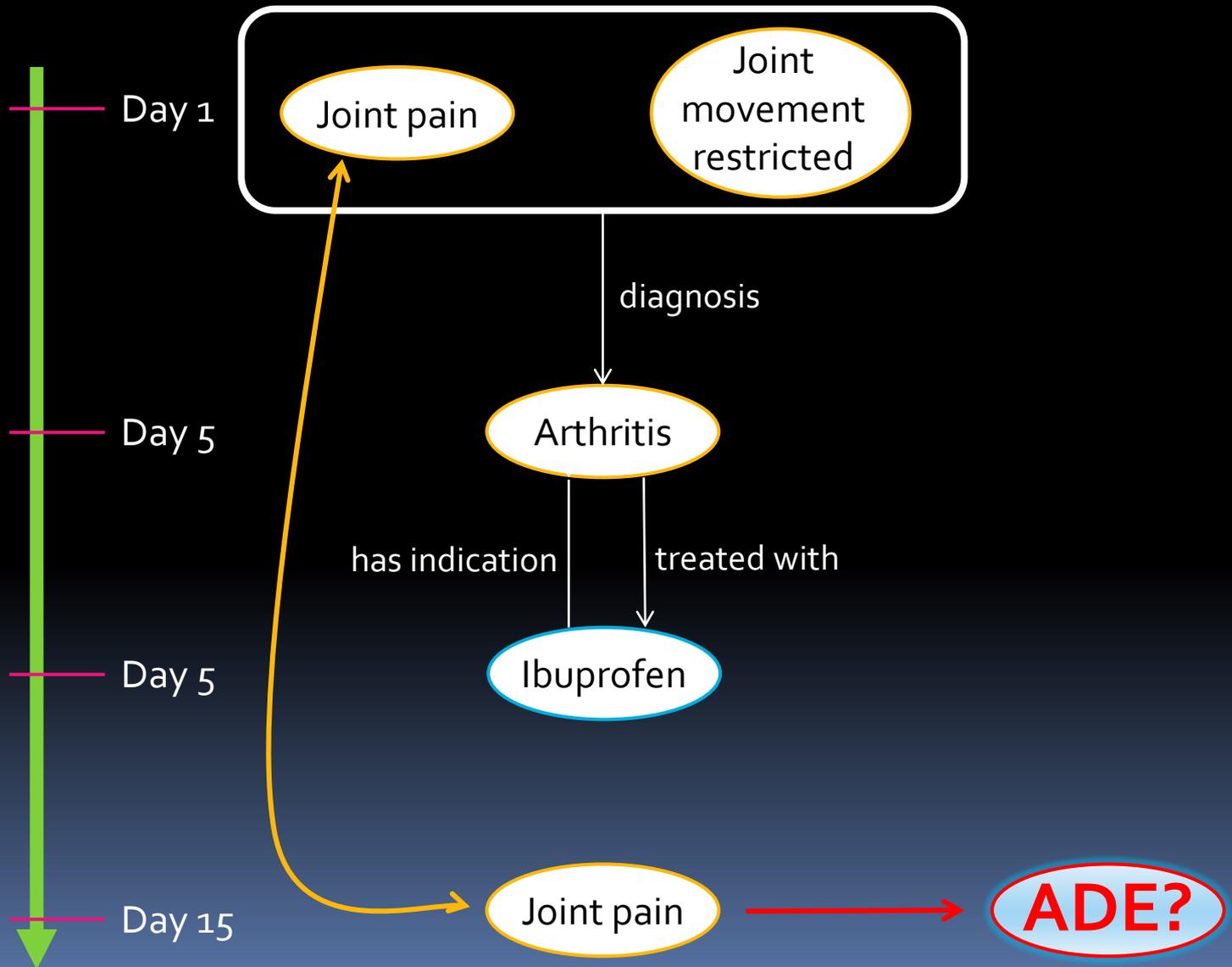
# Adverse drug events



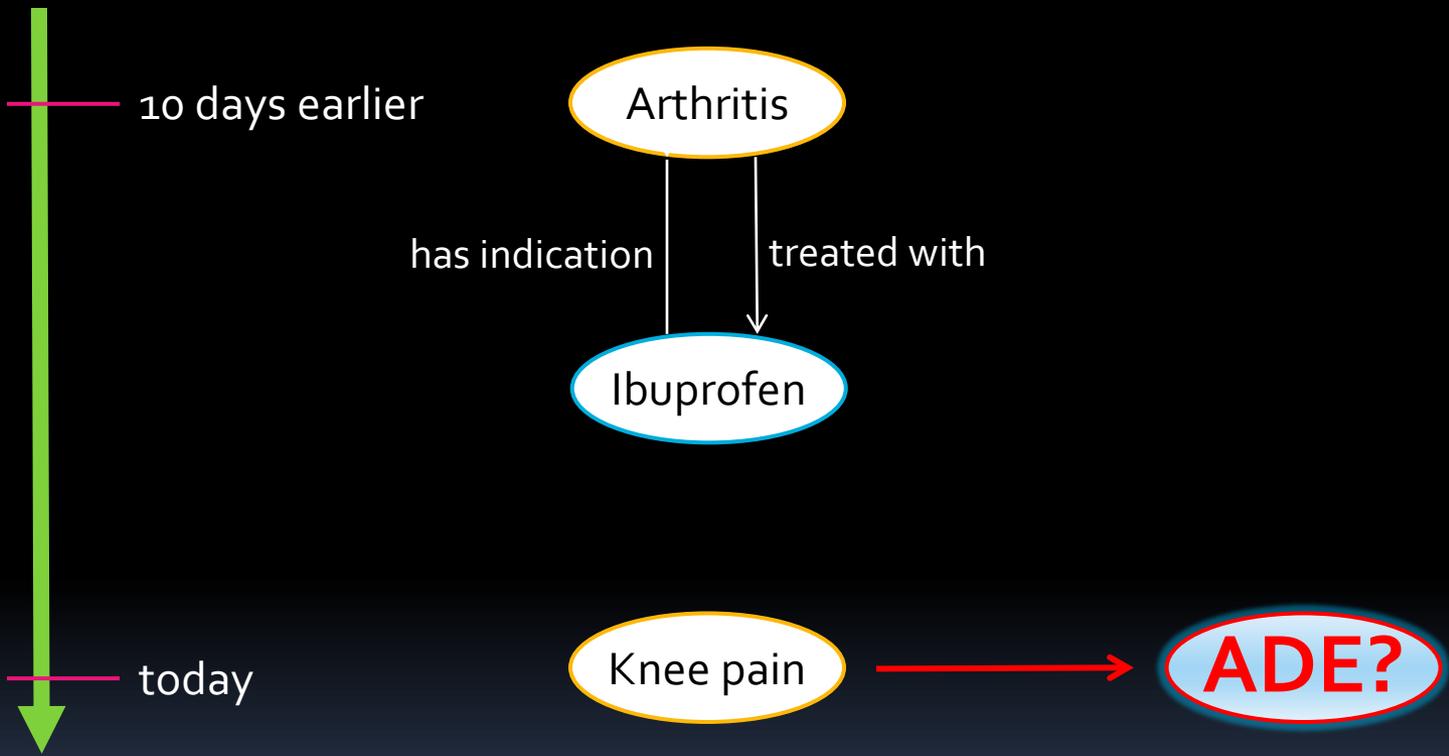
# adverse drug event

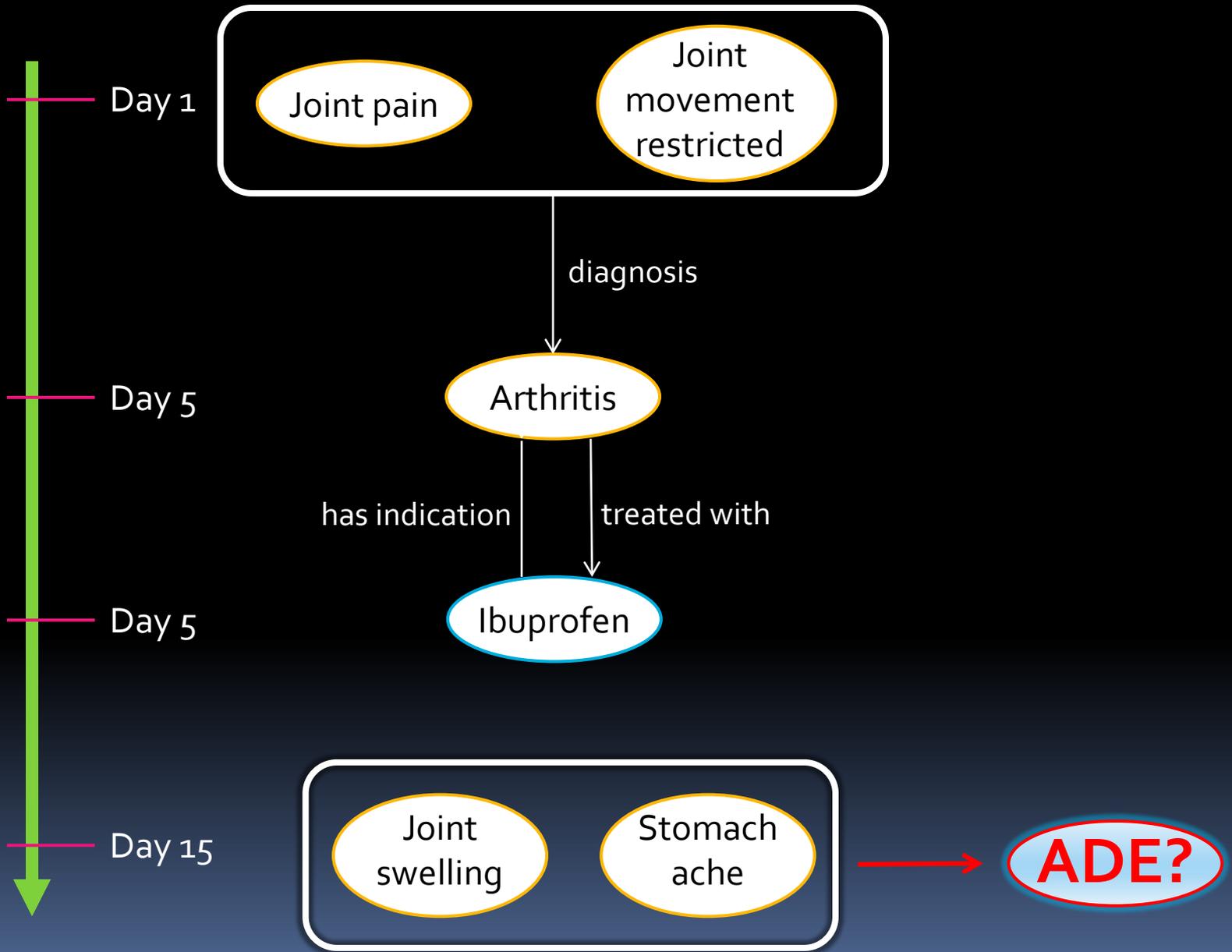
- **Abnormal finding** is **NOT** same as '**Harm**'
- The patient must have **suffered actual harm**  
e.g.: in a patient on **Coumadin**, simply a **high INR** is not a ADE unless excessive bleeding occurs (a **hematoma**).

# Use case



50 yr. old male patient  
presenting with knee pain.  
10 days earlier, patient was  
diagnosed with Arthritis and  
put on Ibuprofen 400 mg bid





KNOW so far

- Diagnosis
- Medication name
- Event time line

Knowledge about Diseases

Signs &  
Symptoms  
Lab findings  
Radiology  
etc.

Arthritis

has indication

treated with

Ibuprofen

*DON'T KNOW* about

- Disease  $\leftrightarrow$  Manifestations
- Drug information

Knowledge about Drugs

Indications  
Contra-indications  
Adverse events  
Drug interactions

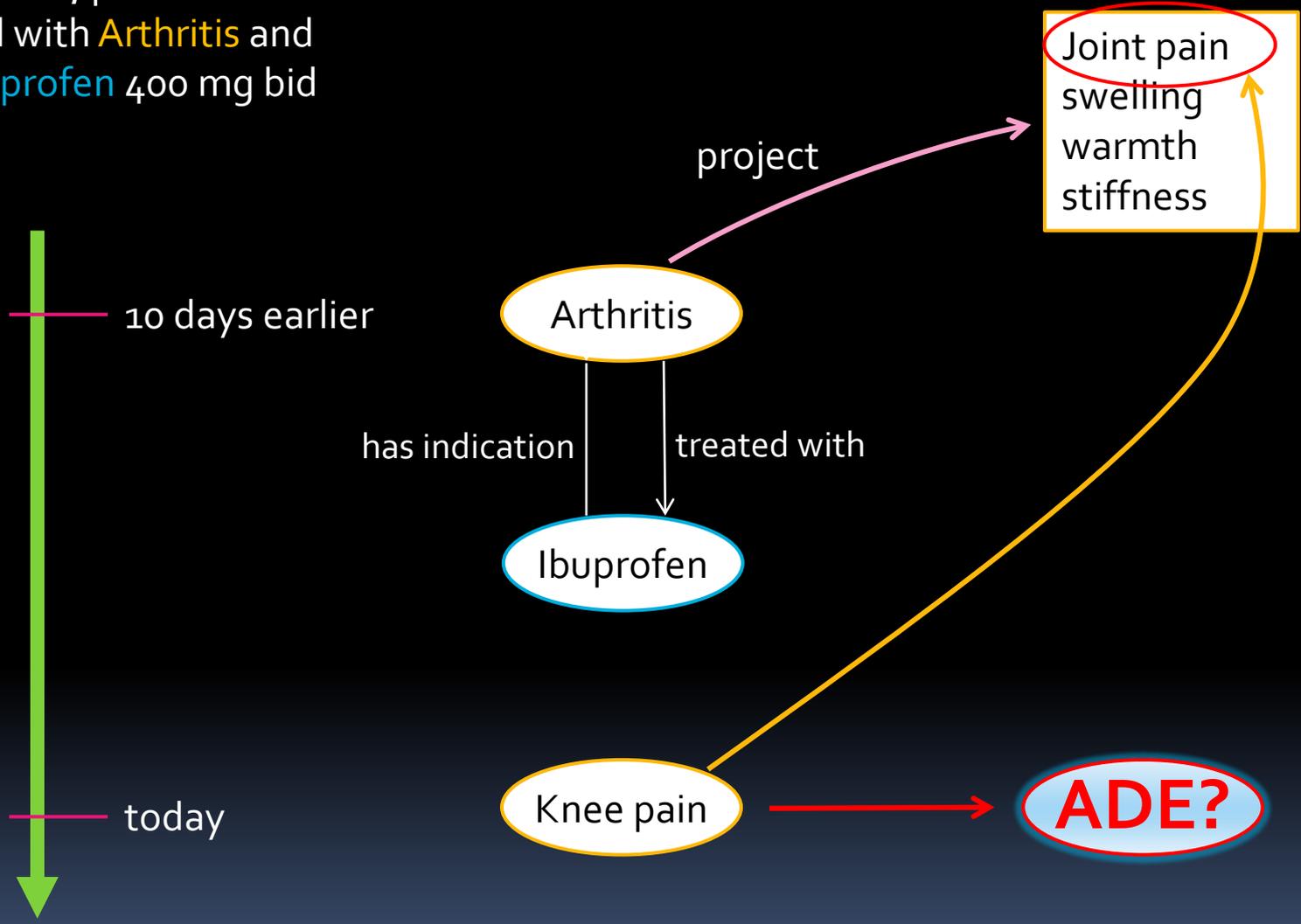
# disease – manifestations

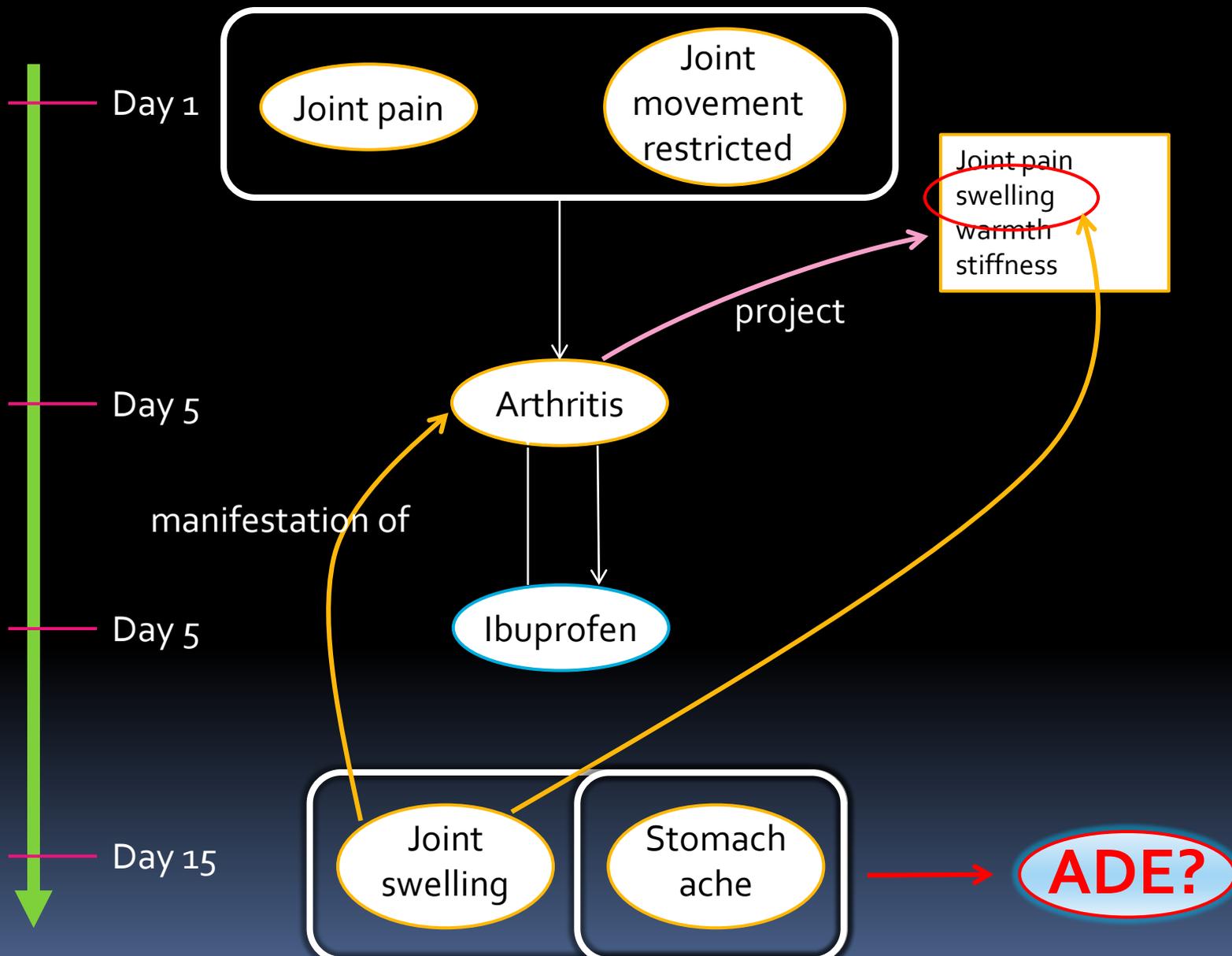
- **Rheumatoid Arthritis:**
  - **Symptoms & Signs:** Joint pain, swelling, warmth, stiffness, extremity weakness, deformities, ....
  - **X-ray:** joint effusion, narrowing of inter-articular space, decalcification, ....

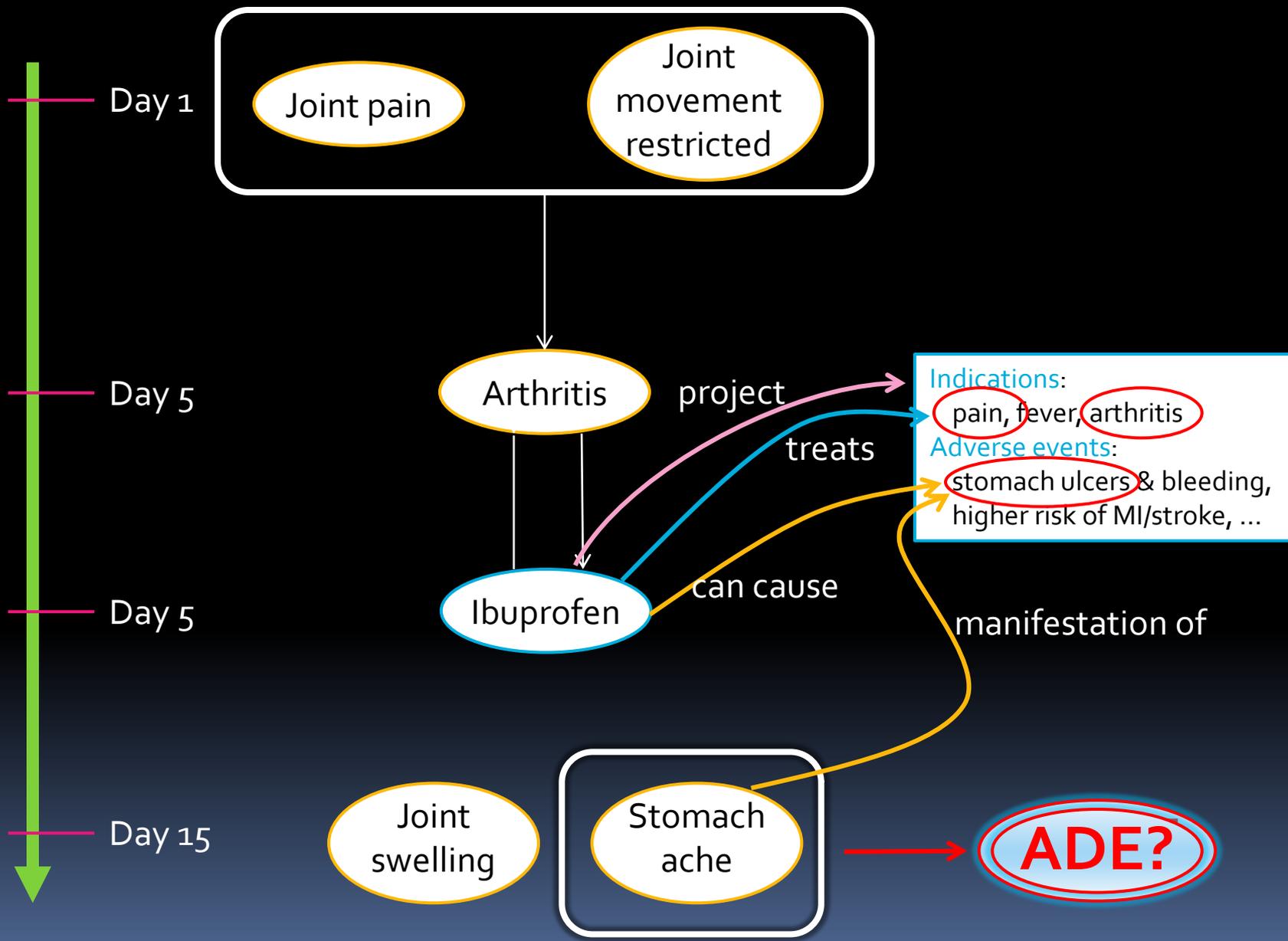
# medications

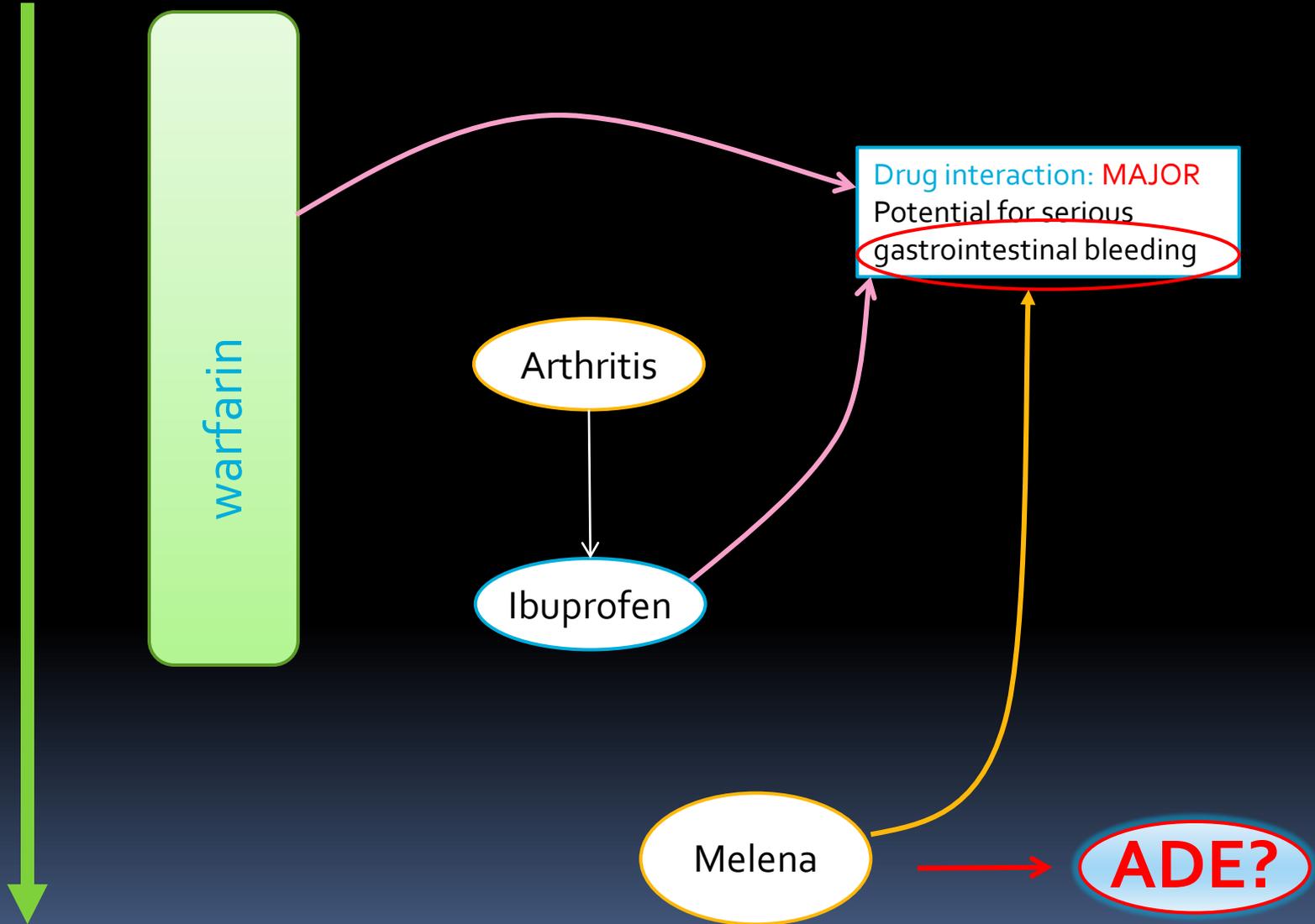
- **Ibuprofen:**
  - **Indications:** pain, fever, arthritis, bursitis, gout, ...
  - **Contra-indications:** asthma, pregnancy, drug allergy, ...
  - **Drug interactions:** aspirin, dicumarol, ketoralac, ...
  - **Adverse events:** stomach ulcers/bleeding, higher risk of MI/stroke, ...

50 yr. old male patient  
presenting with knee pain.  
10 days earlier, patient was  
diagnosed with Arthritis and  
put on Ibuprofen 400 mg bid









warfarin

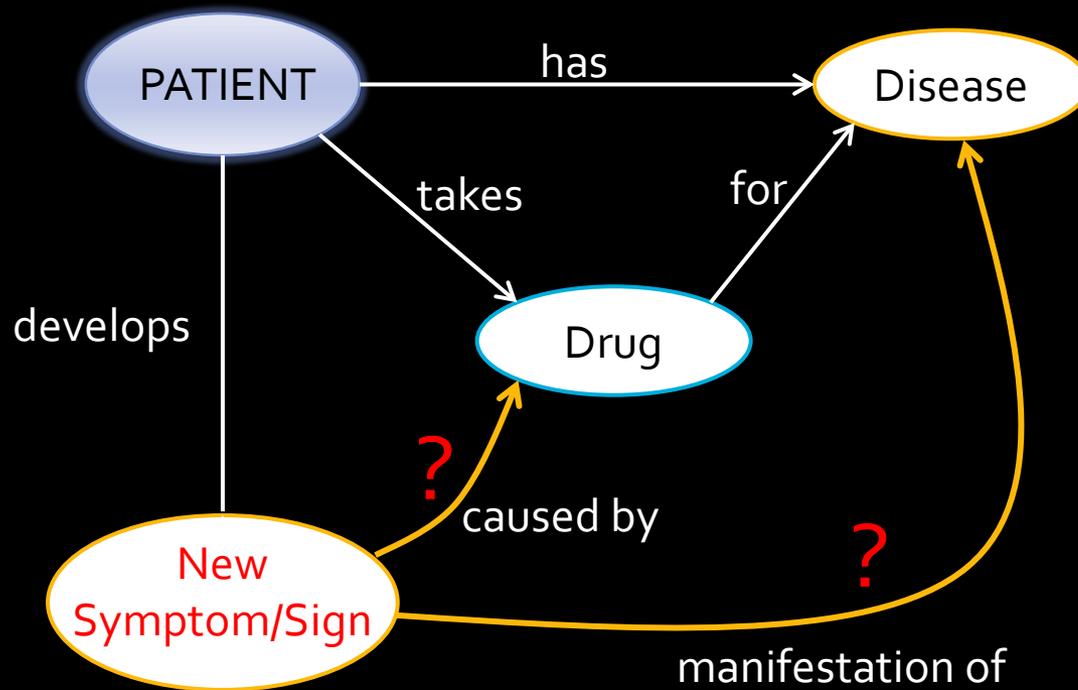
Drug interaction: MAJOR  
Potential for serious  
gastrointestinal bleeding

Arthritis

Ibuprofen

Melena

ADE?



## RULE

IF finding follows drug  
AND finding does not precede drug  
AND finding is not a manifestation of the disease  
AND finding is a known adverse effect of the drug

-> THEN, this is a potential ADE

Where is the **knowledge** to solve this?



Knowledge resources

# knowledge resources

- NEED:
  - Disease → manifestations
  - Drugs → indications/contra-indications, adverse effects
  - Drug – Drug interactions
- Traditional: books, articles, etc.
- Computable sources – some with additional processing
  - Drugs:
    - RxNorm/NDF-RT
    - MedlinePlus, DailyMed, FDA
    - Research databases - ADESSA
  - Clinical Decision Support Systems (*think MYCIN*):
    - Dxplain<sup>®</sup>
    - Quick Medical Reference (QMR)
    - Isabel
    - DiagnosisPro
  - Web pages

# RxNorm/NDF-RT

- Normalized names for medications
- Maps to many drug vocabularies
- Contains brand names, dose forms
- Mechanism of action, Pharmacokinetics
- Indications & Contra-indications
- Drug interactions
- Adverse events???

Kind	Subject	Predicate	
Drug	IBUPROFEN	isa	I [Preparations]
Disease	IBUPROFEN	CI_with	Asthma
Disease	IBUPROFEN	CI_with	Bronchial Hyperreactivity
Disease	IBUPROFEN	CI_with	Drug Hypersensitivity
Disease	IBUPROFEN	CI_with	Pregnancy
Disease	IBUPROFEN	CI_with	Pregnancy Third Trimester
Disease	IBUPROFEN	CI_with	Pregnancy, Abdominal
Disease	IBUPROFEN	CI_with	Rhinitis
Disease	IBUPROFEN	may_prevent	Pain
Disease	IBUPROFEN	may_treat	Arthritis, Juvenile Rheumatoid
Disease	IBUPROFEN	may_treat	Arthritis, Rheumatoid
Disease	IBUPROFEN	may_treat	Bursitis

# MedlinePlus / DailyMed

- NLM websites
- Wealth of drug related information

## Ibuprofen

(eye byoo' proe fen)



[Why is this medication prescribed?](#)

[How should this medicine be used?](#)

[Other uses for this medicine](#)

[What special precautions should I follow?](#)

[What special dietary instructions should I follow?](#)

[What should I do if I forget a dose?](#)

[What side effects can this medication cause?](#)

[What storage conditions are needed for this medicine?](#)

[In case of emergency/overdose](#)

[What other information should I know?](#)

[Brand names](#)

[Brand names of combination products](#)

# D<sub>x</sub> *plain*<sup>®</sup>

- Disease → list of symptoms
- Set of symptoms → differential diagnosis

## Evidence of "ARTHRITIS, RHEUMATOID" (COMMON)

The following clinical manifestations (if present) would support this disease:

[subcutaneous nodule](#)  
[uveitis, bilateral](#)  
[rheumatoid vasculitis](#)  
[Baker cyst](#)  
[carpal tunnel syndrome](#)  
[joint dislocation](#)  
[uveitis](#)  
[fever unknown origin](#)  
[migratory polyarthritis](#)  
[extremity deformity](#)  
[extremity deformity, upper](#)  
[finger deviation, ulnar](#)  
[finger swelling](#)  
[joint effusion](#)  
[joint warmth](#)  
[joint induration](#)  
[knuckle swelling](#)  
[muscle atrophy](#)

# Adverse Drug Event [ADE]

- Actual harm
  - Have suffered an ADE
  - Have potentially suffered an ADE
- Potential for harm
  - Can suffer a known ADE from the medications
  - Can suffer an ADE based on LBD information

LBD = Literature Based Discovery (e.g. Semantic Medline)

# ADE detection: SemRep

- Rofecoxib (Vioxx) – NSAID and a selective COX-2 inhibitor
- Introduced: 1999
- Withdrawn: 2004
  - increased risk of MI & stroke

Could we have detected this using SemRep?

## Citations

Thromboxane A<sub>2</sub>

Select 11595413

[PMID:11595413](#)

Date of Publication: 2001 Jan-Dec

Title: The ever-emerging anti-inflammatories. Have there been any real advances?

### Abstract:

Gastrointestinal (GI) Adverse Drug Reactions (ADRs) from the NSAIDs are a major cause of morbidity and mortality in arthritic patients taking these drugs. The recent much heralded development of COX-2 selective drugs (celecoxib, rofecoxib), the objective of which has been to spare inhibition of the production of COX-1 derived mucosal protective prostaglandins, may have represented an advance in reducing the risk of serious ADRs--ulcers and bleeding--but does not appear to have reduced the incidence of symptomatic side-effects (nausea, vomiting, epigastric pain/heartburn, abdominal discomfort) which are a major reason for withdrawal from NSAID therapy, especially in the long term. The rationale of COX-2 selectivity from these newer drugs is controversial since there may be pharmacokinetic differences from established carboxylate-NSAIDs that accounts for their apparent lower ulcerogenicity. Moreover, concerns have been recently expressed that as COX-2 is important in ulcer healing, control of prostacyclin production and renal function that they may have adverse reactions from these effects.

**Indeed, recent reports of enhanced risk of congestive heart failure with rofecoxib are of importance and may relate to impaired prostacyclin production.**

Moreover, there are other therapeutic strategies that have yielded equally low ulcerogenic NSAIDs (e.g. the prodrug, nabumetone; the established COX-2 inhibitory drug, nimesulide) and even the well-established NSAIDs ibuprofen and diclofenac have relatively low upper GI ulcerogenicity and have been used as benchmark standards in comparative trials of the

tinal Hemor  
Glomerula



# conclusions

- Existing work – identify concepts & relations
- Enhancements
  - Abnormal manifestations
  - Timelines
- Additional Knowledge and Rules needed



# acknowledgment

- Olivier Bodenreider
  - Tom Rindflesch
  - Marcelo Fiszman
  - Jim Cimino
  - Lan Aronson
  - Clem McDonald
  - May Cheh
- 