Detecting Side-effects via Electronic Medical Records with a case study in the Correlation of Parkinson's Disease with the use of Lipophilic Beta-Blockers

He Zhang

Assistant Professor
Dept. of Information Systems & Decision Sciences,
Muma College of Business, University of South Florida, Tampa, FL 33620.

November 08, 2017

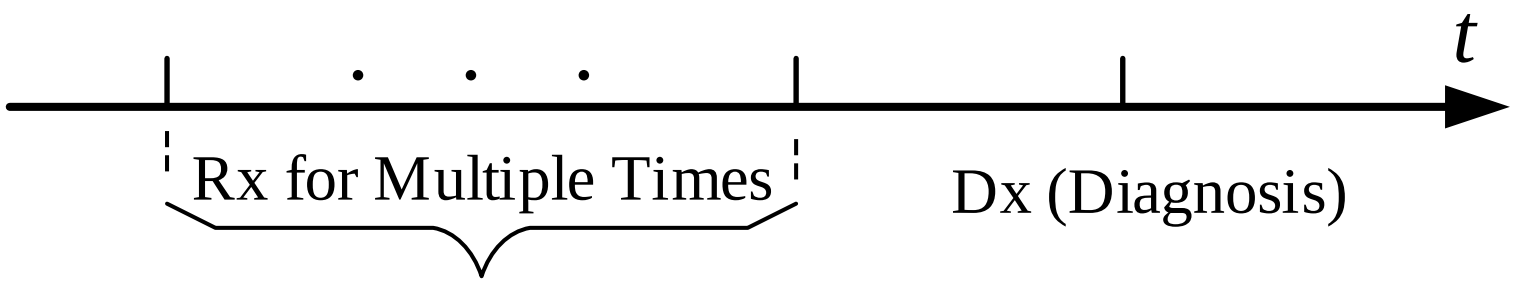
Electronic Medical Records and Pharmacovigilance

- EMRs are a source of information regarding patients' medical history, diagnoses and medications
- Pharmacovigilance (PV or PhV), also known as drug safety, is the pharmacological science relating to the collection, detection, assessment, monitoring, and prevention of adverse effects with pharmaceutical products.
- Existing research in PhV
  - Traditional research depends on the reporting database.
  - Current research uses EHR to detect instant side-effect.
- Our method can be used to detect signals for chronic diseases.

Research Dataset

- 10 years of EHR data (2004 to 2014) from HealthAxis Group with 20,777 adult patients.
  - Patient Diagnosis (Dx) & Prescription (Rx) History
  - Doctors Notes and Messages
  - Patient Demographics
- External Data
  - Dx Code Mappings (`http://www.icd9data.com/`).
  - Treatment Rx for Dx Mappings.
- This work depends on both structured and unstructured data.

Methodology for ADR Detection



- We count the occurrence of each pair (Rx, Dx) if Rx is not a treatment drugs for the Dx.
  - Patient Diagnosis (Dx) & Prescription (Rx) History
  - Doctors Notes and Messages
  - Patient Demographics
- Each patient will be considered as a valid case for the pair (Rx,Dx) if
  - Rx has been prescribed at least five times before Dx
  - The minimum time gap between the fifth prescription and diagnosis is 230 days.

Methodology for ADR Detection

- Odds ratios were generated for each (Rx,Dx) pair according to the following contingent table:

		Case: Developed Disease (DX)		
		Yes	*No*	*Totals*
Exposed: Prescribed Medication (RX)	*Yes*	Rx – Dx	Rx – No Dx	Rx
	No	No Rx - Dx	No Rx – No Dx	No Rx
	Totals	Dx	No Dx	Total Patients

- where odds ratio is defined as:

$$\text{Odds Ratio} = \frac{\frac{\text{Rx}-\text{Dx}}{\text{Rx}-\text{No Dx}}}{\frac{\text{No Rx}-\text{Dx}}{\text{No Rx}-\text{No Dx}}}$$

Methodology for ADR Detection

- We reduce the noise in the results by setting the following filters.
  - The minimum number of patients for a (Rx,Dx) pair was set to 4.
  - The minimum number of patients that were prescribed the Rx drug was set to 200 patients.
- Intuitively, if there exists causal relationship between Rx and Dx, the odds ratio will be high.

ADR Validation

- The detected ADRs was analyzed with the help of clinical experts.
- We select the category Central Nervous Systems (ICD9 330 to 337) for further investigation because we have an expert in this domain.
- For strong signals ((Rx,Dx) pair with high odds ratio), we further refine the results by mining the text messages.
- We conduct the following statistical tests:
  - Test the difference between the two groups, exposed and not exposed.
  - Test the population proportion odds.
  - Fisher's exact test.
  - χ^2 test for independence between the Rx and Dx.
  - Confounder analysis, i.e. demographic info.

ADR Detection for Parkinson's Disease

We focus on the Central Nervous Systems (ICD9 330-337), we only have significant signals for the Parkinson's disease (ICD9 332). The top five ADRs for PD are listed as follows

DRUG NAME	ICD9	Y-Y	Y-_	_-Y	TOTAL	N-_	_-N	Y-N	N-Y	N-N	ODDS RATIO
Coumadin	332.00	5	385	76	20777	20392	20701	380	71	20321	**3.76**
Pravachol	332.00	5	387	76	20777	20390	20701	382	71	20319	**3.74**
Metoprolol	332.00	6	486	76	20777	20291	20701	480	70	20221	**3.61**
Toprol	332.00	6	501	76	20777	20276	20701	495	70	20206	**3.50**
Zocor	332.00	6	533	76	20777	20244	20701	527	70	20174	**3.28**

- Coumadin fell into the Blood Thinner category.
- Pravachol and Zocor fell into the Statin category.
- Metoprolol and Toprol fell into the Lipophilic Beta-Blockers category.

ADR Detection for Parkinson's Disease

- The correlation between Statin and Parkinson is controversial.
  - Risk may declines with increasing blood cholesterol level (Simon et al. 2007).
  - Risk may decline with regular use of statin (reduce blook cholesterol level) (Gao et al 2012).
  - Statin use may be associated with a higher PD risk (Huang et al. 2005).
- For Coumadin, we found out that out of the five target patients, two were on Statins and other three are on other medication which are also on the ADR signal list.
- In the Lipophilic Beta Blocker (LBB) group, there are eight patients, among which five are not on Statins.

LBB and Parkinson's Disease

The contingent table for (LBB, PD) is as follows:

		Case: *Developed Parkinson's Disease after Start of Observation Period*		
		Yes	No	Totals
Exposed: *Prescribed Lipophilic Beta Blockers Prior to Development of PD*	Yes	14	3,034	3,048
	No	42	17,687	17,729
	Totals	56	20,721	20,777

- We combine the results from ADR detection stage and text analysis on the clinical notes.
- Risk for PD in the LBB group is $\frac{14}{3{,}048} = 0.46\%$ and $\frac{42}{17729} = 0.24\%$.

Statistical Conclusion

- The odds of developing PD in the LBB group are estimated to be 1.93 times as large as the odds of developing PD in the No-LBB group (p-value < 0.05, 95% CI $= 1.05$ to 3.55).
- The data are consistent with the hypothesis of non-equal proportions of development of PD in LBB vs No LBB (p-value < 0.05).
- Fisher exact test rejects the hypotheses which was that the exposure to LBB does not affect the diagnosis of PD (p-value < 0.05).
- χ^2 test suggests that exposure to LBB and the development of Parkinson's disease are not independent (p-value < 0.05).

Confounding Analysis

- Stratified analysis was performed to check if patient gender was a confounder.
- Patients ages were compared.
- Analyze (LBB,PD) by excluding Statin.
- Results suggest:
  - There is no significant difference between men and women.
  - There is not significant difference between LBB & PD vs LBB & No PD.
  - All statistical tests support the conclusion when we exclude Statins.

Conclusion and Current Research

- Develop a framework to efficiently detect ADR signals.
- The framework incorporate both data analytics and expert domain knowledge.
- Demonstrate the effectiveness by analyzing a new ADR signal, i.e (LBB,PD).

Current research:

- Developing a clinical decision support system for ADR detection for the EMR systems.
- Analyzing the physician notes and messages for more patterns.

Questions?