FDA Workshop
NLP to Extract Information from Clinical Text

Murthy Devarakonda, Ph.D.
Distinguished Research Staff Member
PI for Watson Patient Records Analytics Project
IBM Research
mdev@us.ibm.com

*This work is a part of the IBM Watson EMRA (Electronic Medical Records Analytics) project*
Goal: Cognitive Insights from Longitudinal Patient Records
Watson EMRA Research Initiative

Goal: Cognitive Insights from Longitudinal Patient Records
**Watson EMRA Problem List Generation**

**EMRA Problem List Accuracy:**
- Recall (Sensitivity) = 0.70
- Precision (Positive Predictive Rate) = 0.75

**Entered Problem List Accuracy**
- Recall (Sensitivity) = 0.55
- Precision (Positive Predictive Value) = 0.28

**EMRA Problem List Generation and Physicians Perspective from a Pilot Study**

*International Journal of Medical Informatics*

*In Press, Accepted Manuscript — Note to users*

*Murthy V. Devarakonda, Neil Mehta, Ching-Huei Tsou, Jennifer J. Liang, Amy S. Nowacki, John Eric Jekowsk*
Problem-Oriented Patient Record Summary

- Uses generated problems list
- Relates medications, labs, procedures, and clinical notes to medical problems
- Organizes lists in a clinical order
- Enable one/two click access to raw data such as Notes, labs over a time line, medication history,…

...also, allergies, social history, and demography
Screen Shot: Research Prototype of Watson Patient Record Summary

<table>
<thead>
<tr>
<th>Problems - All (11)</th>
<th>Medications - Active (14)</th>
<th>Lab Results (35)</th>
<th>Vitals (7)</th>
<th>Orders - All (41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Name</td>
<td>Name</td>
<td>Name</td>
<td>Name</td>
</tr>
<tr>
<td>Date</td>
<td>Prescribed</td>
<td>Value</td>
<td>Value</td>
<td>Date</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------</td>
<td>------------------</td>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>gastroesophageal reflux disease</td>
<td>metformin</td>
<td>hgb a1c</td>
<td>182.58</td>
<td>07/23/2012</td>
</tr>
<tr>
<td>12/26/2010</td>
<td>12/26/2010</td>
<td>7.8</td>
<td>07/15/2013</td>
<td></td>
</tr>
<tr>
<td>diabetes mellitus type 2</td>
<td>insulin regular human u-500</td>
<td>estimated average glucose</td>
<td>167</td>
<td>07/15/2013</td>
</tr>
<tr>
<td>12/26/2010</td>
<td>“concentrated”</td>
<td>124</td>
<td>07/15/2013</td>
<td></td>
</tr>
<tr>
<td>obesity</td>
<td>lixivatide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/26/2010</td>
<td>03/21/2012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hyperlipidemia</td>
<td>amoxicillin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03/03/2011</td>
<td>03/06/2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>morbid obesity</td>
<td>ergocalciferol (vitamin d2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05/08/2011</td>
<td>06/09/2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitamin d deficiency</td>
<td>multivitamin capsule</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06/09/2011</td>
<td>05/08/2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>asthma</td>
<td>aspirin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/26/2010</td>
<td>12/26/2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sleep apnea</td>
<td>lisinopril</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/26/2010</td>
<td>12/27/2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benign essential hypertension</td>
<td>hydralazine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/26/2010</td>
<td>10/26/2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>angina</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/26/2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>backache</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/28/2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Indication or Reason to Use Extraction

2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text

Özlem Uzuner,1 Brett R South,2,3,4 Shuying Shen,2,3,4 Scott L DuVall2,3 JAMIA 2011

Limitations:
(1) Relations could be across sentences
(2) Needs aggregation from instances to universal

Box 1 Relation annotated for the i2b2/VA challenge

1. Medical problem—treatment relations:
   a. Treatment improves medical problem (TrIP). Includes mention where the treatment improves or cures the problem, for example, hypertension was controlled on hydrochlorothiazide.
   b. Treatment worsens medical problem (TrWP). Includes mentions where the treatment is administered for the problem but does not cure the problem, does not improve the problem, or makes the problem worse, for example, the tumor was growing despite the available chemotherapeutic regimen.
   c. Treatment causes medical problem (TrCP). The implied cause of the problem is due directly to the treatment itself.

F1 measure

<table>
<thead>
<tr>
<th>Relation classification</th>
<th>Supervised</th>
<th>Semi-supervised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberts et al23</td>
<td>N</td>
<td>0.737</td>
</tr>
<tr>
<td>deBruijn et al25</td>
<td>N</td>
<td>0.731</td>
</tr>
</tbody>
</table>

Generalized problem to medication relation

• Determine if a medication treats/prevents a problem? (not just in sentence)
  \[\text{Lisinopril, HTN} \rightarrow \text{?} \text{ (Ans: 0.78 out of 1.0 (strong association))}\]

• Ensemble of methods:
  – Based on text books, papers, and dictionaries
    • Method 1: Using distributional semantics and UMLS (DRE)
    • Method 3: Using a small part of Watson question answering (SER)
  – Based on coded data in millions of patient records
    • Method 2: Using statistical measures mined (AD), e.g. Odds Ratio at diagnosis (M,D)

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRE</td>
<td>0.67</td>
<td>0.73</td>
<td>0.70</td>
<td>0.76</td>
</tr>
<tr>
<td>AD</td>
<td>0.75</td>
<td>0.54</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>SER</td>
<td>0.56</td>
<td>0.51</td>
<td>0.54</td>
<td>0.52</td>
</tr>
<tr>
<td>DRE + AD</td>
<td>0.78</td>
<td>0.69</td>
<td>0.73</td>
<td>0.79</td>
</tr>
<tr>
<td>DRE + AD + SER</td>
<td>0.79</td>
<td>0.75</td>
<td>0.77</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Does not find or analyze specific instances of reason/indication in a clinic note. That will come later…
Outcome and/or Adverse Events

Concept/framework research exists, but accurate and robust methods needed

Systematic Review: NLP performs better but still accuracy need to improve

From clinical narratives...
Adverse Events – Recent Work

- Two challenges are on the horizon

- Our approach
  - Classify sentences in a clinic note (is it asserting an ADR/ADE?) and then extract the ADR/ADE using the context
    - Can be applied to outcomes as well
  - A challenge is the definition and gold standard
    - "Lisinopril was discontinued and will start on Norvasc due to hyperkalemia" (Clearly ADR)
    - "...discontinued his amlodipine due to low BP .." (is it ADR?)
Will sentence classification work?

- We adapted two methods to classify “sentences”
  - An SVM-based “assertion” framework
  - Deep learning, a convolution neural network (CNN)

Disease Status Sentences

• From Clinical Note text:
  – as a direct statement or as a discussion of the disease-related vitals/labs

<table>
<thead>
<tr>
<th>Model / Features</th>
<th>Training and Testing</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhanced assertion framework (SVM) / features from text alone</td>
<td>10-fold cross validation using manual labels</td>
<td>83% Controlled or not? 77%</td>
</tr>
</tbody>
</table>

401.9 HYPERTENSION NOS
Comment: not controlled today.
Plan: Increase Lisipnopril from 20 to 40 mg. He is not on thiazide. Is also on metoprolol.
He will get a home machine and monitor BP and bring it in next time.

401.9 HYPERTENSION NOS
Comment: BP still running high 160/85.
Plan: increase metoprolol succinate XL from 25 to 50mg daily
OFFICE VISIT: February 27, 2014
Name: Mr. XXXX YYYY
PCP: CCCC, M.D

Active problems:
HTN
Hyperlipidemia
Smoking

ASSessment:

1. HTN – started taking home BPs sporadically. Running 135-150/90-95. Admits to not taking his meds consistently. Have reinforced the importance of controlling his BP due to the cumulative risks for CV events. I explained how to use the feedback from his BP device to help reinforce the importance of his taking his med regularly.
Patient agrees and says he will start taking his BPs regularly and taking his meds.
F: 1. Check home BPs daily; report repeated BPs over 140/90
   2. Reinforced importance of taking meds consistently
   3. May increase meds if home BPs consistently elevated when he is taking his meds regularly
   4. RTC 3 mo

2. Smoking – talked to patient about the combination of smoking and HTN and his risk of MI, stroke. Encouraged him to follow his wife’s advice about joining the smoking cessation class. Enroll in smoking cessation class.

3. Hyperlipidemia – will check LDL. Has been under control in the past.
Recommendations: 1. Lipid panel
   2. COMP METABOLIC PANEL.

<table>
<thead>
<tr>
<th>Method</th>
<th>Plan Sentences (Positive Class)</th>
<th>Positive &amp; Negative Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision Recall F1</td>
<td>Micro F1 Macro F1</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.868 0.230 0.364</td>
<td>0.472 0.600</td>
</tr>
<tr>
<td>SVM All Features</td>
<td>0.952 0.855 0.901</td>
<td>0.858 0.834</td>
</tr>
<tr>
<td>CNN non-static</td>
<td>0.953 0.884 0.917</td>
<td>0.880 0.849</td>
</tr>
</tbody>
</table>
Medication Disposition Sentences

“Add metoprolol 25 mg BID Tues, Thurs, Sat, Sun (days she doesn't have dialysis)”
positive, asserts disposition

“He is now going to gym regularly”
“Reported cough when she was on Lisinopril”
negative, no disposition

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Disposition</td>
<td>0.94</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>Disposition</td>
<td>0.79</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>Macro Average</td>
<td>0.87</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>Micro Average</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Other Medication Related Tasks

• Confounder
  – Will require additional analysis

• Code validation
  – Our NLP stack maps medical concepts to SNOMED CT (CORE), ICD 9, ICD 10, LOINC, and RxNORM

• Clinical trials
  – IBM has an offering for Clinical Trial Mapping (matches form-based input to on-going clinical trials)
Acknowledgements

• Cleveland Clinic
  – For a multi-year research collaboration
  – De-identified patient records with IRB approval

• IBM Watson Researchers and Engineers, especially to the following, whose work is represented here:
  – Ching-Huei Tsou, Bharath Dandala, Jennifer Liang MD, John Prager, Partha Suryanarayanan, Sreeram Joopudi, Ananya Poddar, Diwakar Mahajan, Tong-Haing Fin, Michele Cestone, and Preethi Raghavan

• Paul Tang, MD (IBM)
  – For leadership and education in practice of medicine

• Cleveland Clinic and IBM self funded the joint research
• No other financial disclosures
Summary

• Medications related information extraction from patient record text is clearly possible
• More research is needed for focused and advanced information extraction
• Inconsistencies exist between semi-structured (coded) data and text narratives in patient record
• Challenges/Risks: Availability of patient record data and gold standard