Improving Search for Evidence-based Practice using Information Extraction

Jin Zhao, Min-Yen Kan, Paula M. Procter, Siti Zubaidah, Wai Kin Yip, Goh Mien Li
Evidence-based Practice (EBP)

- Advantages
  - Reliable
  - Efficient
  - Updated
Two Stages of EBP

- **Stage 1: Search and Appraise**
  - EBP Resources → Research Articles → Interventions
  - Crucial yet difficult!

- **Stage 2: Apply and Evaluate**
  - Interventions → Application → Evaluation
Usability Gap between Search Engines and Healthcare Professionals

• EBP-related information
  – PICO elements
    ➢ Patient
      • Sex, CO-morbidity, Race, Age, Pathology (SCORAP)
    ➢ Intervention
    ➢ Condition
    ➢ Outcome
  – Strength of evidence
    ➢ Systematic review of Randomized Control Trials (RCTs)
    ➢ At least one RCT
    ➢ Pseudo-RCT, Cohort studies, …
Usability Gap between Search Engines and Healthcare Professionals

- EBP-related information missing in search results

**Effects of Intermittent Electrical Stimulation on Superficial Pressure, Tissue Oxygenation, and Discomfort Levels for the Prevention of Deep Tissue Injury.**

Solis LR, Gyawali S, Seres P, Curtis CA, Chong SL, Thompson RB, Musshahwar VK. Rehabilitation Science Program, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada.

**Abstract**

The overall goal of this project is to develop effective methods for the prevention of deep tissue injury (DTI). DTI is a severe type of pressure ulcer that originates at deep bone-muscle interfaces as a result of the prolonged compression of tissue. It affects individuals with reduced mobility and sensation, particularly those with spinal cord injury. We previously proposed using a novel electrical stimulation paradigm called intermittent electrical stimulation (IES) for the prophylactic prevention of DTI. IES-induced contractions mimic the natural repositioning performed by intact individuals, who subconsciously reposition themselves as a result of discomfort due to prolonged sitting. In this study, we investigated the effectiveness of various IES paradigms in reducing pressure around the ischial tuberosities, increasing tissue oxygenation throughout the gluteus muscles, and reducing sitting discomfort in able-bodied volunteers. The results were compared to the effects of voluntary muscle contractions and conventional pressure relief maneuvers (wheelchair push-ups). IES significantly reduced pressure around the ischial tuberosities, produced significant and long-lasting elevations in tissue oxygenation, and significantly reduced discomfort produced by prolonged sitting. IES performed as well or better than both voluntary contractions and chair push-ups. The results suggest that IES may be an effective means for the prevention of DTI.
Bridging the Gap…

• Automated extraction of EBP-related information

We performed an open, prospective, randomized clinical trial in 51 patients receiving mechanical ventilation for more than 72 h, in order to evaluate the impact of using noninvasive (quantitative endotracheal aspirates [QEA]) diagnostic method on the morbidity and mortality of ventilator-associated pneumonia (VAP).

• Integration into the search process
  – Filtering and sorting function
  – Re-ranking based on key information
Two-stage Extraction Pipeline

Key Sentences:
- Patient
- Intervention
- Result
- Study Design
- Research Goal

Keywords:
- Sex
- Race
- Age
- Condition
- Intervention
- Study Design
Extraction via Supervised Classification

- **Binary Maximum Entropy Classifiers**
  - One for each class
  - Extracted if the corresponding classifier reports positive
  - Text features only

<table>
<thead>
<tr>
<th>Feature Group</th>
<th>Definition</th>
<th>Feature Group</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Token</td>
<td>The N-grams of the sentence</td>
<td>Token</td>
<td>The word itself, its stem and its part-of-speech</td>
</tr>
<tr>
<td>Sentence</td>
<td>The length and position of the sentence</td>
<td>Phrase</td>
<td>The position of the word in the phrase and the head noun in the phrase</td>
</tr>
<tr>
<td>Named Entity</td>
<td>Whether the sentence contains person, location, organization names</td>
<td>Named Entity</td>
<td>Whether the word is part of person, location, organization names</td>
</tr>
<tr>
<td>MeSH</td>
<td>Whether the sentence contains a MeSH term of a particular category</td>
<td>MeSH</td>
<td>Whether the word is part of a MeSH term of a particular category</td>
</tr>
<tr>
<td>Lexica</td>
<td>Whether the sentence contains a word from the manually compiled wordlist for sex/age/race.</td>
<td>Lexica</td>
<td>Whether the word is part of an entry in the manually compiled wordlist for sex/age/race.</td>
</tr>
</tbody>
</table>
Evaluation

- **Corpus development**
  - 19,893 medical abstracts and full text articles from 17 journal websites
  - 2,000 randomly selected sentence annotated
  - All (6,754) words in sentences belonging to Patient, Intervention and Study Design annotated

<table>
<thead>
<tr>
<th>Patient</th>
<th>Intervention</th>
<th>Result</th>
<th>Study Design</th>
<th>Research Goal</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>8%</td>
<td>4.8%</td>
<td>22.5%</td>
<td>10.2%</td>
<td>3.5%</td>
<td>66.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th>Race</th>
<th>Condition</th>
<th>Intervention</th>
<th>Study Design</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8%</td>
<td>2.6%</td>
<td>0.2%</td>
<td>5.4%</td>
<td>5.5%</td>
<td>3.8%</td>
<td>82.4%</td>
</tr>
</tbody>
</table>
Evaluation

• Metrics
  – Precision, Recall and F$_1$-Measure
    - Precision (P) = TP / (TP+FP)
    - Recall (R) = TP / (TP+FN)
    - F$_1$-Measure (F) = 2 * P * R / (P + R)

where TP: true positive, FP: false positive, FN: false negative.
Results for Sentence Extraction

<table>
<thead>
<tr>
<th>Class</th>
<th>All</th>
<th>No Token</th>
<th>No Sentence</th>
<th>No Named Entity</th>
<th>No MeSH</th>
<th>No Lexica</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F</td>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td>Patient</td>
<td>.70</td>
<td>.24</td>
<td>.36</td>
<td>-13 +.08 +.05</td>
<td>-10 - .04 - .06</td>
<td>+.02 - .01 - .01</td>
</tr>
<tr>
<td>Intervention</td>
<td>.90</td>
<td>.28</td>
<td>.43</td>
<td>-.40 - .10 - .16</td>
<td>-.14 - .05 - .07</td>
<td>-.04 - .03 - .04</td>
</tr>
<tr>
<td>Result</td>
<td>.78</td>
<td>.56</td>
<td>.65</td>
<td>-.17 - .23 - .22</td>
<td>-.02 +.02 +.01</td>
<td>-.01 0 0</td>
</tr>
<tr>
<td>Study Design</td>
<td>.89</td>
<td>.39</td>
<td>.54</td>
<td>-.56 -.29 -.39</td>
<td>-.09 -.04 -.06</td>
<td>-.01 -.05 -.05</td>
</tr>
<tr>
<td>Research Goal</td>
<td>.92</td>
<td>.27</td>
<td>.42</td>
<td>-.47 -.08 -.15</td>
<td>+.03 -.01 -.01</td>
<td>-.05 +.02 +.02</td>
</tr>
</tbody>
</table>

- Precise but much room for improvement in recall
- Difficulty in handling structural variation and short information
- Important features: Token and Sentence
Results for Keywords Extraction

<table>
<thead>
<tr>
<th>Class</th>
<th>All</th>
<th>No Token</th>
<th>No Phrase</th>
<th>No Named Entity</th>
<th>No MeSH</th>
<th>No Lexica</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F</td>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td>Sex</td>
<td>.98</td>
<td>1</td>
<td>.99</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Condition</td>
<td>.74</td>
<td>.62</td>
<td>.67</td>
<td>+.03</td>
<td>-.03</td>
<td>-.15</td>
</tr>
<tr>
<td>Race</td>
<td>1</td>
<td>.86</td>
<td>.92</td>
<td>-.08</td>
<td>0</td>
<td>-.04</td>
</tr>
<tr>
<td>Age</td>
<td>.84</td>
<td>.75</td>
<td>.79</td>
<td>-.04</td>
<td>-.07</td>
<td>-.06</td>
</tr>
<tr>
<td>Intervention</td>
<td>.73</td>
<td>.55</td>
<td>.63</td>
<td>0</td>
<td>-.01</td>
<td>-.01</td>
</tr>
<tr>
<td>Study Design</td>
<td>.91</td>
<td>.71</td>
<td>.80</td>
<td>-.15</td>
<td>-.23</td>
<td>-.21</td>
</tr>
</tbody>
</table>

- Better performance in general
- Difficulty in handling lexical variations and diverse vocabulary
- Positive contribution from all feature groups
Future Work

• Joint inference
  – Allows interactions between two extraction steps

• Bigger knowledge source
  – Addresses diverse vocabulary problem

• Integration of extracted information into search process
Conclusion

- Two-stage pipeline for key information extraction
  - Extraction using binary Maximum Entropy classifiers and text features

- To improve extraction performance and integrate extracted information into search process