Natural language processing approaches to extracting patient functioning from clinical data

Denis Newman-Griffis

Stanford Center for Population Health Sciences
March 16, 2018
Outline

Motivation
   Motivating example
   Key questions

Defining the problem

Pilot annotation study: Mobility

Challenges of functioning data
Motivating example: Disability adjudication support

New claimant applying for disability
Motivating example: Disability adjudication support

New claimant applying for disability

Application includes

- Allegations
- *(Optionally)* Self-collected medical evidence
Motivating example: Disability adjudication support

New claimant applying for disability

Application includes
- Allegations
- *(Optionally)* Self-collected medical evidence
Motivating example: Disability adjudication support

SSA needs further evidence to decide the case
Motivating example: Disability adjudication support

SSA needs further evidence to decide the case

Adjudicator contacts care facilities for related claimant records

- Here, assume 3 different ERs
Motivating example: Disability adjudication support

SSA needs further evidence to decide the case

Adjudicator contacts care facilities for related claimant records

- Here, assume 3 different ERs

Care facilities send back hundreds of pages of records
Motivating example: Disability adjudication support

Adjudicator now has all necessary information

- Hundreds of pages of records
Motivating example: Disability adjudication support

Adjudicator now has all necessary information
  - Hundreds of pages of records

Only a small subset are relevant to the case
  - Much relevant information in free text observations
  - This is where NLP comes in!
Subfield of AI: looks at processing information that is described in human language

- Broadly: going from unstructured text to structured data
- Clinical applications in phenotyping, DDI, ADE detection, etc
Subfield of AI: looks at processing information that is described in human language

- Broadly: going from unstructured text to structured data
- Clinical applications in phenotyping, DDI, ADE detection, etc

Medication list:
- 300 mg aspirin p.o. tid
- 150 ml ethanol p.o. daily
- 2 tab methotrexate MWF
NLP in 30 seconds

Subfield of AI: looks at processing information that is described in human language

▶ Broadly: going from unstructured text to structured data
▶ Clinical applications in phenotyping, DDI, ADE detection, etc

Medication list:
- 300 mg aspirin p.o. tid
- 150 ml ethanol p.o. daily
- 2 tab methotrexate MWF

<table>
<thead>
<tr>
<th>Medication</th>
<th>Dose</th>
<th>Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>300mg</td>
<td>3x daily</td>
</tr>
<tr>
<td>Ethanol</td>
<td>150ml</td>
<td>1x daily</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>2tab</td>
<td>MWF</td>
</tr>
</tbody>
</table>
Functioning and Diagnostic information

Diagnostic

- Concerned with detecting, describing, and treating conditions
- Diagnoses, symptoms, procedures, measurements, etc
- Focus of most clinical NLP

Functioning

- Concerned with evaluating, describing, and rehabilitating impact of health conditions
- Activities (mobility, self care, communication, domestic life, etc), participation, rehabilitative care, goals, etc
Goal: Identify functioning data in free text medical records relevant to alleged impairments

Key research questions:
**Goal:** Identify functioning data in free text medical records relevant to alleged impairments

**Key research questions:**

1. What does functioning information look like in text?
**Goal:** Identify functioning data in free text medical records relevant to alleged impairments

**Key research questions:**

1. What does functioning information look like in text?

2. What new and existing NLP techniques can we apply to get it?
Filtering to relevant data

**Goal:** Identify functioning data in free text medical records relevant to alleged impairments

Key research questions:

1. What does functioning information look like in text?

2. What new and existing NLP techniques can we apply to get it?

3. What challenges does functioning information pose?
Outline

Motivation

Defining the problem
  Functioning as an information domain
  The language of functioning

Pilot annotation study: Mobility

Challenges of functioning data
Individuals engage in different tasks and life situations in daily life, and operate in different environments.

*Functioning* (roughly) describes the ability to do these things in a given health condition.
Individuals engage in different tasks and life situations in daily life, and operate in different environments.

*Functioning* (roughly) describes the ability to do these things in a given health condition.

We consider this at the **individual** level!

But there are lots of interesting population-level questions as well.
Health condition: severe back pain when standing/walking for more than 10 minutes

Effects on three hypothetical cases:

1. Retiree in assisted living facility, has existing transportation needs to church, store, etc

2. Packing manager in warehouse

3. Computer programmer who backpacks on the weekends
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

Health condition (disorder or disease)

Body functions and structure

Activity

Participation

Environmental factors

Personal factors

Contextual factors
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

- Physiological, psychological functions; Anatomical structures
- Health condition (disorder or disease)
- Body functions and structure
- Activity
- Participation
- Environmental factors
- Personal factors

Contextual factors
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

```
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

- Body functions and structure
- Activity
- Participation

- Environmental factors
- Personal factors

Execution of task or action

Contextual factors
```
The ICF Framework

International Classification of Functioning, Disability, and Health

▶ Companion to ICD-N
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

![Diagram of the ICF Framework]

- Health condition (disorder or disease)
- Body functions and structure
- Activity
- Participation
- Environmental factors
- Personal factors

*Contextual factors*
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N
The ICF Framework

International Classification of Functioning, Disability, and Health

- Companion to ICD-N

![Diagram of the ICF Framework]

- Health condition (disorder or disease)
- Body functions and structure
- Activity
- Participation
- Environmental factors
- Personal factors
- Contextual factors
Corpus analysis of functioning language

**Goal:** describe linguistic patterns in documentation of functioning

**Primary resource:** BTRIS
- 155,000 documents from throughout NIH Clinical Center
- 68,000 from Rehabilitation Medicine Department
- 85,000 from various other CC departments
- Automatically deidentified before use

**Auxiliary resources**
- 450,000 documents from Ohio State Wexner Medical Center
- 2 million documents from MIMIC-III
Two levels of data classification: Domain and Discipline.

**Domain**

- **Functioning** - concerned primarily with patient functioning
- **Diagnostic** - concerned primarily with diagnosis/treatment of health conditions
Two levels of data classification: *Domain* and *Discipline*.

**Discipline**

- **Therapy** - documents related to therapeutic encounters (phys ther, occ ther, pulm ther, etc)
- **Medical** - non-therapeutic medical documents (majority of records)
- **Ancillary** - ancillary care encounters, including psychological evaluation and social work
- **Other** - primarily administrative documents
## Document classes within corpora

<table>
<thead>
<tr>
<th></th>
<th>BTRIS</th>
<th>MIMIC-III</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic</td>
<td>68,501</td>
<td>2,075,079</td>
</tr>
<tr>
<td>Function</td>
<td>59,532</td>
<td>8,101</td>
</tr>
<tr>
<td><strong>Discipline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>71,799</td>
<td>2,074,112</td>
</tr>
<tr>
<td>Therapy</td>
<td>49,055</td>
<td>5,431</td>
</tr>
<tr>
<td>Ancillary</td>
<td>6,496</td>
<td>2,670</td>
</tr>
<tr>
<td>Other</td>
<td>683</td>
<td>967</td>
</tr>
</tbody>
</table>

- Labels assigned heuristically, based on document type in source EHR
- OSUWMC shows similar splits to MIMIC-III
Vocabularies are highly distinctive

Approx 50% overlap between Functional and Diagnostic vocabularies (BTRIS)
Vocabularies are highly distinctive

Similar small overlaps at Discipline level
Vocabularies are highly distinctive

Example subset-specific words:

**Domain**

- **Functional**: amusing, care/cleaning, antipsychotic-induced
- **Diagnostic**: hernia/cyst, prebronchodilator, ovulated

**Discipline**

- **Therapy**: activities/interventions, self-advocacy
- **Ancillary**: youngster, downplaying
- **Medical**: accentuated, leukoencephalopathy

Can distinguish documents with very high accuracy by keyword frequency
Manual review findings

Stratified random sample of 75 documents (sampled from each class)

Evaluated cTAKES performance on SBD, POS tagging, NER
  ▶ Used to represent go-to clinical NLP tools in practice

Findings:
  ▶ Sentence splitting overeager; separates some information
  ▶ NER and normalization consistently poor on functioning information
    ▶ Many critical concepts missed entirely
    ▶ bed → “Bornholm eye disease”
  ▶ Reliable performance on Diagnostic/Medical data
Outline

Motivation

Defining the problem

Pilot annotation study: Mobility
  Information structure for Mobility
  Annotation development
  Establishing baselines for Mobility recognition

Challenges of functioning data
Pilot study: Mobility

Activities and Participation:
- Learning and Applying Knowledge
- General Tasks and Demands
- Communication
- Mobility
- Self-Care
- Domestic Life
- Interpersonal Interactions and Relationships
- Major Life Areas
- Community, Social and Civic Life
Mobility has several advantages as an information domain:

- Self-contained; does not rely heavily on environmental factors
- Manageable in size: we use 13 3-digit ICF codes
  - Capture health outcome evaluations
- Correlated to work disability
Mobility coding

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>d410</td>
<td>Changing basic body position</td>
</tr>
<tr>
<td>d415</td>
<td>Maintaining a body position</td>
</tr>
<tr>
<td>d420</td>
<td>Transferring oneself</td>
</tr>
<tr>
<td>d430</td>
<td>Lifting and carrying objects</td>
</tr>
<tr>
<td>d435</td>
<td>Moving objects with lower extremities</td>
</tr>
<tr>
<td>d440</td>
<td>Fine hand use</td>
</tr>
<tr>
<td>d445</td>
<td>Hand and arm use</td>
</tr>
<tr>
<td>d450</td>
<td>Walking</td>
</tr>
<tr>
<td>d455</td>
<td>Moving around</td>
</tr>
<tr>
<td>d460</td>
<td>Moving around in different locations</td>
</tr>
<tr>
<td>d470</td>
<td>Using transportation</td>
</tr>
<tr>
<td>d475</td>
<td>Driving</td>
</tr>
<tr>
<td>d480</td>
<td>Riding animals for transportation</td>
</tr>
</tbody>
</table>

Example usage:

He is independent with bed mobility supine to sit and sit to supine [d410]. He can move with min A from sit on raised surface to stand w/o a.d. [d410] He experienced frequent LOB and able to self-correct 50% of the time. He was provided with r.w. and was able to walk in the room with it [d450].
Structure of Mobility information

Mobility span

The patient ambulates with modified independence for 300 ft.

with modified independence

Assistance (Modifier)

Ambulates

Action (Head)

300 ft.

Quantification (Modifier)
Structure of Mobility information
Initial annotations

Subset of Physical Therapy documents selected for annotation

250 documents fully annotated

- Mobility mentions: 2,978
- Actions: 2,867
- Assistance: 1,671
- Quantification: 1,227
- Score definition: 157

[He has resumed playing basketball, [jogging $d455$], [doing squats $d410$]]

*Pt’s husband states* [she is [mostly independent $ASST$] for [ambulation $d450$]]

*and* [has about [[10 steps $QUANT$] up to each level, which she is able to climb $d455$]], *but*

- Mentions may be overlapping (see third example)
- Mobility mentions may not have Action *or* Assistance *or* Quantification
Two baseline methods:
- Stanford NER (CRF-based model)
- LSTM-CRF (Lample et al., 2016)

Predict unique combinations of tags:
- she-0 is-0 mostly-M/ASST independent-M/ASST
- for-M ambulation-M/ACT

Evaluation settings:
- 3-fold cross-validation
- Evaluating based on partial match and exact match
Both baseline models doing quite well on Mobility NER

- 80-90% F-measure on exact matches
- >90% F-measure on partial matches
- Score Definition and Quantification are easiest, Mobility is hardest
Both baseline models doing quite well on Mobility NER

- 80-90% F-measure on exact matches
- >90% F-measure on partial matches
- Score Definition and Quantification are easiest, Mobility is hardest

Some caveats

- Pretty high linguistic regularity in our data (common training for all providers)
- Expect to see decreased performance on more diverse data
Current successes and failures

Significant portion of perfect matches (even complex cases)

[He evals as [ind w/ [walking ACT] and self-care w/o the use of assistive devices ASST]]

[Pt should continue with core strengthening exercises and her [jogging ACT]/[hiking ACT]]

Some complete misses

[currently 3/10 while [lying supine ACT] on bed]

[Pt reported she led w/ left LE during both [ascent and descent ACT]]
Current successes and failures

Some evidence of lexical memorization

- Gold: [w/o the need for gait aids ASST]
- Predicted: [[w/o the need ASST] for [gait ACT] aids]

Generally, CRF model does better on multi-label cases

- More mistakes on Mobility-only tokens
- Overextends some sub-entity bounds

Seems to perform better on Mobility mentions at start of sentence than mid-sentence
Outline

Motivation

Defining the problem

Pilot annotation study: Mobility

Challenges of functioning data
  Short-term challenges
  Long-term challenges
**Technical challenges**

**The big one:** lack of standardized terminology / ontology for functioning information

- Many clinical NLP successes rely on SNOMED, UMLS, etc.
- ICF is a conceptual framework, not a robust terminology
  - $d450$ is “walking”, not “ambulation”

**Approaches**

- Develop a terminology! *(big ask)*
- Investigating methods for learning representations of concepts from text + seed set of terms
Another big one: lack of annotated data

- Many successful NLP methods require huge amounts of labeled data
- Haven’t yet established common standards for what data structures/relationships at application level should look like

Approaches

- Drawing on existing data from other domains
  - Pre-training text representations on large web corpus and tuning on BTRIS improves NER performance
- Starting to look at unsupervised/semi-supervised methods to transfer data from related tasks; help increase training data
Document formatting

Form fields/options in templates → concepts that aren’t actually observed

- Difficulty walking X
- Difficulty sitting
- Difficulty xfer chair to bed

Therapy notes often include patient goals

- Functional data, but hypothetical
- Walk 300’ (2 weeks)
Linguistic challenges

Document content
Self-reported functional observations

- Often not practical to evaluate in clinical setting
- Highly different linguistic structure
- pt spontaneously weight shifting ‘my bottom hurts’

Relevant descriptions of current environment/occupation

- May often imply functional requirements
- Plans to go to son’s house; 6 steps to enter one level apartment

Document structure varies highly between institutions
The moonshot question

Many patients have minimal access to rehabilitative care

- Means most medical evidence will be entirely diagnostic

How can we develop methods to infer functional status from diagnostic observations?
Ongoing research support comes from the Intramural Research Program at the National Institutes of Health, Clinical Research Center and through an Inter-Agency Agreement with the US Social Security Administration.
Thank you!

Contact me at:
denis.griffis@nih.gov
drgriffis.github.io