Automated Summarisation for Evidence Based Medicine

Diego Mollá

Centre for Language Technology,
Macquarie University

HAIL, 22 March 2012

Contents

About Us: Research Group on Natural Language Processing of Medical Texts

http://web.science.mq.edu.au/~diego/medicalnlp/

Active Members

Diego Mollá	Senior lecturer at Macquarie University.
Cécile Paris	Senior principal research scientist at CSIRO ICT Centre.
Abeed Sarker	PhD student at Macquarie University.
Sara Faisal Shash	Masters student.

Past Members

María Elena Santiago-Martínez	Research programmer.
Patrick Davis-Desmond	Masters student.
Andreea Tutos	Masters student.

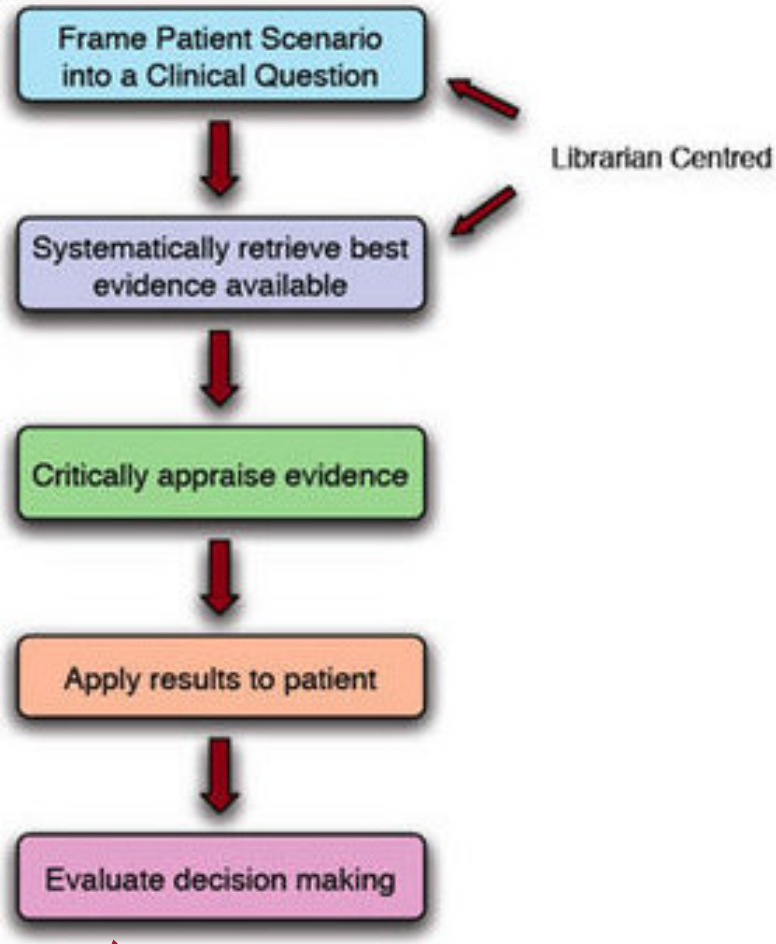
Contents

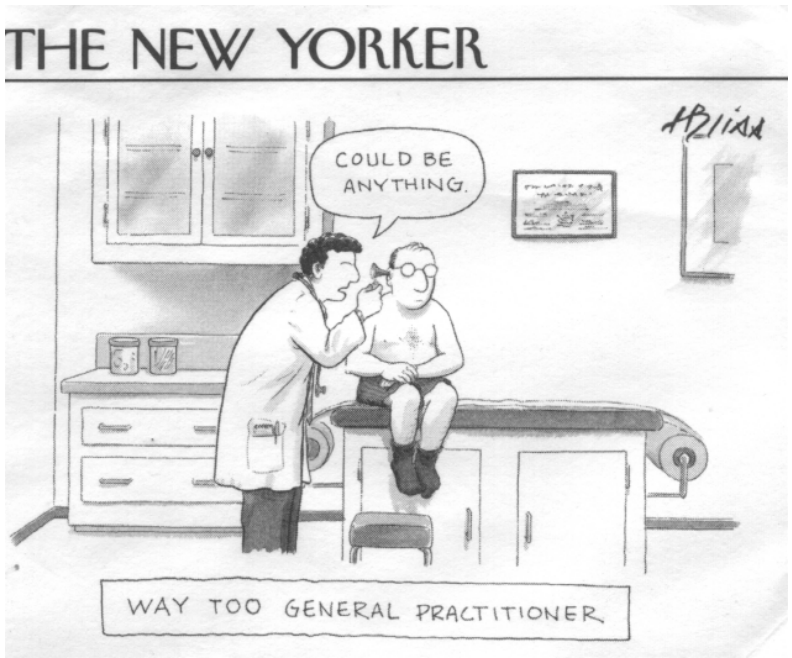
Evidence Based Medicine

http://laikaspoetnik.wordpress.com/2009/04/04/evidence-based-medicine-the-facebook-of-medicine/

EBM and Natural Language Processing

http://hlwiki.slais.ubc.ca/index.php?title=Five_steps_of_EBM





PICO for Asking the Right Question

	1	2	3	4
	Patient or Problem	**Intervention** (a cause, prognostic factor, treatment, etc.)	**Comparison Intervention** (if necessary)	**Outcomes**
Tips for Building	Starting with your patient, ask "How would I describe a group of patients similar to mine?" Balance precision with brevity.	Ask "Which main intervention am I considering?"	Ask "What is the main alternative to compare with the intervention?" Again, be specific	Ask "What can I hope to accomplish?", or "What could this exposure really affect?" Again, be specific
Example	"In patients with heart failure from dilated cardiomyopathy who are in sinus rhythm..."	"...would adding anticoagulation with warfarin to standard heart failure therapy..."	"...when compared with standard therapy alone..."	"...lead to lower mortality or morbidity from thromboembolism. Is this enough to be worth the increased risk of bleeding?"

Where to search for external evidence?

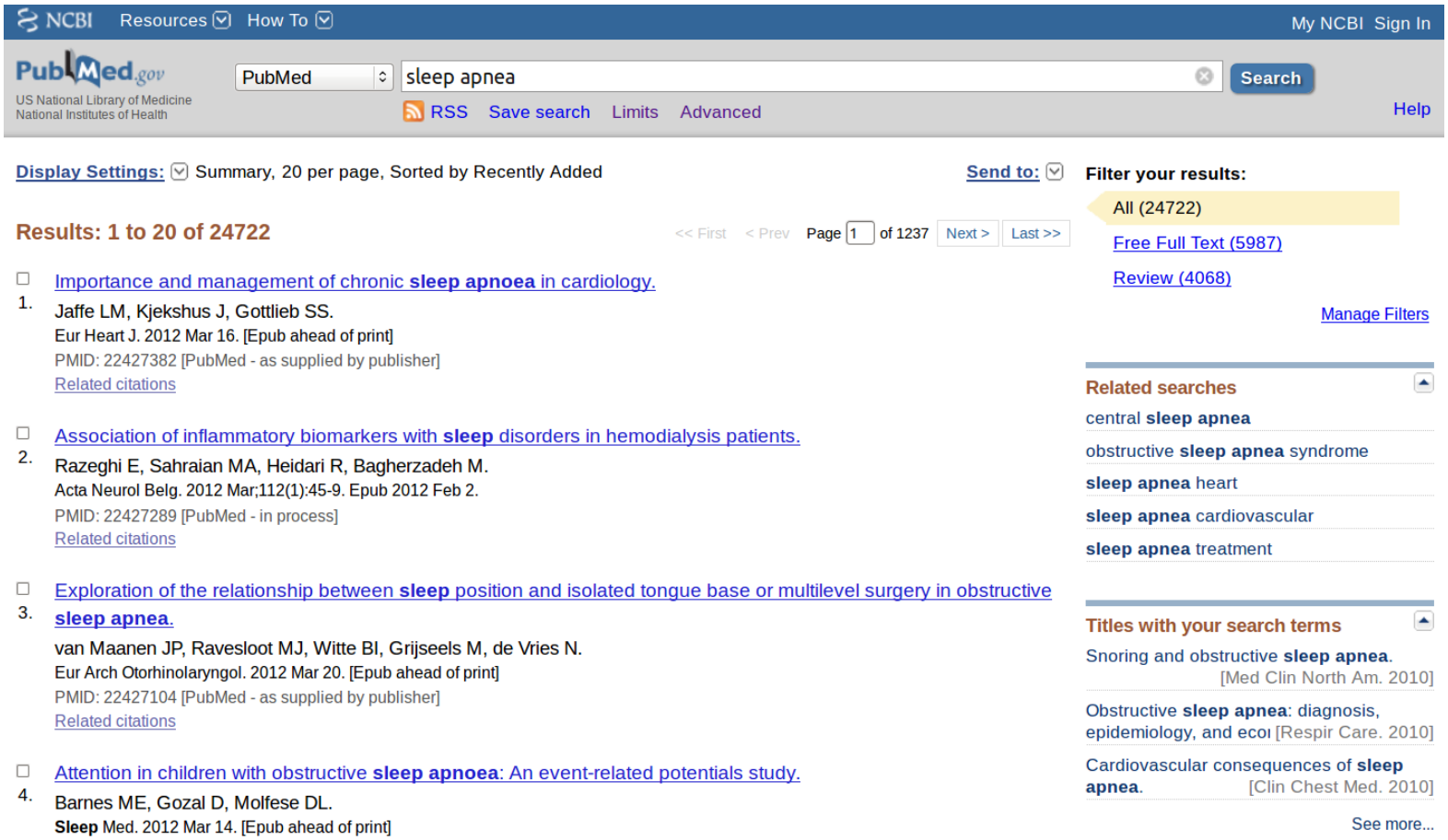
1. Evidence-based Summaries (Systematic Reviews):
   - EBM Online (http://ebm.bmj.com).
   - UptoDate (http://www.uptodate.com).
   - The Cochrane Library (http://www.thecochranelibrary.com/).
   - ...

Where to search for external evidence?

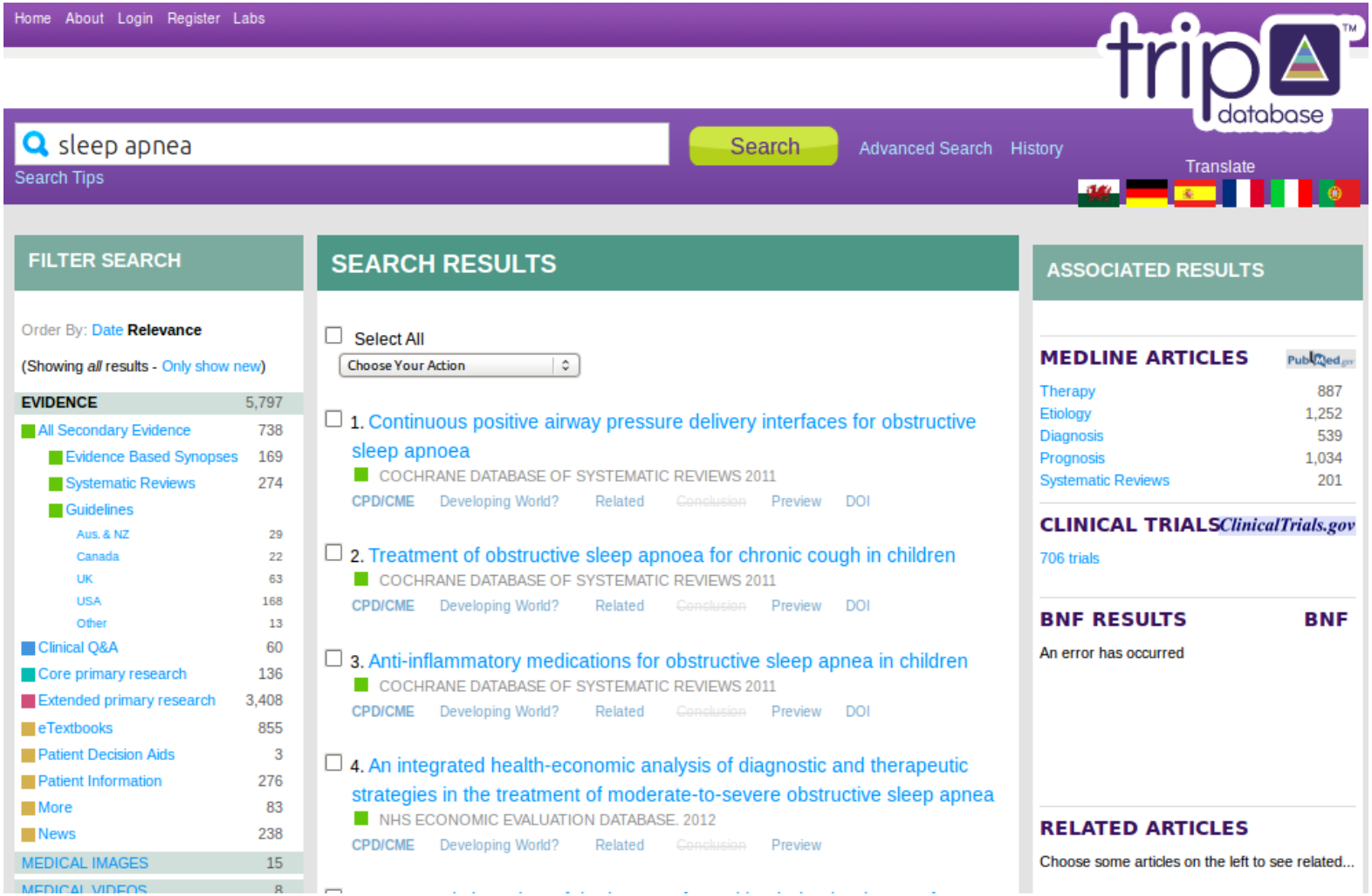
1. Evidence-based Summaries (Systematic Reviews):
   - EBM Online (http://ebm.bmj.com).
   - UptoDate (http://www.uptodate.com).
   - The Cochrane Library (http://www.thecochranelibrary.com/).
   - ...
2. Search the Medical Literature:
   - E.g. PubMed (http://www.ncbi.nlm.nih.gov/pubmed/).

Searching Cochrane

Searching PubMed



Searching the Trip Database



Appraising the Evidence

The SORT Taxonomy

Level A Consistent and good-quality patient-oriented evidence.

Level B Inconsistent or limited-quality patient-oriented evidence.

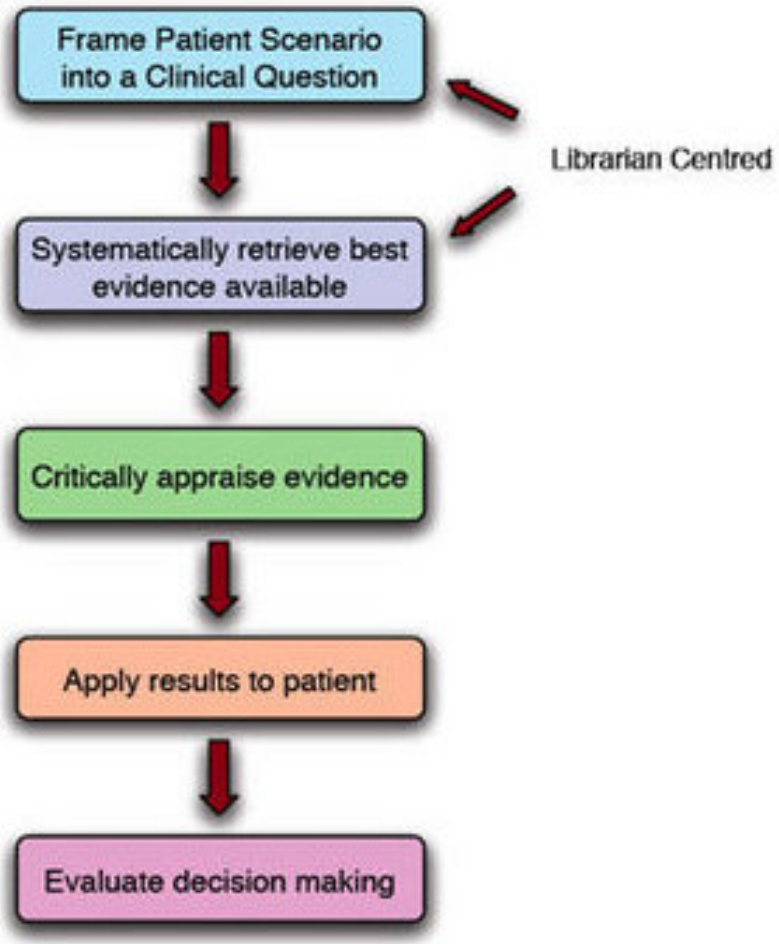
Level C Consensus, usual practise, opinion, disease-oriented evidence, or case series for studies of diagnosis, treatment, prevention, or screening.

Levels of Evidence

Study quality	Diagnosis	Treatment / prevention / screening	Prognosis
Level 1: good-quality patient-oriented evidence	Validated clinical decision rule; SR/meta-analysis of high-quality studies; high-quality diagnostic cohort study	SR/meta-analysis of RCTs with consistent findings; high-quality individual RCT; all-or-none study	SR/meta-analysis of good-quality cohort studies; prospective cohort study with good follow-up
Level 2: limited-quality patient-oriented evidence	Unvalidated clinical decision rule; SR/meta-analysis of lower-quality studies or studies with inconsistent findings; lower-quality diagnostic cohort study or diagnostic case-control study	SR/meta-analysis of lower-quality clinical trials or of studies with inconsistent findings; lower-quality clinical trial; cohort study; case-control study	SR/meta-analysis of lower-quality cohort studies or with inconsistent results; retrospective cohort study or prospective cohort study with poor follow-up; case-control study; case series
Level 3: other evidence	Consensus guidelines, extrapolations from bench research, usual practice, opinion, disease-oriented evidence (intermediate or physiologic outcomes only), or case series for studies of diagnosis, treatment, prevention, or screening		

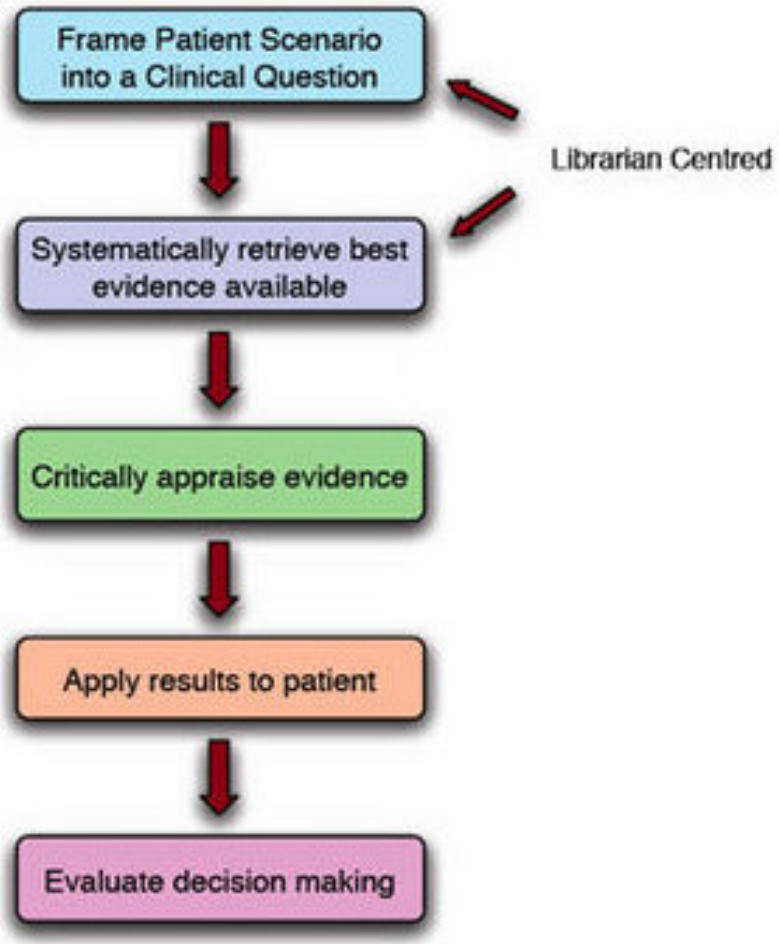
Where can NLP Help?

- Questions:
  - Help to formulate answerable questions.
  - Question analysis and classification.



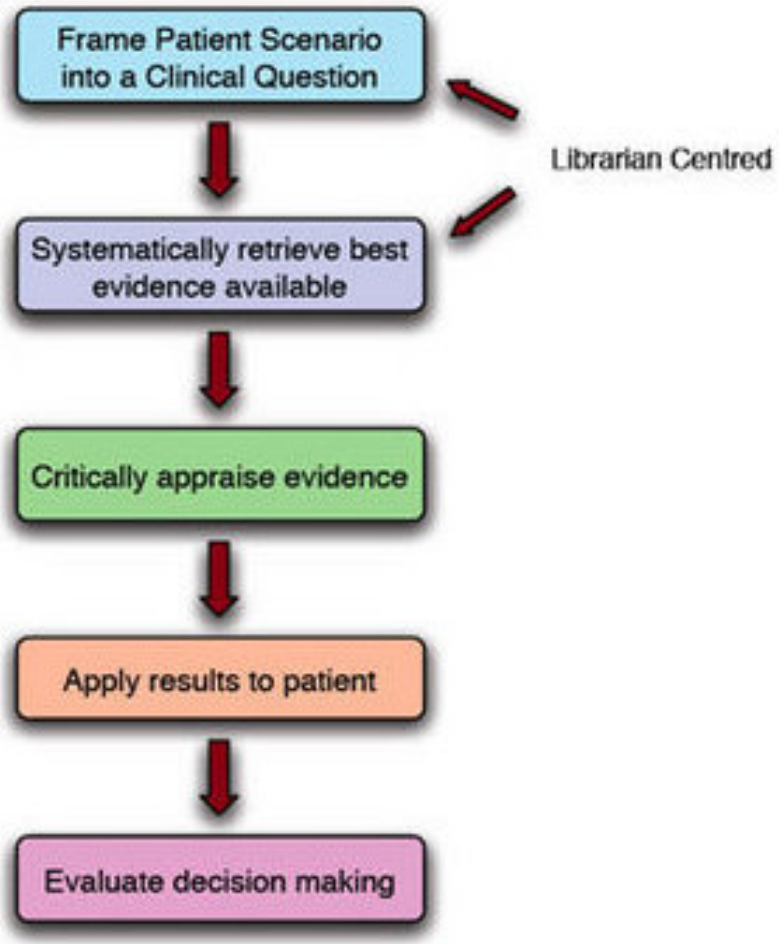
Where can NLP Help?

- Questions:
  - Help to formulate answerable questions.
  - Question analysis and classification.
- Search:
  - Retrieve and rank relevant literature.
  - Extract the evidence-based information.
  - Summarise the results.



Where can NLP Help?

- Questions:
  - Help to formulate answerable questions.
  - Question analysis and classification.
- Search:
  - Retrieve and rank relevant literature.
  - Extract the evidence-based information.
  - Summarise the results.
- Appraisal: Classify the evidence.



Contents

Where's the Corpus for Summarisation?

Summarisation Systems

- CENTRIFUSER/PERSIVAL: Developed and tested using user feedback (iterative design).
- SemRep: Evaluation based on human judgement.
- Demner-Fushman & Lin: ROUGE on original paper abstracts.
- Fiszman: Factoid-based evaluation.

Where's the Corpus for Summarisation?

Summarisation Systems

- CENTRIFUSER/PERSIVAL: Developed and tested using user feedback (iterative design).
- SemRep: Evaluation based on human judgement.
- Demner-Fushman & Lin: ROUGE on original paper abstracts.
- Fiszman: Factoid-based evaluation.

Corpora

- Several corpora of questions/answers available.
- Answers lack explicit pointers to primary literature.
- Medical doctors want to know the primary sources.

Contents

Journal of Family Practice's "Clinical Inquiries"

Which treatments work best for hemorrhoids?

Evidence-based answer

Excision is the most effective treatment for thrombosed external hemorrhoids (strength of recommendation [SOR]: **B**, retrospective studies). For prolapsed internal hemorrhoids, the best definitive treatment is traditional hemorrhoidectomy (SOR: **A**, systematic reviews). Of nonoperative techniques, rubber band ligation produces the lowest rate of recurrence (SOR: **A**, systematic reviews).

Evidence summary

External hemorrhoids originate below the dentate line and become acutely painful with thrombosis. They can cause perianal pruritus and excoriation because of interference with perianal hygiene. Internal hemorrhoids become symptomatic when they bleed or prolapse (TABLE).

For thrombosed external hemorrhoids, surgery works best

Few studies have evaluated the best treatment for thrombosed external hemorrhoids. A retrospective study of 231 patients treated conservatively or surgically found that the 48.5% of patients treated surgically had a lower recurrence rate than the conservative group (number needed to treat [NNT]=2 for recurrence at mean follow-up of 7.6 months) and earlier resolution of symptoms (average 3.9 days compared with 24 days for conservative treatment).[1]

Another retrospective analysis of 340 patients who underwent outpatient excision of thrombosed external hemorrhoids under local anesthesia reported a low recurrence rate of 6.5% at a mean follow-up of 17.3 months.[2]

A prospective, randomized controlled trial (RCT) of 98 patients treated nonsurgically found improved pain relief with a combination of topical nifedipine 0.3% and lidocaine 1.5% compared with lidocaine alone. The NNT for complete pain relief at 7 days was 3.[3]

Conventional hemorrhoidectomy beats stapling

Many studies have evaluated the best treatment for prolapsed hemorrhoids. A Cochrane systematic review of 12 RCTs that compared conventional hemorrhoidectomy with stapled hemorrhoidectomy in patients with grades I to III hemorrhoids found a lower rate of recurrence (follow-up ranged from 6 to 39 months) in patients who had conventional hemorrhoidectomy (NNT=14).[4] Conventional hemorrhoidectomy showed a nonsignificant trend in decreased bleeding and decreased incontinence.

A second systematic review of 25 studies, including some that were of lower quality, showed a higher recurrence rate at 1 year with stapled hemorrhoidectomy than with conventional surgery.[5]

Nonoperative techniques? Consider rubber band ligation

A systematic review of 3 poor-quality trials comparing rubber band ligation with excisional hemorrhoidectomy in patients with grade III hemorrhoids found that excisional hemorrhoidectomy produced better long-term symptom control but more immediate postoperative complications of anal stenosis and hemorrhage.[6] Rubber band ligation had the lowest recurrence rate at 12 months compared with the other nonoperative techniques of sclerotherapy and infrared coagulation.[7]

Fiber supplements help relieve symptoms

A Cochrane systematic review of 7 RCTs enrolling a total of 378 patients with grade I to III hemorrhoids evaluated the effect of fiber supplements on pain, itching, and bleeding. Persistent hemorrhoid symptoms decreased by 53% in the group receiving fiber.[8]

When surgical hemorrhoidectomy is recommended

The American Society of Colon and Rectal Surgeons recommends adequate fluid and fiber intake for all patients with symptomatic hemorrhoids. For grade I to III hemorrhoids, the society states that banding is usually most effective. When office treatments fail, the society recommends surgical hemorrhoidectomy (SOR: **B**).

The society recommends excision of thrombosed hemorrhoids less than 72 hours old and expectant treatment with hemorrhoids that present early. Surgical hemorrhoidectomy should be reserved for when conservative treatment fails and for patients with symptomatic grade III and IV hemorrhoids.[10] ■

TABLE

Classification of sympto internal hemorrhoid

GRADE	DESCRIPTION
I	Hemorrhoids do not protru
II	Hemorrhoids protrude wit reduce spontaneously
III	Hemorrhoids protrude and by hand
IV	Hemorrhoids are permane

Source: Madoff RD, et al. *Gastroenterology.* 2004.[10]

References

1. Greenspon J, Williams SB, Young HA, et al. Thrombosed external hemorrhoids: outcome after conservative or surgical management. *Dis Colon Rectum.* 2004;47:1493-1498.
2. Jongen J, Bach S, Stubinger SH, et al. Excision of thrombosed external hemorrhoids under local anesthesia: a retrospective evaluation of 340 patients. *Dis Colon Rectum.* 2003;46:1226-1231.
3. Perrotti P, Antropoli C, Molino D, et al. Conservative treatment of acute thrombosed external hemorrhoids with topical nifedipine. *Dis Colon Rectum.* 2001;44:405-409.
4. Jayaraman S, Colquhoun PH, Malthaner RA. Stapled versus conventional surgery for hemorrhoids. *Cochrane Database Syst Rev.* 2006;(4):CD005393.
5. Tjandra JJ, Chan MK. Systematic review on the procedure for prolapse and hemorrhoids (stapled hemorrhoidopexy). *Dis Colon Rectum.* 2007;50:878-892.
6. Shanmugam V, Thaha MA, Rabindranath KS, et al. Systematic review of randomized trials comparing rubber band ligation with excisional haemorrhoidectomy. *Br J Surg.* 2005;92:1481-1487.
7. Johanson JF, Rimm A. Optimal nonsurgical treatment of hemorrhoids: a comparative analysis of infrared coagulation, rubber band ligation, and injection sclerotherapy. *Am J Gastroenterol.* 1992;87:1600-1606.
8. Alonso-Coello P, Guyatt G, Heels-Ansdell D, et al. Laxatives for the treatment of hemorrhoids. *Cochrane Database Syst Rev.* 2005(4):CD004649.

The XML Contents I

```
<record id="7843">
<url>http://www.jfponline.com/Pages.asp?AID=7843&amp;issue=September_2009&amp;UID=</url>
<question>Which treatments work best for hemorrhoids?</question>
<answer>
  <snip id="1">
    <sniptext>Excision is the most effective treatment for thrombosed
 external hemorrhoids.</sniptext>
    <sor type="B">retrospective studies</sor>
    <long id="1_1">
      <longtext>A retrospective study of 231 patients treated
       conservatively or surgically found that the 48.5% of patients
       treated surgically had a lower recurrence rate than the
       conservative group (number needed to treat [NNT]=2 for
       recurrence at mean follow-up of 7.6 months) and earlier
       resolution of symptoms (average 3.9 days compared with 24 days
       for conservative treatment).</longtext>
      <ref id="15486746" abstract="Abstracts/15486746.xml">Greenspon
       J, Williams SB, Young HA ,et al. Thrombosed external
       hemorrhoids: outcome after conservative or surgical
       management. Dis Colon Rectum. 2004; 47: 1493-1498.</ref>
    </long>
    <long id="1_2">
      <longtext>A retrospective analysis of 340 patients who underwent
       outpatient excision of thrombosed external hemorrhoids under
       local anesthesia reported a low recurrence rate of 6.5% at a
```

The XML Contents II

```
      mean follow-up of 17.3 months.</longtext>
      <ref id="12972967" abstract="Abstracts/12972967.xml">Jongen J,
      Bach S, Stubinger SH ,et al. Excision of thrombosed external
      hemorrhoids under local anesthesia: a retrospective evaluation
      of 340 patients. Dis Colon Rectum. 2003; 46: 1226-1231.</ref>
    </long>
    <long id="1_3">
      <longtext>A prospective, randomized controlled trial (RCT) of 98
      patients treated nonsurgically found improved pain relief with a
      combination of topical nifedipine 0.3% and lidocaine 1.5% compared
      with lidocaine alone. The NNT for complete pain relief at 7 days was
      3.</longtext>
      <ref id="11289288" abstract="Abstracts/11289288.xml">Perrotti P,
      Antropoli C, Molino D ,et al. Conservative treatment of acute
      thrombosed external hemorrhoids with topical nifedipine. Dis
      Colon Rectum. 2001; 44: 405-409.</ref>
    </long>
  </snip>
</answer>
</record>
```

Components of the Corpus

Question direct extract from the source.

Answer split from the source and manually checked.

Evidence extracted from the source.

Additional text manually extracted from the source and massaged.

References PMID looked up in PubMed (automatic and manual procedure).

Contents

Annotation of Text Justifications

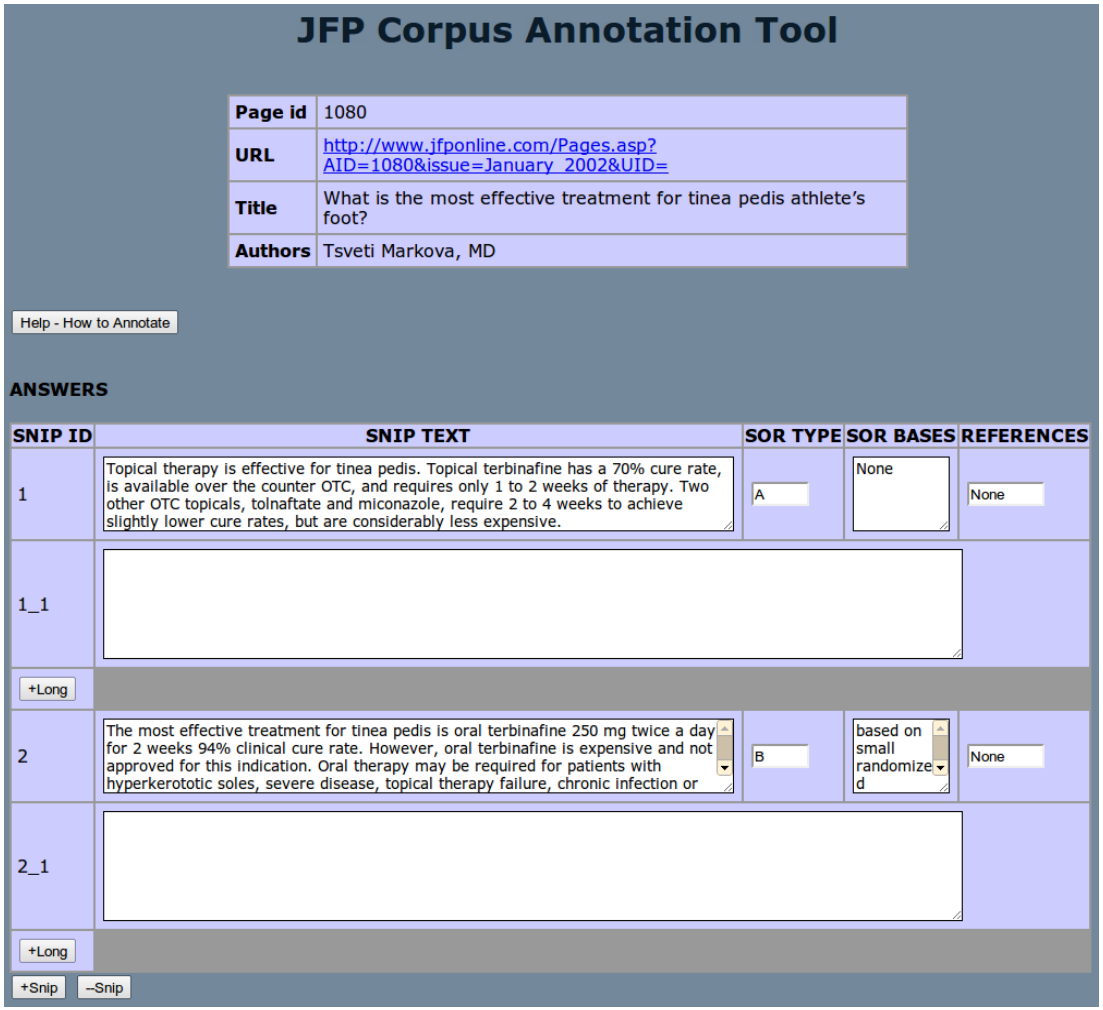
Goal

- Identify the text justifications.
- Align the text justifications with the answer parts.

Method

- Three annotators (members of the research group).
- Annotation tool contains pre-zoned text:
  - answer summary;
  - body text;
  - recommendations;
  - references.
- Annotators need to copy and paste (and massage) the text.

Annotation Tool I



JFP Corpus Annotation Tool

Page id	1080
URL	http://www.jfponline.com/Pages.asp?AID=1080&issue=January_2002&UID=
Title	What is the most effective treatment for tinea pedis athlete's foot?
Authors	Tsveti Markova, MD

Help - How to Annotate

ANSWERS

SNIP ID	SNIP TEXT	SOR TYPE	SOR BASES	REFERENCES
1	Topical therapy is effective for tinea pedis. Topical terbinafine has a 70% cure rate, is available over the counter OTC, and requires only 1 to 2 weeks of therapy. Two other OTC topicals, tolnaftate and miconazole, require 2 to 4 weeks to achieve slightly lower cure rates, but are considerably less expensive.	A	None	None
1_1				
+Long				
2	The most effective treatment for tinea pedis is oral terbinafine 250 mg twice a day for 2 weeks 94% clinical cure rate. However, oral terbinafine is expensive and not approved for this indication. Oral therapy may be required for patients with hyperkerototic soles, severe disease, topical therapy failure, chronic infection or	B	based on small randomized	None
2_1				
+Long				

+Snip –Snip

Annotation Tool II

SUMMARY

The Cochrane Database of Systemic Reviews, reported 72 placebo-controlled trials of topical agents that yielded the following cure rates: undecenoic acid, 72%; allylamines terbinafine, naftifine, butenafine, 70%; tolnaftate, 64%; azoles miconazole, clotrimazole, ketoconazole, econazole, oxiconazole, 47%. A meta-analysis of 11 RCTs suggests that allylamines are slightly more effective than azoles. (REF:1,2).

Orally administered antifungal agents are expensive and can have systemic side effects. Griseofulvin and ketoconazole are approved for oral therapy, but product labels clearly state that they should be used only after topical agents have failed. Griseofulvin has been used for more than 30 years, is well tolerated, and efficacious in treating dermatomycoses in the range of 60%. Ketoconazole's cure rate is similar, but its use in cutaneous infections is limited by multiple drug interactions and serious side effects. Three placebo-controlled RCTs of itraconazole of varying doses and duration of treatment suggested favorable clinical cure of moccasin-type tinea pedis 51%-85% . The most effective itraconazole regimen was 200 mg twice daily for 1 week. In a large double-blind multicenter study of all forms of tinea pedis, De Keyser et al compared 2 weeks of terbinafine at 250 mg/day to 2 weeks of itraconazole at 100 mg/day. After 8 weeks they found terbinafine superior to itraconazole for clinical cure 94.1% vs 72.4%. In a single multicenter open study the cure rate for fluconazole 150 mg was 77% when used once weekly for 3 weeks. (REF:3,4).

RECOMMENDATIONS

American Academy of Dermatology Guidelines recommend topical therapy for initial treatment of tinea pedis. Oral therapy may be required to treat patients with hyperkeratotic soles, disabling or extensive disease, topical therapy failure, chronic infection, or immunosuppression. Surgical therapy is not indicated. (REF:5).

REFERENCES

ID	PUBMED	CORRECT PUBMED	SOR TYPE	PUB TYPE	CITATION
1	19040832				Crawford F, Hart R, Bel-Syer S, Togerson D, Young P, Russell I . Cochrane Review. In: The Cochrane Library, Issue 3, 2001. Oxford: Update Software.
2	20685791				Hart R, Sally E, Bell-Syer S, Crawford F, Togerson D, Young P, Russell I . BMJ 1999; 319: 79-82.
3	20967420				Pierard G, Arrese J, Pierrard-Franchimont C . Drugs 1996; 52: 209.
4	None				De Keyser P, De Backer M, Massart DL, Westelnick KJ . Br J Dermatol 1994; 130: 22-5.
5	20947703				Drake LA, Dinehart SM, Farmer ER ,et al. J Am Acad Dermatol 1996; 34: 282-6.

Annotating Answer Justifications

Conventions for text massaging

1. Remove/edit connecting phrases.
2. Remove irrelevant introductory text.
3. If a paragraph has several references, attempt to split the paragraph.
   - May need to massage the text of resulting splits.
4. If a paragraph has no references, attempt to merge with previous or next paragraph.

Finding PubMed IDs

Method

1. Split the reference text into sentences.
2. Remove author and pagination text:
   - Use simple regexps.
3. Perform a sequence of searches with all combinations of sentences.

Example I

Collins NC . Is ice right? Does cryotherapy improve outcome for acute soft tissue injury? Emerg Med J. 2008; 25: 65-68.

- Collins NC .
- Is ice right?
- Does cryotherapy improve outcome for acute soft tissue injury
- Emerg Med J. 2008; 25: 65-68.

Example II

list	search	ID	title	match %
1, 2, 3	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury? Emerg Med J	18212134	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury?	92
1, 2	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury?	18212134	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury?	100
1, 3	Is ice right? Emerg Med J	18212134	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury?	39
2, 3	Does cryotherapy improve outcome for acute soft tissue injury? Emerg Med J	18212134	Is ice right? Does cryotherapy improve outcome for acute soft tissue injury?	82
1	Is ice right?	None	None	0
2	Does cryotherapy improve outcome for acute soft tissue injury?	15496998	Does Cryotherapy Improve Outcomes With Soft Tissue Injury? 78	
3	Emerg Med J	None	None	0

Using Amazon Mechanical Turk I

Mechanics

- AMT was used to find the correct IDs.
- An AMT hit had 10 references:
  - 2 known references for checking quality of annotation.
- Each hit was assigned to 5 Turkers.
- There was a preliminary training session.

Using Amazon Mechanical Turk II

Approving and rejecting hits

Reject hit if there are two or more "bad" IDs, i.e. one of:

- A known ID is wrong.
- The ID is invalid:
  - Not found in PubMed;
  - No title is returned.
- The title of the ID does not match the title of our reference:
  - threshold: 50% match.
- The ID does not agree with majority.

Using Amazon Mechanical Turk III

Checking validity for final annotation

- Majority wins automatically except when:
  - majority is a "bad" ID;
  - majority is the "nf" ID;
  - the other two are agreeing ("full house").
- Manual check is done in all other cases.

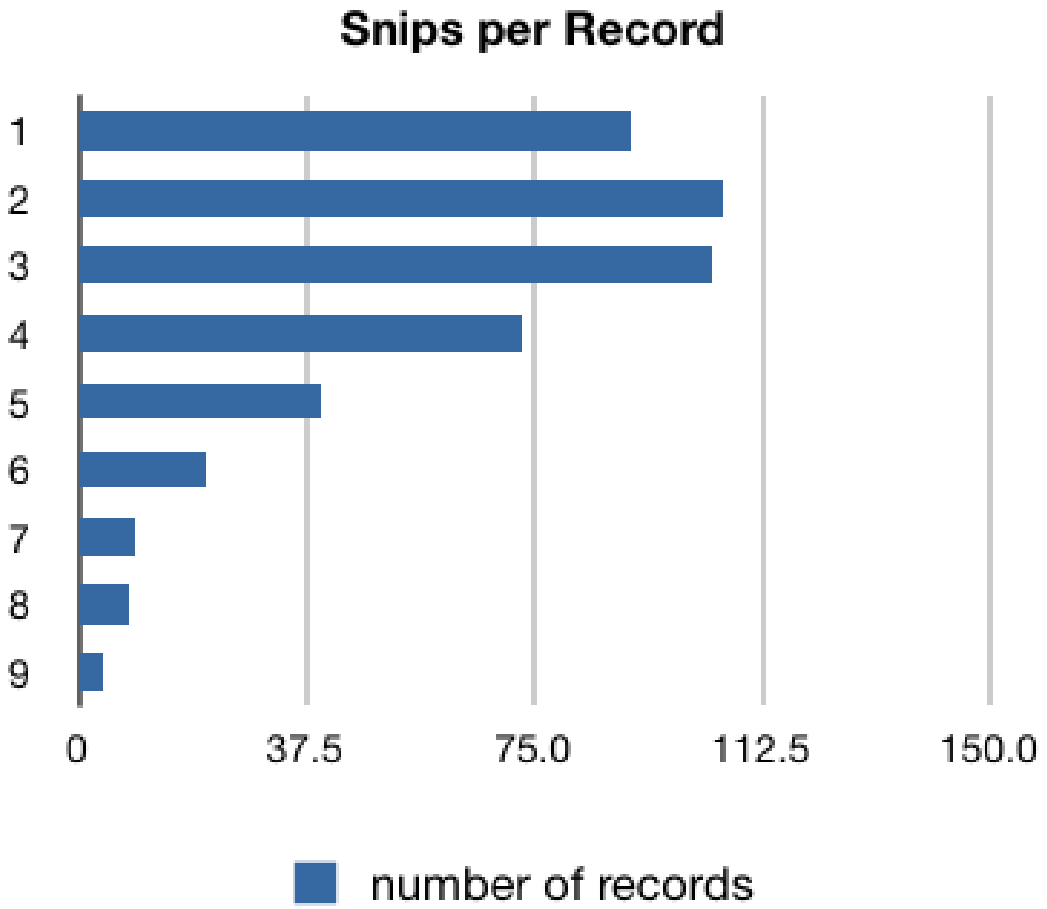
Contents

Corpus Statistics

Size

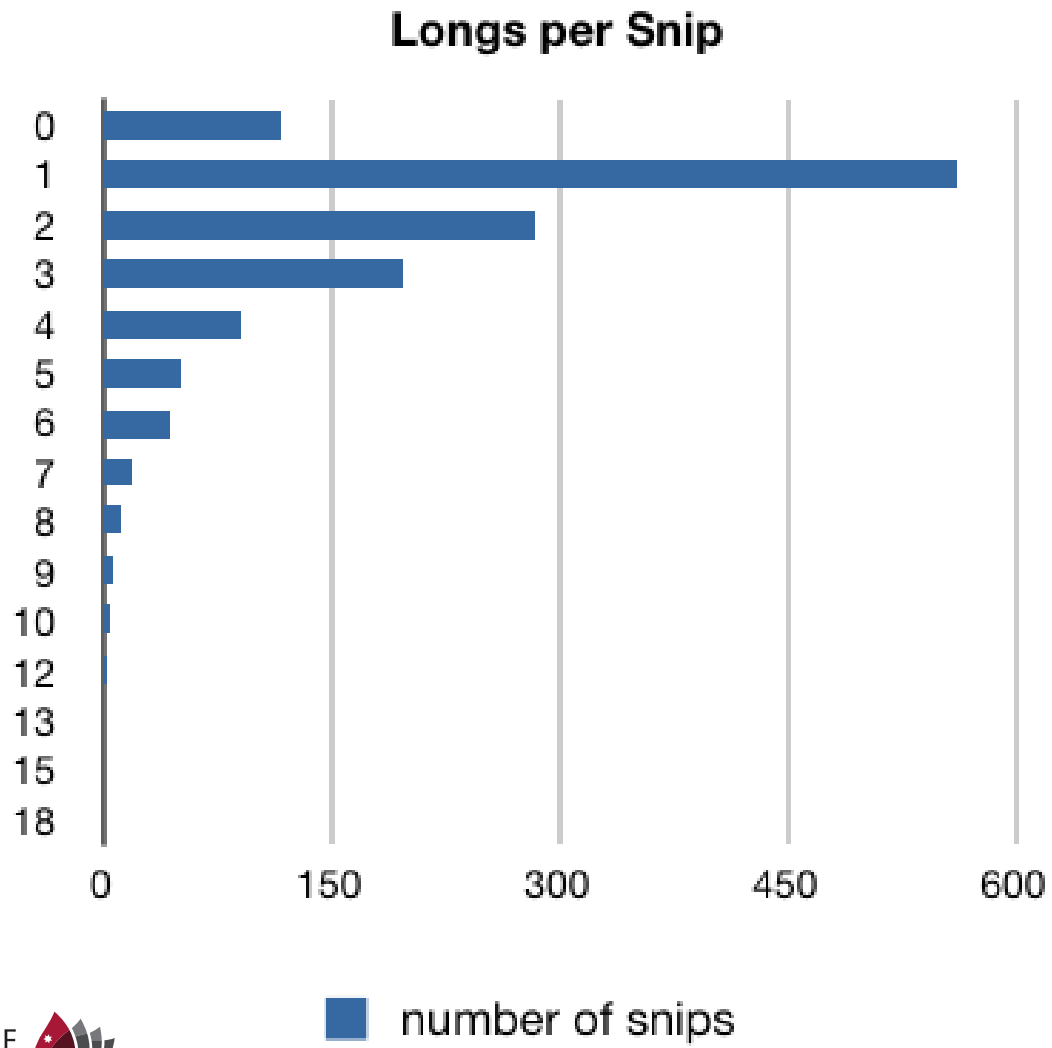
- 456 questions ("records").
- 1,396 answers ("snips").
- 3,036 text explanations ("longs").
- 3,705 references:
  - 2,908 unique references.
  - 2,657 XML abstracts from PubMed.

Answers per Question



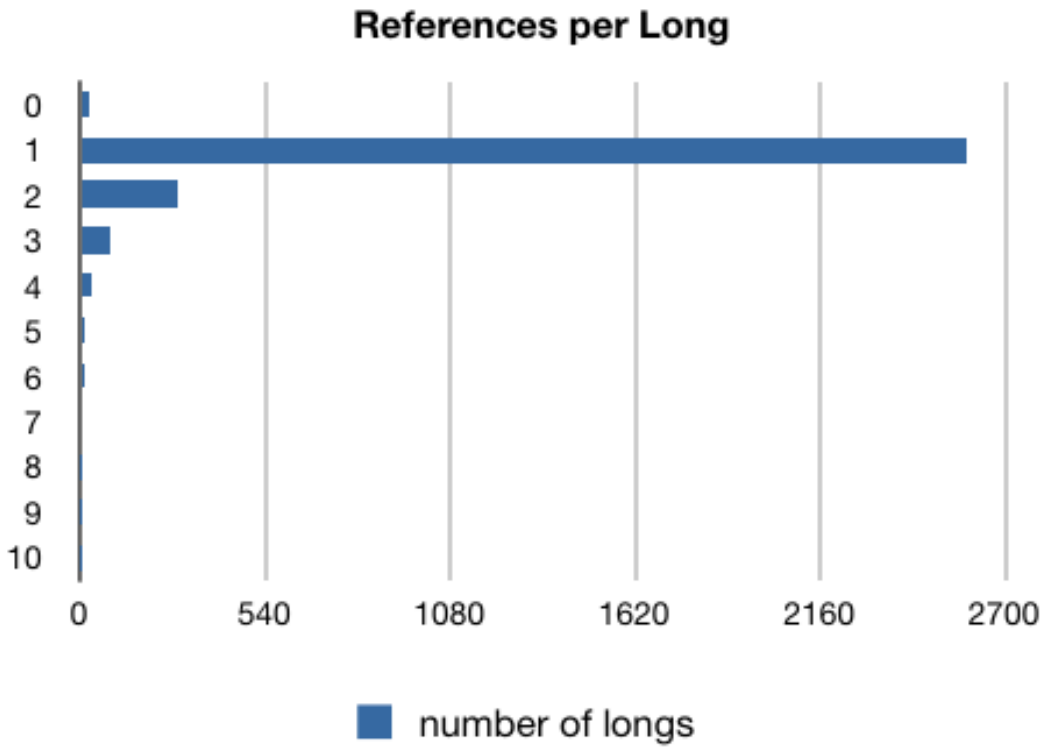
Avg=3.06

Answer justifications per answer



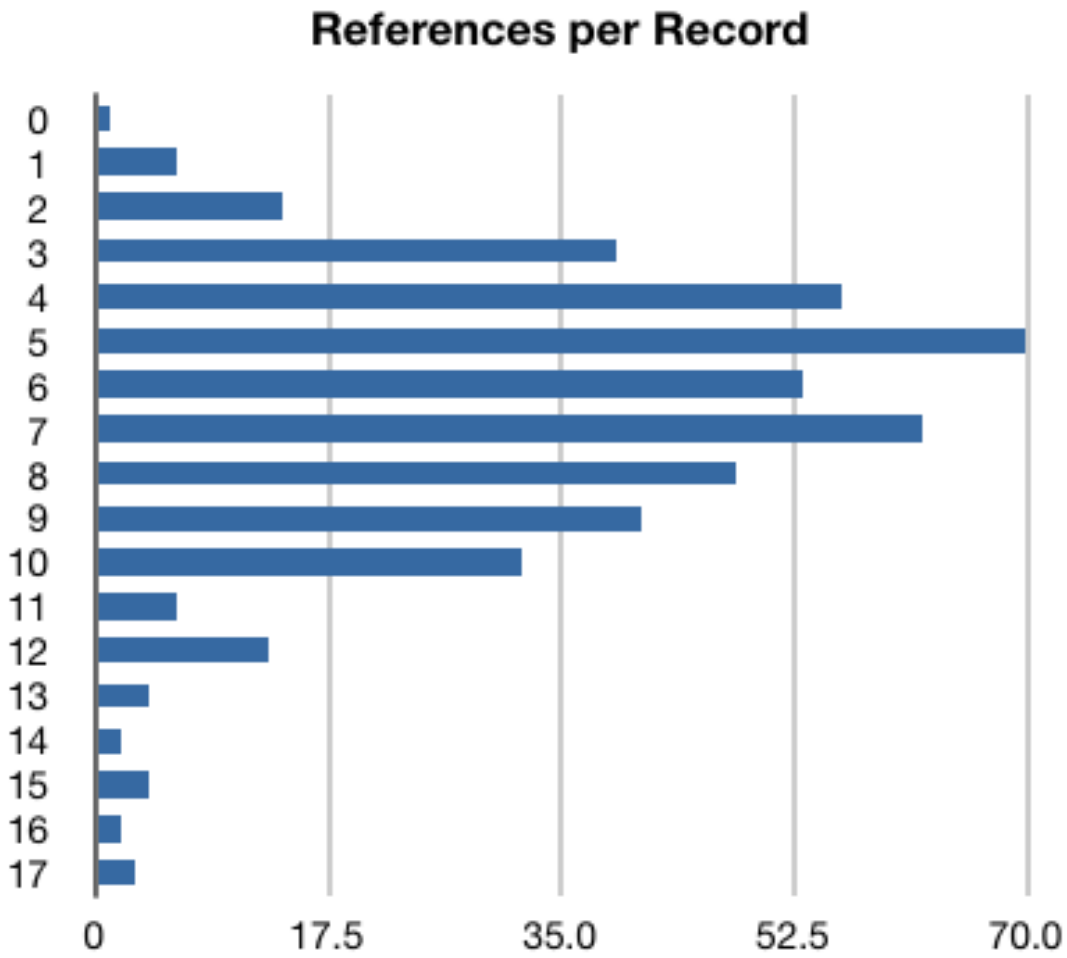
Avg=2.17

References per answer justification



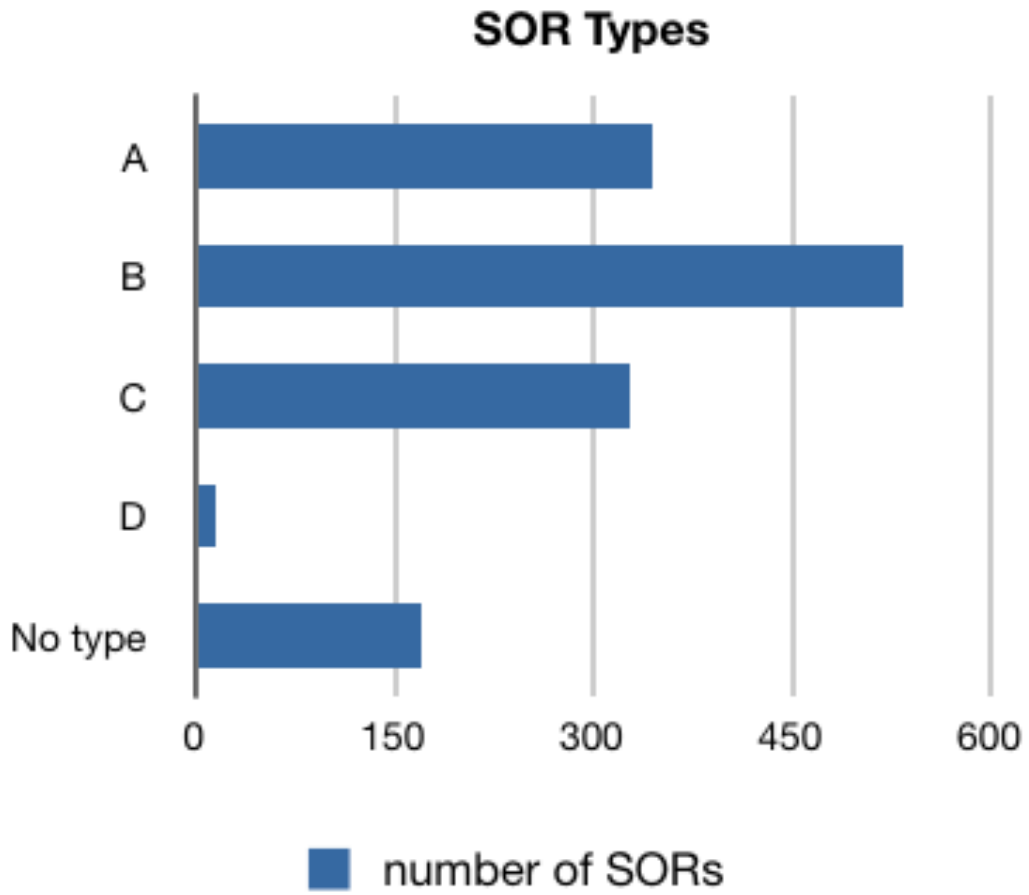
Avg=1.22

References per question

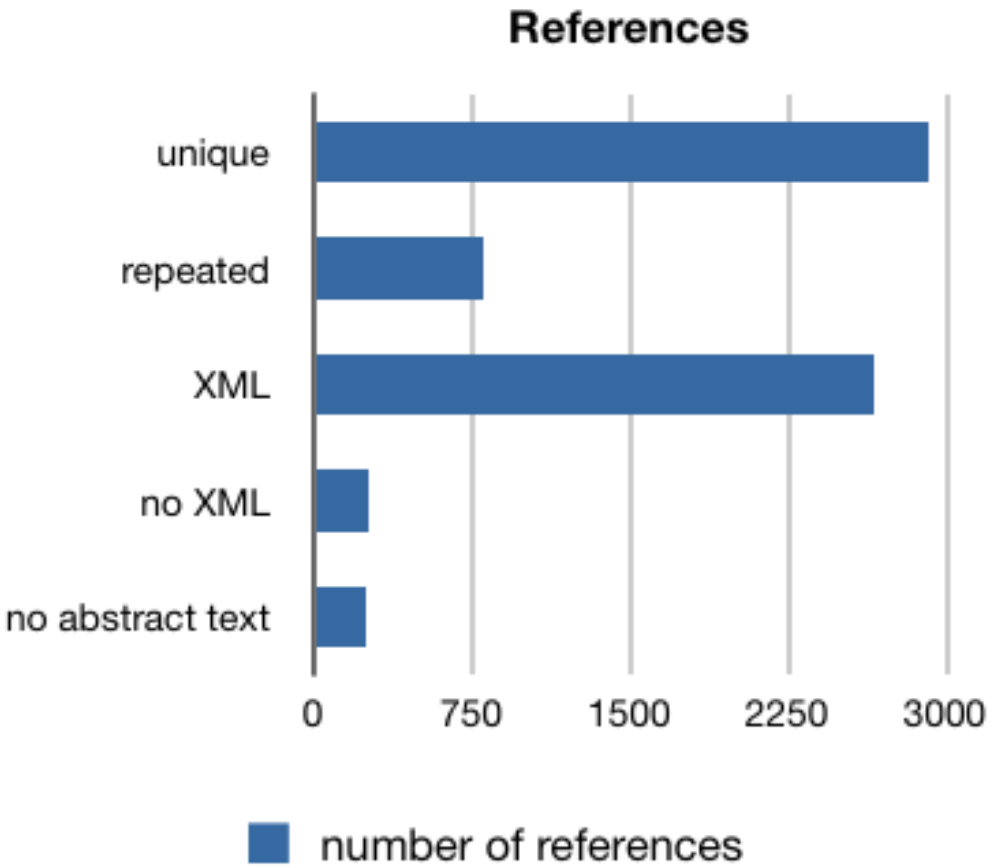


Avg=6.57

Evidence Grade



References


References
unique
repeated
XML
no XML
no abstract text
0
750
1500
2250
3000
number of references

Contents

Contents

Evidence-based Summarisation

Single Document Summarisation

Input: Question, reference.

Target: Text explanation.

Evidence-based Summarisation

Single Document Summarisation

Input: Question, reference.

Target: Text explanation.

Multi-document Summarisation

Input: Question, group of relevant references.

Target: Answer parts (optional: plus text explanation).

Appraisal, Clustering

Text Classification for Appraisal

Input: Group of references.

Target: Evidence-based grade.

Appraisal, Clustering

Text Classification for Appraisal

Input: Group of references.

Target: Evidence-based grade.

Clustering

Input: Question, group of relevant references.

Target: Cluster groupings (optional: plus answer parts).

Retrieval?

Possible task

Input: Question.

Target: List of references.

Retrieval?

Possible task

Input: Question.

Target: List of references.

However...

- Some of the references are old.
- The references are likely not exhaustive.

Contents

Input, Output

Input

- Question.
- Document Abstract.

Output

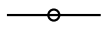
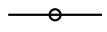
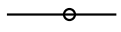
- Extractive summary that answers the question.
- Target summary is the annotated evidence text ("long").
- Evaluated using ROUGE-L with Stemming.

Baselines

plain Return the last *n* sentences.

keywords Return the last *n* sentences that share any non-stop words with the question.

umls Return the last *n* sentences that share any UMLS concepts with the question.

System	F	Conf Interval	
baseline plain	0.193	[0.190–0.196]	
baseline keywords	0.195	[0.192–0.198]	
baseline umls	0.194	[0.190–0.197]	

Using the Abstract Structure

Preselect sentences and then:

Abstract	
Section 1	S1.1 S1.2
Section 2	S2.1
Section 3	S3.1 S3.2
Section 4	S4.1 S4.2
Section 5	S5.1 S5.2

Summary

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).
2. Select the first *n* sentences of the last "conclusions" section.

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary
S5.1 S5.2

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).
2. Select the first n sentences of the last "conclusions" section.
3. If we have less than n sentences, fill from the first sentences of the previous "conclusions" section, and so on until all "conclusions" sections are used up.

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary
S5.1 S5.2 S4.1 S4.2

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).
2. Select the first n sentences of the last "conclusions" section.
3. If we have less than n sentences, fill from the first sentences of the previous "conclusions" section, and so on until all "conclusions" sections are used up.
4. If we have less than n sentences, fill from the "results" sections.

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary
S5.1 S5.2 S4.1 S4.2 S3.1

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).
2. Select the first n sentences of the last "conclusions" section.
3. If we have less than n sentences, fill from the first sentences of the previous "conclusions" section, and so on until all "conclusions" sections are used up.
4. If we have less than n sentences, fill from the "results" sections.
5. If we still have less than n sentences, fill from the "methods" sections.

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary
S5.1 S5.2 S4.1 S4.2 S3.1

Using the Abstract Structure

Preselect sentences and then:

1. Use PubMed's section tags (background, conclusions, methods, objective, results).
2. Select the first n sentences of the last "conclusions" section.
3. If we have less than n sentences, fill from the first sentences of the previous "conclusions" section, and so on until all "conclusions" sections are used up.
4. If we have less than n sentences, fill from the "results" sections.
5. If we still have less than n sentences, fill from the "methods" sections.
6. If the abstract has no structure, return the last n sentences.

Abstract	
Background	S1.1 S1.2
Methods	S2.1
Results	S3.1 S3.2
Conclusions	S4.1 S4.2
Conclusions	S5.1 S5.2

Summary
S5.1 S5.2 S4.1 S4.2 S3.1

Results

The F is calculated using ROUGE-L with stemming.

System	F	Conf Interval
baseline plain	0.193	[0.190–0.196]
baseline keywords	0.195	[0.192–0.198]
baseline umls	0.194	[0.190–0.197]
structure plain	**0.196**	[0.193–0.199]
structure keywords	0.193	[0.190–0.197]
structure umls	0.192	[0.189–0.195]

ROUGE-L with Stemming for All 3-Sentence Subsets I

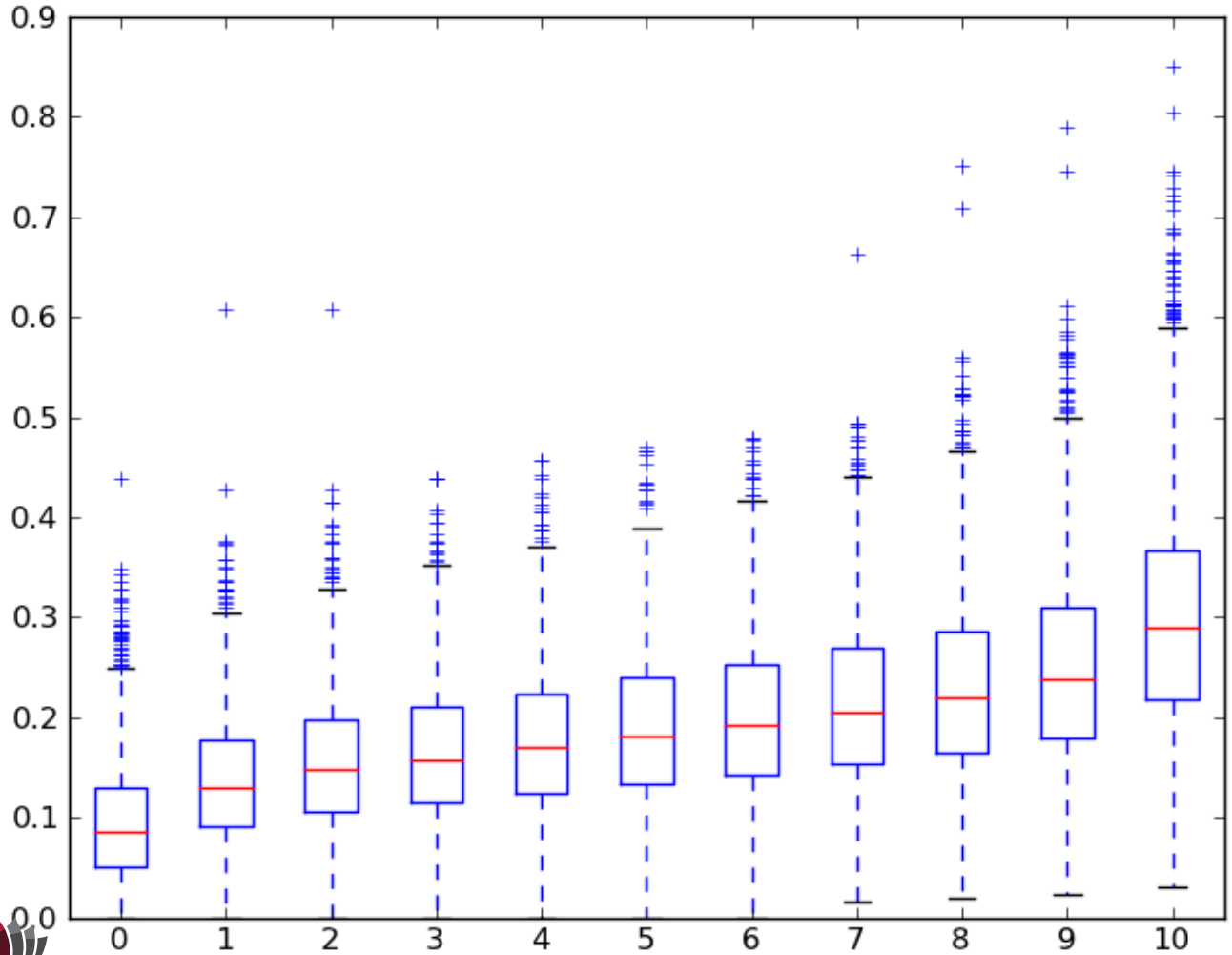
Process

1. Compute the ROUGE-L of all 3-sentence subsets in each abstract.
2. Find the decile boundaries in each abstract.
3. Find the distribution of decile boundaries.

	0	1	2	3	4	5
Mean	0.094	0.136	0.153	0.164	0.176	**0.188**
Std Dev	0.060	0.062	0.065	0.067	0.070	**0.073**

	6	7	8	9	10
Mean	0.200	0.213	0.229	0.249	0.299
Std Dev	0.076	0.081	0.087	0.094	0.112

ROUGE-L with Stemming for All 3-Sentence Subsets II



Contents

ALTA 2011 Shared Task

The ALTA Shared Tasks

- Competitions where all participants are evaluated on the same data.
- The ALTA 2011 shared task was based on evidence grading.

The Data

- Clusters of abstracts.
- The SOR grade of each cluster.

Data Sample

Fragment

```
41711 B 10553790 15265350
53581 C 12804123 16026213 14627885
53583 B 15213586
52401 A 15329425 9058342 11279767
```

Words as Features

Abstract *n*-grams

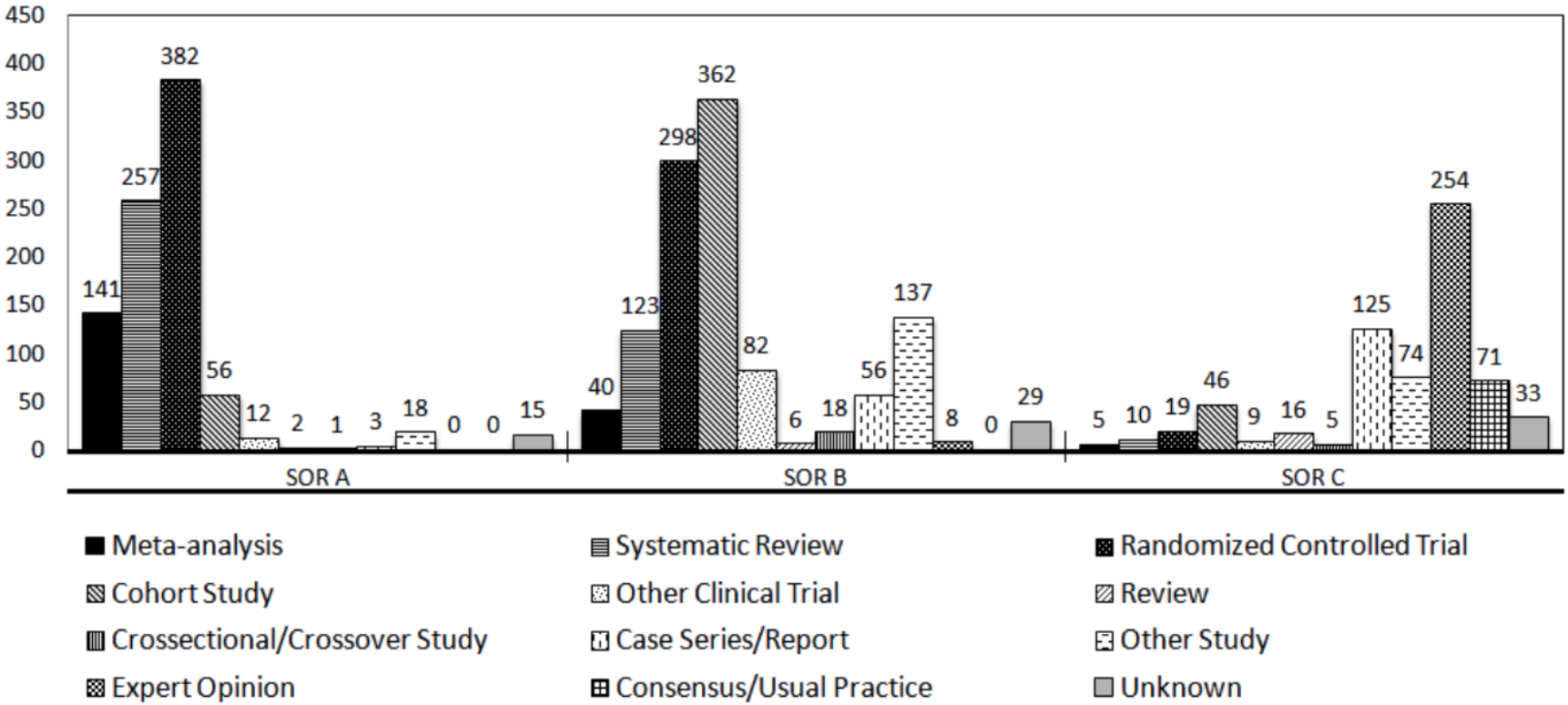
- Generated *n*-grams ($n = 1, 2, 3, 4$) for each of the abstracts.
- Replaced specific medical concepts with generic 'sem_type' tags using UMLS.
- Stemmed, lowercased, stop words removed.

Title *n*-grams

- Generated *n*-grams ($n = 1, 2$) for each title.
- Processed in the same way as abstract *n*-grams.

Publication Types as Features I

Distribution of publication types in a different corpus.



Publication Types as Features II

Publication types

- Rule-based classifier to detect publication types.
- Simple regular expressions that identify major publication types.
- Used the publication types marked up by PubMed when available.
- If an article has several possible publication types, choose the one with highest quality.

Cascaded Classification

Process: Cascaded SVMs

1. Default class: B.
2. SVMs with abstract *n*-grams to identify A and C.
3. SVMs with publication types to identify A and C.
4. SVMs with title *n*-grams to identify A and C.

Results

Method	Accuracy	Confidence Intervals
Majority (B)	48.63%	41.5 – 55.83
Cascaded SVMs	62.84%	

Questions?

Further Information

`http://web.science.mq.edu.au/~diego/medicalnlp/`