Fever Detection in Clinic Visit Notes Using a General Purpose Processor

George Hripcsak, MD, MS, Akiva Bamberger, Carol Friedman, PhD

Biomedical Informatics, Columbia University; New York City Dept of Health & Mental Hygiene

OBJECTIVE

To assess the performance of an unmodified, general purpose natural language processing system to detect fever, and to assess the feasibility of parsing visit notes for syndromic surveillance.

BACKGROUND

With increased penetration of clinical information system products and increased interest in clinical data exchange, a variety of clinician's notes are becoming available for surveillance. Chief complaints have been studied extensively, and emergency department notes have received attention [1], but narrative clinic visit notes have gotten little attention.

METHODS

We selected one year of each of the following from the Columbia University Medical Center clinical data repository: emergency department (ED) visit notes, clinic initial visit notes, and clinic follow up visit notes. These notes were typed by clinicians, although use of the electronic system was optional. We defined current acute fever as fever related to the current visit, that occurred within the past week, and that was not part of a chronic pattern (e.g., intermittent fevers for six months). The clinician's judgment of whether a fever was present took precedence over numeric values, but measured temperatures of 38C or over were considered to be fever if no other mention was made. Because vital signs were logged elsewhere, measured temperatures were only included in notes when the clinician chose to type them in.

MedLEE [2] is a general purpose natural language processing system that uses a lexicon and a semantic grammar to parse and encode medical documents. MedLEE creates a finding (e.g., fever) and modifiers (e.g., certainty of 'negated,' status of 'resolved'). It also indicates when a parse is uncertain. It has been validated on discharge summaries, radiograph reports, and other note types. We did not modify MedLEE for this surveillance study, but simply took its output at face value. We wrote a rule that classifies notes as having a current acute fever, as not having a current acute fever, or as equivocal (those notes for which MedLEE was uncertain). We used stratified sampling and manual review to assess correctness and to characterize how fever is referred to in notes. We used non-parametric estimates of receiver operating characteristic curve area.

RESULTS

Clinic initial visit notes had the lowest performance with the worst specificity (Table 1). ED notes compared well with previous work [1], which achieved an ROC area of .97 (sens .98, spec .89). As shown in Table 2, notes varied in how often a fever occurred, how often there was negation, how long the notes were, and how often a mentioned fever occurred in a previous visit. In general, initial visit notes were more complex and more difficult to accurately determine that current acute fever was present.

Table 1	*	Sensitivity	Specificity	ROC
ED note	−	.76 (.62-.87)	.96 (.94-.97)	**.97**
	+	.94 (.85-.99)	.94 (.92-.96)	
Clinic initial visit	−	.83 (.59-.96)	.89 (.86-.92)	**.94**
	+	.94 (.73-1.00)	.86 (.82-.89)	
Clinic follow up	−	.64 (.46-.79)	.97 (.96-.98)	**.98**
	+	1 (.90-1.00)	.93 (.91-.94)	

*equivocal cases counted as negative (−) or positive (+)

Table 2	ED note	Initial Visit	Follow-up
Current acute fever (%)	**4.34**	**1.81**	**2.16**
Chronic fever	.08	.49	.06
Fever on previous visit	1.34	**2.37**	1.45
Conditional fever (e.g., part of a plan)	.84	.50	.31
Fever stated as being absent or unlikely	10.86	**35.89**	**27.28**
Confused concept (e.g., Scarlet Fever)	.86	1.63	.01
No fever or temp mentioned	81.67	57.31	68.68
Characters per note (N)	809	2623	1766

CONCLUSIONS

(1) Detecting fever in initial and follow up clinic visit notes is feasible, although initial visit notes appear to be more difficult to classify due to length, frequent negation, and frequent reference to previous visits.

(2) Use of a general purpose but sophisticated natural language processing system performed similarly to a special purpose system designed for the task.

REFERENCES

[1] Chapman WW, Dowling JN, Wagner MM. Fever detection from free-text clinical records for biosurveillance. J Biomed Inform. 2004 Apr;37(2):120-7.

[2] Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated encoding of clinical documents based on natural language processing. J Am Med Inform Assoc 2004;11(5):392-402.

Funding: NLM grants LM06910, LM06274, LM8635, & LM7659.