



Natural language processing to identify patient symptoms during and prior to cancer therapy



Julian C. Hong, M.D., M.S.¹, Donna Niedzwiecki, Ph.D.², Manisha Palta, M.D.¹, Jessica D. Tenenbaum, Ph.D.²
 Departments of Radiation Oncology¹ and Biostatistics & Bioinformatics², Duke University, Durham, NC

Purpose/Objective

10-20% of patients with cancer undergoing outpatient radiation therapy (RT) or chemoradiation (CRT) require emergency department (ED) evaluation or hospitalization due to symptoms from treatment, disease, or comorbidities.

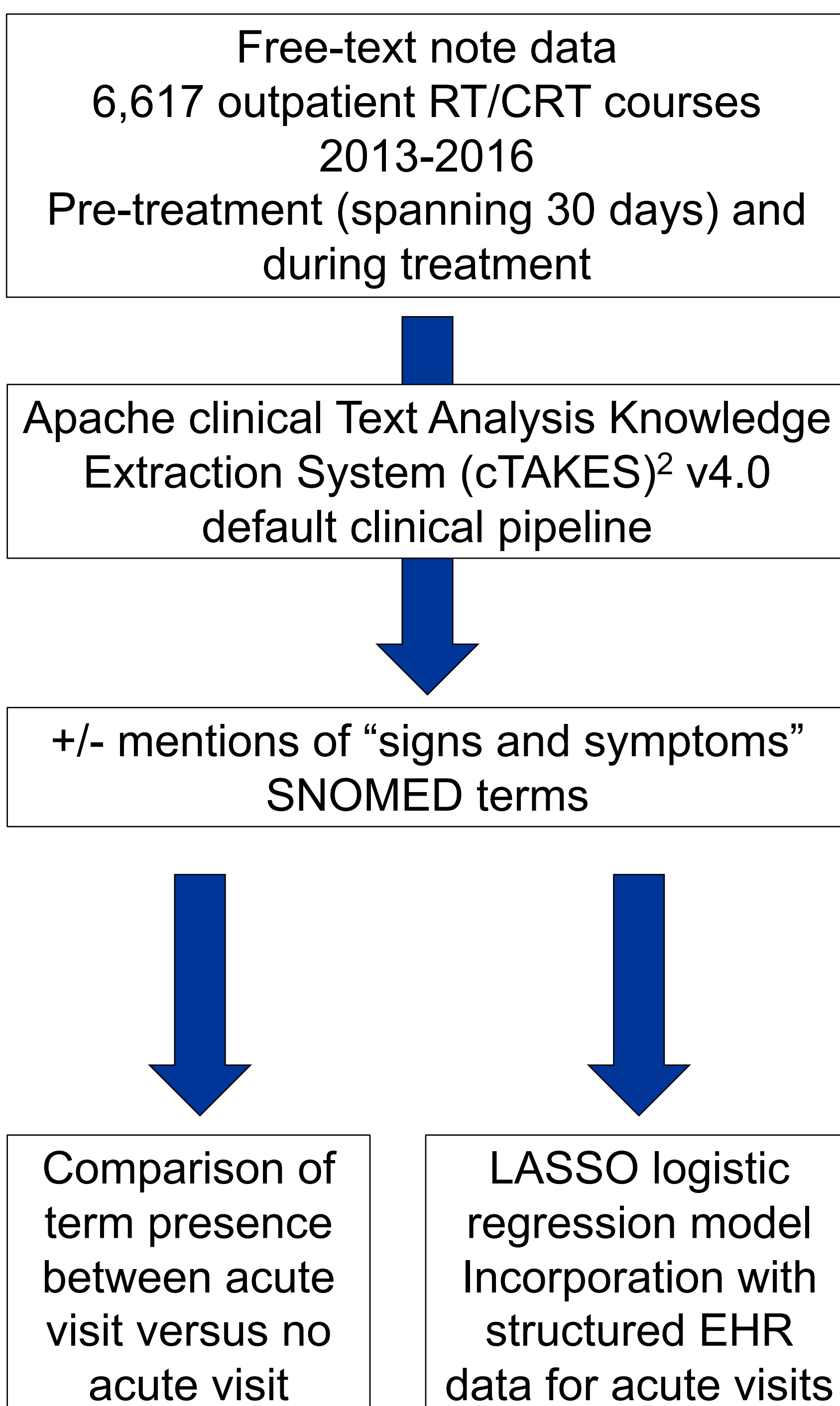
This can impact treatment outcomes and healthcare costs.

Early identification may direct supportive care to prevent these events; an estimated 50% of emergency visits during outpatient cancer therapy are potentially preventable.

We previously developed a pre-treatment machine learning (ML) algorithm to predict ED visits and hospitalization during treatment.¹

We present preliminary analysis of extracting symptom data from notes before and over the course of treatment.

Methods



Results

Differential extracted terms prior to treatment

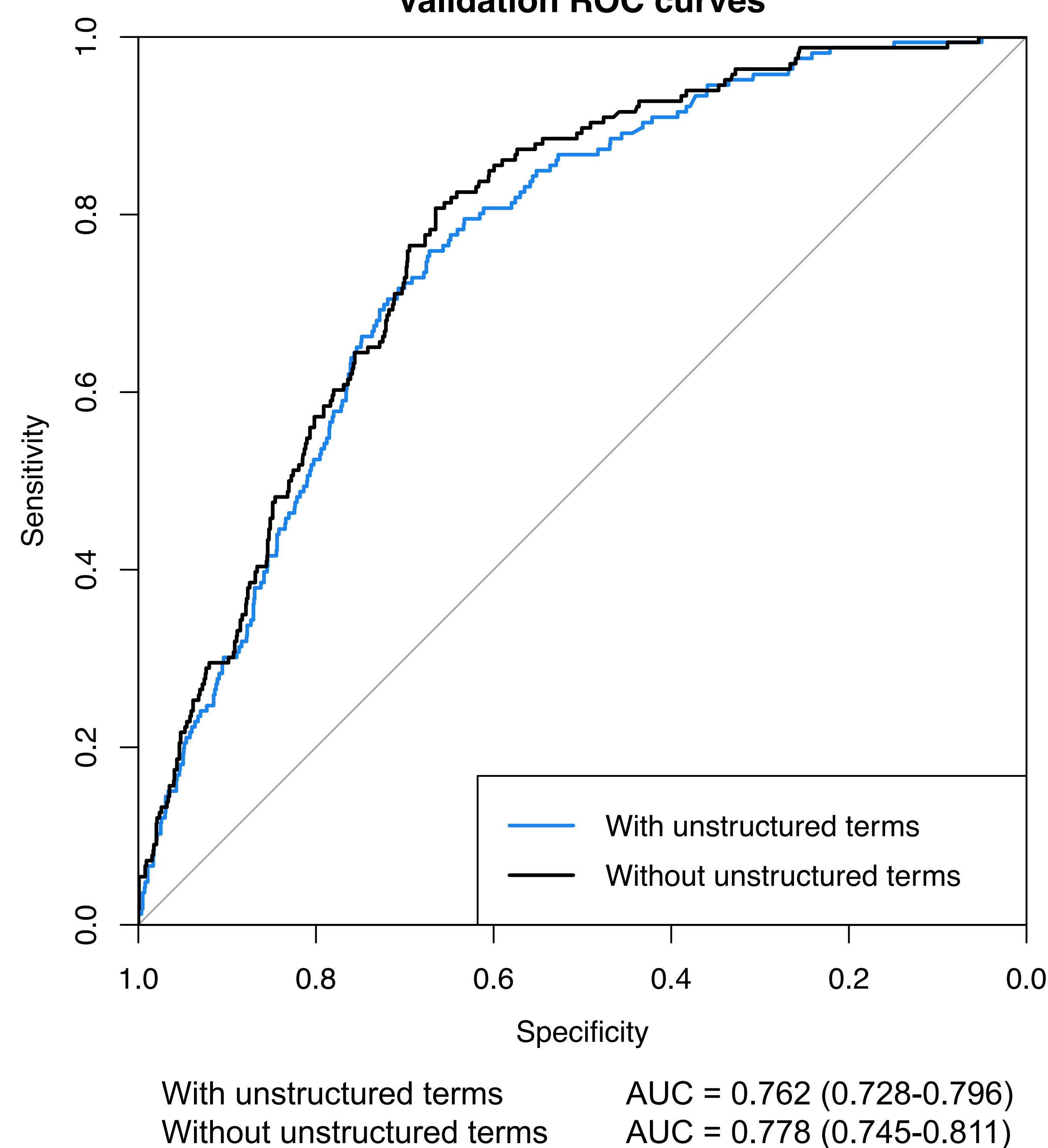
	% of events	% of non-events
+ Nutrition function	40	21
+ Comprehension	76	57
- Jugular venous engorgement	58	40
+ Dilated	42	25
+ Drug Tolerance	42	25
+ Immune Tolerance	42	25
+ Smoker	83	66
+ Patient date of birth	48	32
+ Chest Pain	52	35
- Complication	52	36
+ Weight decreased	52	36
+ Sore to touch	64	49
+ Mental state	43	27
+ Able (finding)	66	51
+ Related personal status	54	39
+ Hemorrhage	54	39
+ Unable	45	30
+ Marriage, life event	74	59
+ Illness (finding)	78	63
+ Skin turgor	67	52

Differential extracted terms during treatment

	% of events	% of non-events
+ Exertion	84	17
+ Affect (mental function)	77	11
+ Smoker	93	27
+ Oriented to person	71	7
+ Bowel sounds	87	23
+ Chills	79	16
+ Oriented to place	89	26
+ Chief complaint (finding)	92	30
+ Chest Pain	78	17
+ Abdominal Pain	73	12
- Wheezing	78	17
+ Family history	92	31
+ Fever	87	27
- Rales	70	11
+ Activities of Daily Living	66	7
+ Used by	82	24
+ Dyspnea	85	27
+ Independently able	64	6
+ Illness (finding)	86	28
- Heart murmur	87	29

(+) denotes documented positive mention
 (-) denotes documented negative mention

Validation ROC curves



Conclusions

Biomedical term extraction can be used to identify “signs and symptoms” before and during the course of outpatient cancer therapy.

Preliminary analysis suggests symptoms experienced during cancer therapy differ in patients who require ED evaluation or admission compared to patients who did not require acute care. Increased documentation of normal findings was associated with acute events.

Extracted pre-treatment terms did not improve predictive performance of a LASSO logistic regression model above structured data alone.

Ongoing work focuses on evaluating the accuracy of term extraction against physician chart review, specifically for Common Terminology Criteria for Adverse Events (CTCAE) v4.0.

Predictive value of term extraction prior to and during treatment will continue to be explored to improve discriminative and predictive power prior to ED visits and admissions.

References

- Hong JC, Niedzwiecki D, Palta M, Tenenbaum JD. Predicting toxicity in radiation and chemoradiation: An internally validated pre-treatment machine learning algorithm for emergency visits and hospital admissions. *JCO Clin Cancer Inform.* 2018 Aug 30; (2):1-11.
- Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc.* 2010 Sep;17(5):507-13.