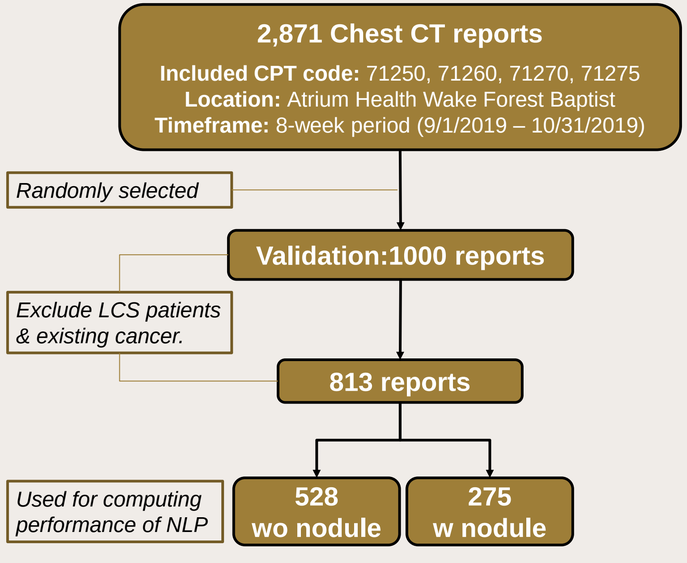
Natural Language Processing (NLP) provides pathway to more reliably identify Incidental Lung Nodules

T. L. Dotson M.D.[1], A. Gasimova[2], J. Watkins[2], Q. Chometon[2], C. R. Bellinger M.D.[1]
[1]Pulmonary and Critical Care Medicine, Wake Forest University School of Medicine, NC, US, [2]Optellum Ltd., Oxford, UK

RATIONALE

- Approximately 31% of chest CT examinations report an incidental lung nodule (ILN) and studies estimate that up to 60% of patients with ILNs are lost to follow-up.
- One proposed mechanism for the high proportion of patients without recommended nodule follow-up is lack of personnel to manage longitudinal surveillance. Health care systems need to automatically and accurately identify ILNs for appropriate follow-up.
- **Objective**: Evaluate an automated method of identifying lung nodule patients from free-text CT radiology reports. Once identified, patients can be tracked by a dedicated nodule navigator.

DATA COLLECTION



METHODS

- The NLP algorithm – Patient Discovery – was developed by Optellum Ltd. for use in their Virtual Nodule Clinic platform.
- The NLP was designed to flag radiology reports containing a measured pulmonary nodule in order to aid in patient identification and tracking.
- A positive report was defined as a report with at least one measured pulmonary nodule.
- Two independent reviewers annotated the reports, and a third physician examined the disagreements to provide a tie-breaker. The annotation was solely made on the radiology report, the CT images were not reviewed as part of the process.

RESULTS

- 813 radiology reports evaluated and 275 (33.8%) had at least one measured nodule Of these, 149 were from outpatient / 75 from emergency / 44 from inpatient / 7 unknown.
- NLP algorithm achieved sensitivity of 1.00, specificity of 0.98, PPV of 0.96, and NPV of 1.00 across all reports.

		NLP prediction	
		Neg	Pos
Actual	Neg	527	11
	Pos	1	274

- ER subgroup, the NLP identified 79 reports with nodules (4 FP). 58% of the reports with a nodule were discovered on a CTA. 43/79 patients (54%) had a reported nodule between 6 and 30 mm and 21/79 patients (27%) have a reported nodule between 8 and 30 mm.

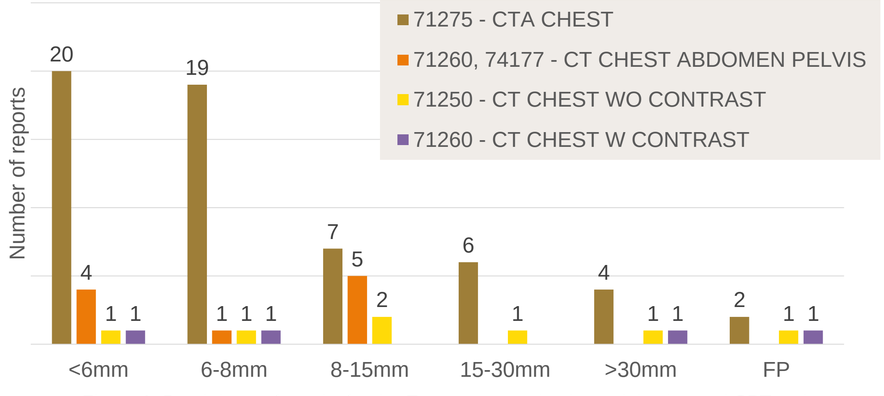


Figure 1: Breakdown of results for the Emergency setting by nodule size and CPT codes.

	Positive (%)	Negative (%)	Sensitivity	Specificity	PPV	NPV
All reports	275 (33.8%)	538 (66.2%)	1.00	0.98	0.96	1.00
Care Setting						
Emergency	75 (22.0%)	266 (78.0%)	1.00	0.98	0.95	1.00
Inpatient	44 (24.7%)	134 (75.3%)	1.00	0.99	0.96	1.00
Outpatient	149 (55.6%)	119 (44.4%)	0.99	0.96	0.97	0.99
Exam description						
CTA Chest	79 (22.7%)	269 (77.3%)	1.00	0.99	0.98	1.00
CT Chest WO contrast	155 (51.0%)	149 (49.0%)	1.00	0.95	0.96	1.00
CT Chest W contrast	30 (39.0%)	47 (61.0%)	0.97	0.96	0.94	0.98
CT Chest Abd Pelvis	11 (13.3%)	72 (86.7%)	1.00	1.00	1.00	1.00

Table 1: Breakdown of results by care setting and exam description (see link to CPT codes in Figure 1). PPV and NPV are positive predictive value and negative predictive value

CONCLUSIONS

- The NLP algorithm was able to identify patients with incidental lung nodules from CT radiology reports with high sensitivity and specificity.
- NLP has the potential to create a more streamlined process to identify patients across different care settings, reduce lost to follow-up nodules, and improve clinical guideline adherence.

DISCLOSURES & ACKNOWLEDGEMENTS

Funding: Optellum Ltd