CLINICAL IMPLEMENTATION OF AN ARTIFICIAL INTELLIGENCE-DRIVEN PULMONARY NODULE DETECTION QUALITY ASSURANCE PROGRAM IN THE EMERGENCY DEPARTMENT SETTING

Irene D. O. Santo, MD¹, Joseph J Cavallo MD, MBA¹, Jonathan L Mezrich, MD, JD, MBA¹, Howard P Forman, MD, MBA¹,²,³

1. Yale Department of Radiology and Biomedical Imaging, Yale New Haven Hospital; 2. Yale School of Management, Yale University; 3. Yale School of Public Health, Yale University
**INTRODUCTION**

- Incidental lung nodules:
  - 8.5% in trauma patients; 31% general adult population
  - **Not time sensitive in acute care setting** but important to overall care
- Partnership with our vendor AIDOC (Tel Aviv-Yafo, Israel)

---

**Goal:** Design & implement an ED QA program to address nodule detection in a time-efficient & nonintrusive manner

| Increase detection rate of incidental pulmonary nodules | Ensure **appropriate outpatient follow-up** of clinically significant pulmonary nodules | Minimizing interruptions to ED Radiologist workflow | Operate autonomously in background (no radiologist involvement necessary at time of interpretation) |
Vision-based AI algorithm to analyze images

- Deep Convolution Neural Network (CNN)
- Trained on tens of thousands CTs
- CT scanners from multiple medical centers around the world
- Produces a 3D segmentation map
- Region proposals are generated from segmentation map
- Second stage classifies each region as positive or negative
- Detects all types of pulmonary nodules (solid, sub-solid and ground glass)
- Trained on 6 mm to 3 cm nodules

Natural language processing tool (NLP) to analyze radiology reports

- Analyzes finalized radiology reports
- Groups them by report characteristics, as positive or negative for pulmonary nodules
- Set of structured rules by domain experts
- Report provided as a raw text file and transformed to a standard form
- Removal of special characters and duplicates, word tokenization, section parsing, and sentence tokenization
- Binary classification (mentions nodule or not)
- If + by the AI algorithm and - by the finalized report (by NLP) = suspected discrepancies

Semi-automated email notification system to alert physicians & track relevant studies

- Suspected discrepancies are then processed using a semi-automated email notification and tracking system
- Allows for a secondary review of images
- Issuance of addendum (if necessary) by the radiologist of record
- No explicit instructions were given for the secondary review, but radiologists within the department customarily utilize the Fleischner criteria to determine when follow-up imaging should be recommended

Emergency setting at our academic tertiary center radiology department between October 1, 2021, and June 1, 2022

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED CT scan with any visible lung anatomy</td>
<td>Motion artifact</td>
</tr>
<tr>
<td>Performed on CT scanner with ≥ 64 detectors</td>
<td>Severe metal artifact</td>
</tr>
<tr>
<td>Patient age ≥ 18</td>
<td>Inadequate field of view</td>
</tr>
</tbody>
</table>
**PROCESS MAP OF THE AI-DRIVEN LUNG NODULE DETECTION QUALITY ASSURANCE PROGRAM IN THE EMERGENCY DEPARTMENT SETTING**

### During Emergency Department Visit
- **ED patient undergoes a CT scan which includes the lungs**
- **ED radiologist reviews the study**
- **Report Issued to provider and patient**

### After Patient Encounter
- **AIDOC NPL tools assess whether algorithm-detected lung nodule has been reported**
- **For exams with no nodule reported, AI-Based Imaging Software analyzes images for the presence of pulmonary nodules**
- **Discrepancies between NLP output and Image analysis trigger QA email to the attending**

### Following QA Email Alert
- **Secondary review of the relevant images by radiologist**
- **If necessary, addendum is issued with or without new imaging recommendations.**
- **Response to QA supervisor closes the loop**
19,246 CT scans met inclusion/exclusion criteria

15,201 identified as negative by NLP of final report

Images were analyzed by image-based AI algorithm

50 suspected discrepancies (0.26%)

Addenda issued in 34 patients (68%)

Mean time from original report to addendum: 3 days

1 too small, 4 stable, 1 reported in concurrent study

16 exams not addended

Average size of nodule in which an addendum was issued: 5.6 mm

10 FP: 9 by the AI nodule detection software, 1 by NLP

Additional imaging recommended in 20

13 explicit, 6 optional, 1 already scheduled for a FU

Pearson's chi-squared test of independence: significant association between CT protocol and addendum rate (p = 0.009012, p<0.05): >150% increase in addendum rate for CT Abdomen Pelvis protocols compared to Chest protocols
# Patient Demographics

## Gender
- Male: 22
- Female: 28

## Malignancy History
- Yes: 9
- No: 41

## Smoking History
- No: 19
- Previously: 23
- Current: 8

## Age
- <50: 11
- 51-60: 14
- 61-70: 13
- 71+: 12

## Ethnicity
- Caucasian: 34
- African American: 7
- Asian: 2
- Other/Not Specified: 7
**RESULTS**

**STUDIES ANALIZED**

<table>
<thead>
<tr>
<th>Anatomical Region (CT)</th>
<th>Processed Exams</th>
<th>Secondarily Reviewed</th>
<th>Addenda Issued</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdomen Pelvis</td>
<td>9572</td>
<td>38</td>
<td>29</td>
</tr>
<tr>
<td>Chest</td>
<td>7189</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Chest Abdomen Pelvis</td>
<td>2136</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Chest Abdomen</td>
<td>336</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**LUNG NODULES**

<table>
<thead>
<tr>
<th>Size (mm)</th>
<th>Nodules</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub 6mm</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>6 - 8</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>&gt; 8</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Breakdown of pulmonary nodules that resulted in addenda being issued after secondary image review, stratified according to the Fleischer Society Guidelines. Also shown is the number of nodules in each category that resulted in new follow up imaging recommendations.
• **Non-intrusive**: only 0.26% of all CTs being classified as discrepant and requiring secondary radiologist review

• 19 patients will receive the **recommended follow-up imaging** (would have otherwise been missed)

• Approximately one third of missed nodules may be clinically significant

• **Diminish medicolegal risk** to radiologists and hospitals

• Although a minority of incidentally identified nodules turn out to be malignancies, the potential for missing cancer at an early stage, both in terms of patient outcome and malpractice judgment awards, can be enormous

• **Time saved** by a QA program that functions almost entirely in the background are substantial

• An estimated total of **3531 unnecessary notifications were avoided**
Odds of a nodule being initially missed by the radiologist was **higher for studies without dedicated CT Chest imaging**

Average time to addendum of **4330 minutes (around 3 days)**: acceptable for nonurgent finding

Elevated SD of the time to addendum (15018 minutes) was attributed to a single outlier due to an **incorrect email address**

---

**Discussion**

- Clinical significance of nodules is not yet known (brief period of this evaluation: 7 months)
- No universal secondary reads = no data on true sensitivity of AI software (still under investigation)
- Nodules missed by the radiologist & image-based AI software → Yes/No? How many?
- NLP system cannot discriminate between nodules of potentially varied significance
- Inherent risk of losing a patient to follow up once they leave the ED following their acute care episode
- Evaluation for stability was limited to previous studies available in our system
• **Sustainable**: No foreseeable obstacle to sustaining this QA program in our institution

• Our goal to eventually study the **impact of this program on patient outcomes**

• Minimal time requirement devoted to supervising notification system (**less than 10 minutes per week**)

• **Generalizable to other institutions**

• Combined AI/NLP program such as the one described herein provides a **nonobtrusive means for QA**

• Such time-shifted applications of AI may have advantages throughout radiology, and even in other medical specialties, provided that they are deployed within appropriate systems to **ensure adequate patient communication** and **follow-up**