More Bang for the Healthcare Buck: Addressing the Emergency and Urgent Care Elephant in Pulmonary Nodule Follow-up Loss with Smart HL7 Workflow

Kyle Jackson
Valeria Makeeva, MD
Michelle Mott, MS
Nabile Safdar, MD, MPH
Peter Harri, MD

Abstract: M5B-QI-5
LIST ALL DISCLOSURES FROM PAST 12 MONTHS

**Presenter:** Kyle Jackson  
Company Name: Emory Healthcare

**Co-author:** Valeria Makeeva  
Company Name: Emory Healthcare

**Co-author:** Peter Harri  
Company Name: Emory Healthcare

**Co-author:** Nabile Safdar  
Company Name: Emory Healthcare

Patent pending on related technique for listed authors
Background

• The overwhelming **majority of incomplete** imaging incidental findings **follow-up** results from **emergency (ED)** and **urgent care (UC) visits**.
• While there have been extensive talks about the need for automated imaging tracking and provider notification systems in healthcare, the **complexity of the ED/UC setting has prevented hospitals from addressing this majority** contributor to lost follow-up.

Objective

• We present **ED/UC-capable system architecture managing HL7 workflow for pulmonary nodule incidental findings** that may be paired with a knowledge-based AI or be used in conjunction with a nurse navigator, depending on hospital infrastructure.
• We address lack of high-level performance evaluation of published tools in the clinical setting by **evaluating clinical performance** within a large academic medical center.
METHODS AND MATERIALS

Follow-up loss

Monitoring at a southeastern academic medical center interpreting 1.2 million imaging exams annually revealed most lost pulmonary nodule follow-up exams were from ED and UC settings.

Automated System Configuration

Novel ADT, ORU, ORM and SIU data streams were designed to configure to all types of clinical recommendations.
AI Study Design

- We previously reported on a LSTM AI model recognizing pulmonary nodule follow-up recommendations (1,970 retrospectively annotated reports, 60/20/20% training/validation/test split, accuracy 94%).

- To generate ED/UC-capable workflow, we paired this LSTM with a knowledge-based AI to assign patients into follow-up intervals: high risk (≤ 3 months), medium risk (> 3 months but ≤ 6 months), and low risk (> 6 months and ≤ 12 months).

- Clinical performance was assessed by retrospective manual validation of 200 reports followed by key word inclusion/exclusion modifications. Real-time clinical performance was again assessed by prospective manual validation of 200 reports.

- Fleischner Society Guidelines and logic governing NLP system-to-EMR communication were leveraged to develop risk-stratified ED/UC-capable HL7-based workflow.
AI performance evaluation

- F0.5 score was selected to optimize clinical practice performance by penalizing false positives more than false negatives, where false reminders would undermine provider system use.
- The macro-averaged F0.5 score for retrospective performance was 0.79 (accuracy by risk class: high risk 94.5%, medium risk 99%, low risk 97%) and for prospective performance was 0.93 (accuracy by risk class: high risk 98.5%, medium risk 100%, low risk 95%).
- Knowledge-based AI modifications resulted in 14% improvement in prospective performance.
Strengths

- Prospective study design
- Real world performance evaluation for AI models
- Explainable AI (XAI) as opposed to "black box" approach of typical machine learning models (e.g. deep neural networks)
- AI system automation with human-in-the-loop reminders
  - Physicians receive reminders and can verify accurate predictions for pulmonary nodule follow-up recommendations originating from ED and UC patient care settings.

Limitations

- Small sample size
- Single healthcare system

Conclusion

- There are no published automated system workflows addressing the safety gap presented by ED/UC settings.
- The proposed ED/UC-capable HL7 logic and knowledge-based AI can help hospitals implement tracking and reminder systems capable of capturing the overwhelming majority of lost follow-up and meet ACR-issued Closing the Recommendations Follow-Up Loop measures.
- High-level clinical performance evaluation underscored the validity of this approach.