

INTRODUCTION

- Identification and follow-up of incidental pulmonary nodules (IPNs) is important for early detection of lung cancer.
- IPNs represent the majority of clinically relevant nodules, outside of lung cancer screening programs yet, the majority are not followed up.
- Developing an accurate automated IPN identification tool is imperative to ensure appropriate follow-up and ultimately, improve early diagnosis of lung cancer.

METHODS

- As part of the **Identifying Early Lung Cancer In A Diverse Population (IDEAL)** study, a multi-centre national trial, the Vancouver site employed a **Medical Named Entity Recognition Model (NER)**- SapienNER v3a (Fig. 1) to identify pulmonary nodules in reports of all ‘lung-in-view’ CT scans completed between November 2023-July 2024 within 3 large hospitals in the Greater Vancouver area.

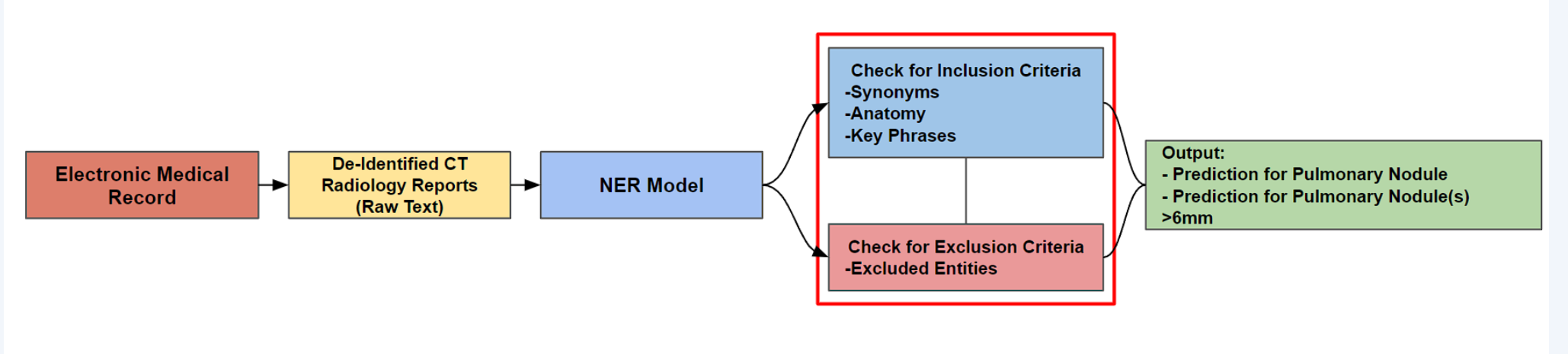


Fig 1. Summary of the automated data workflow for analyzing de-identified CT radiology reports to extract relevant entities and make predictions on pulmonary nodules using the Pulmonary Nodule Inclusion/Exclusion Criteria

- The reports were categorized by age ≥ 50 to 80 years, and nodules ≥ 6 mm then reviewed by clinical staff to ensure the nodules were truly incidental.
- Patients were contacted to arrange nodule follow-up.

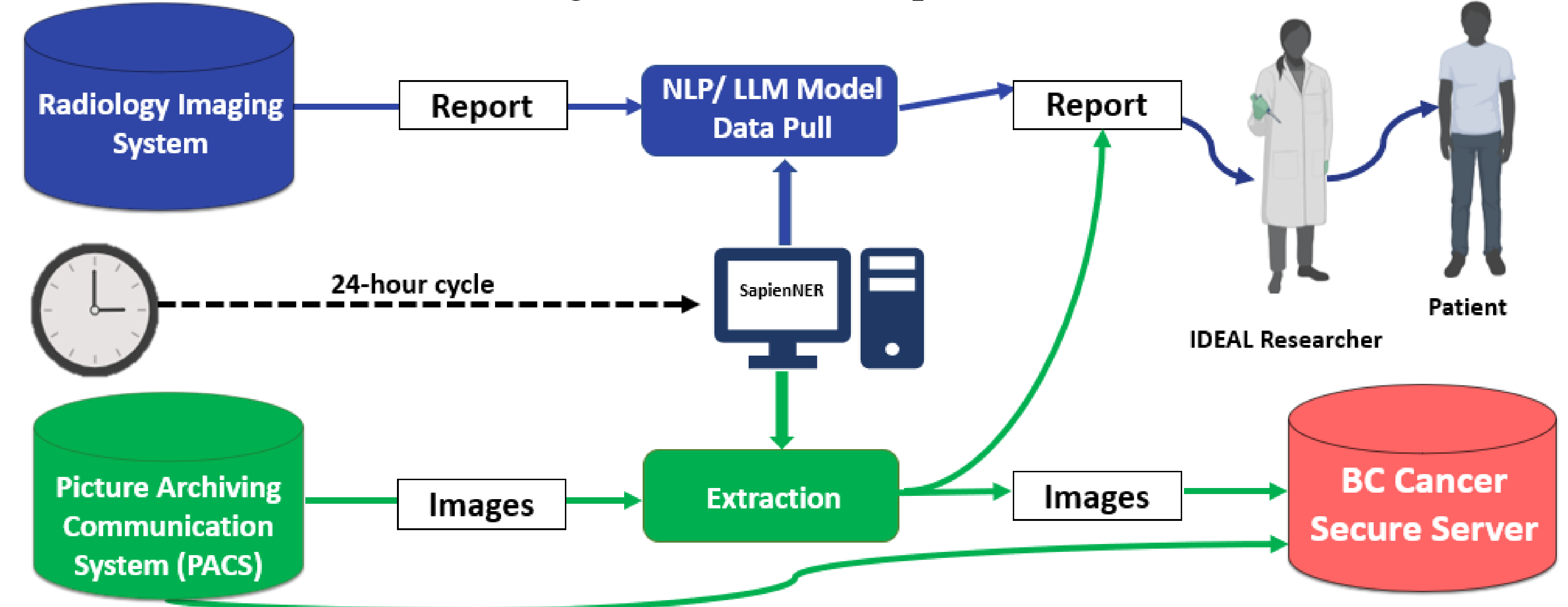


Fig 2. Workflow of CT acquisition from Hospital to BC Cancer

RESULTS

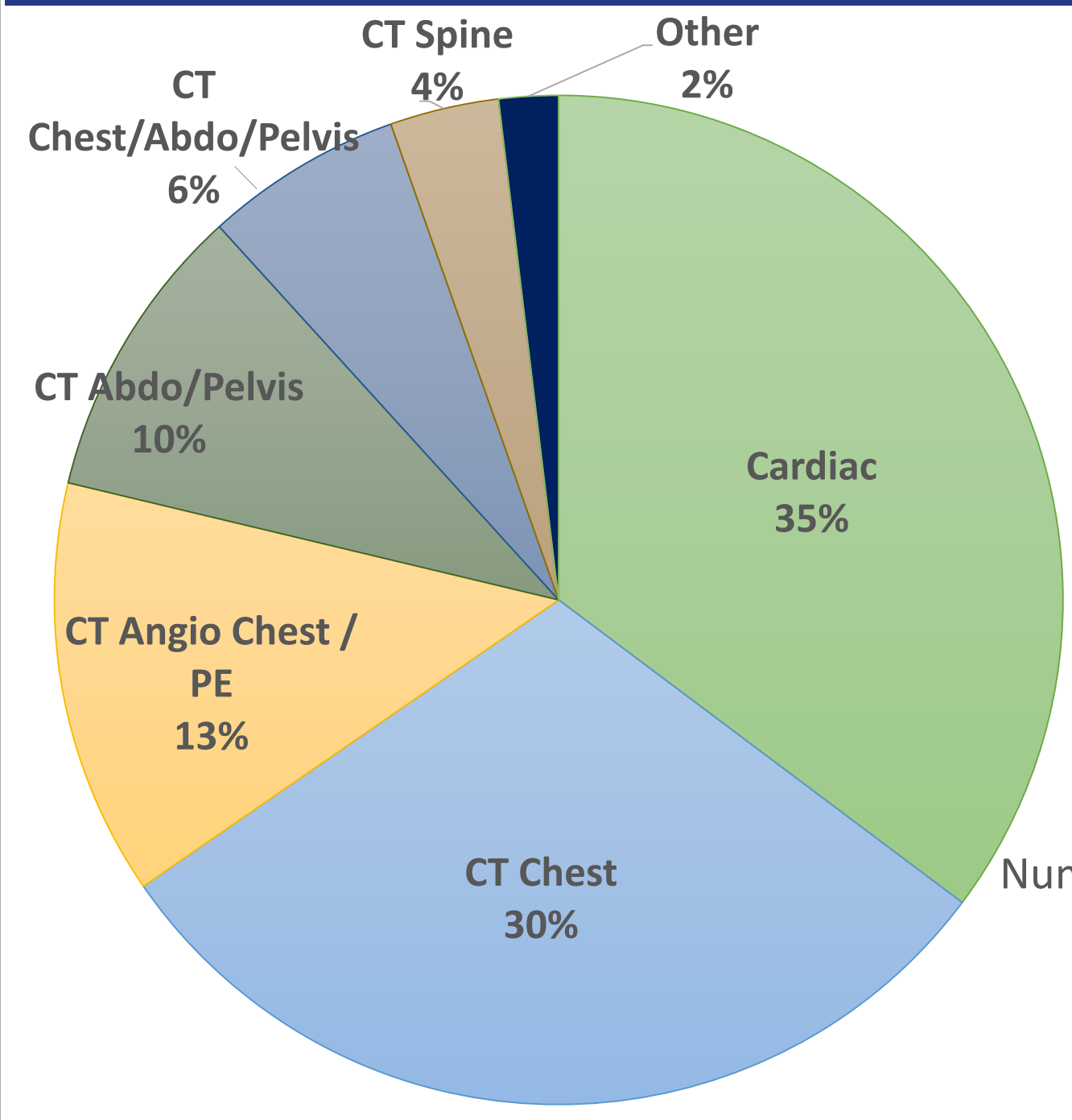


Fig 3. Incidental pulmonary nodule baseline scan type

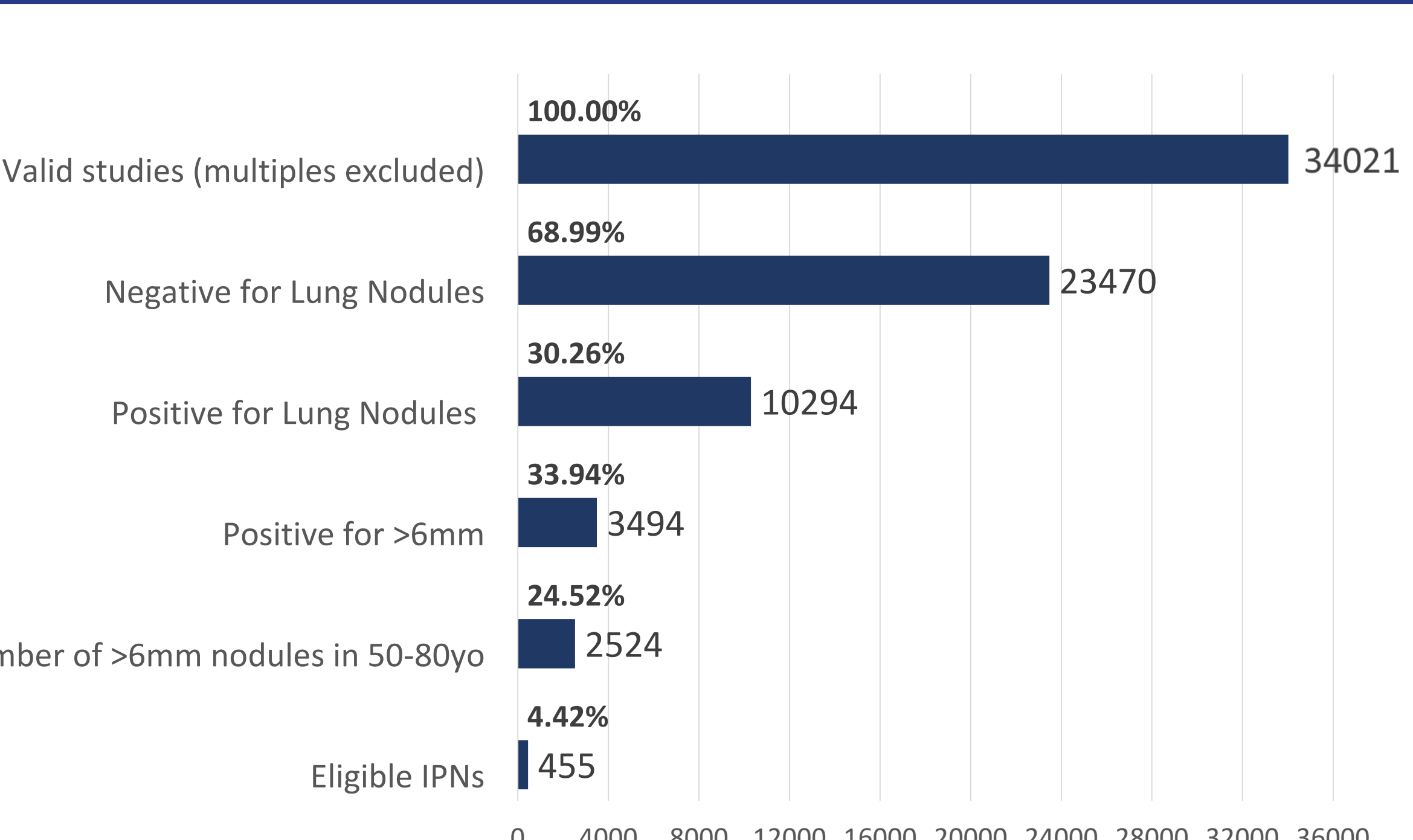


Fig 4. Classification of Pulmonary nodules from all lung in view scans

RESULTS

- In the first 8 months, **4.4%** of scans had clinically relevant (≥ 6 mm) incidental nodules among individuals aged 50-80 yrs.
- The majority of patients were without guideline-based follow-up planned outside of the IPN study.
- Cardiac scans were an important source of IPNs, highlighting potential pathways for collaboration to improve IPN management.

CONCLUSION

- This study demonstrates how a *Natural Language Processing (NLP)* system can be easily deployed in a large healthcare system to reliably and efficiently screen large numbers of exams to identify patients with clinically significant IPNs.
- A *Large Language Model* that can compare prior imaging studies and extract relevant medical history is being developed to improve the accuracy and efficiency of data extraction to identify patients with true IPNs who require clinical follow-up.

Acknowledgement

