Evaluating artificial intelligence for lung cancer detection on chest radiographs: multi-vendor comparison in a retrospective real-world UK population

A Maiter[1,2], P Metherall[1], J Taylor[1], S Matthews[1], E Burton[1], K Hocking[1], S Burns[1], H Reilly[1], E Anderson[1], S Alabed[1,2,3], K Dwivedi[1,2,3], MA Kamil[1], O Evans[1], W Tindale[1], AJ Swift[1,2,3], CS Johns[1]

[1] Sheffield Teaching Hospitals NHS Foundation Trust, [2] NIHR Sheffield Biomedical Research Centre, [3] The University of Sheffield

Background

- Automated detection of lung cancer on chest radiographs could improve the efficiency of "straight-to-CT" and other diagnostic pathways
- Many commercially available AI-based software options are available, but 1) there is limited evidence to support their clinical deployment and 2) their performance has not been directly compared [1]

Objective

- To evaluate and compare the diagnostic accuracy of commercial AI-based software from seven vendors (*Annalise.ai*, *Gleamer*, *Infervision*, *Milvue*, *Oxipit*, *Qure.ai* and *Rayscape*) [a] for the detection of lung cancer on chest radiographs

Methods

Inclusion criteria	The first chest radiograph acquired for an adult patient at Sheffield Teaching Hospitals between July 2020 and February 2021 which was requested from primary care for any indication
Data collection	Retrieval of DICOM image, patient demographic information and radiologist report from PACS and RIS
Reference standard labelling	Labelling of radiologist reports for detection of nodule/mass and suspicion for cancer Assessment of lung MDT meeting records between July 2020 and August 2021 for diagnosis of cancer
Image analysis by software	Each radiograph analysed by commercially available software from the seven vendors Coordinated by *Blackford Analysis*
Evaluation of diagnostic accuracy	Binary software outputs compared against radiologist reports and cancer diagnosis by MDT decision Calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy Failure analysis of incorrectly classified cases

Results [b,c]

Patient population

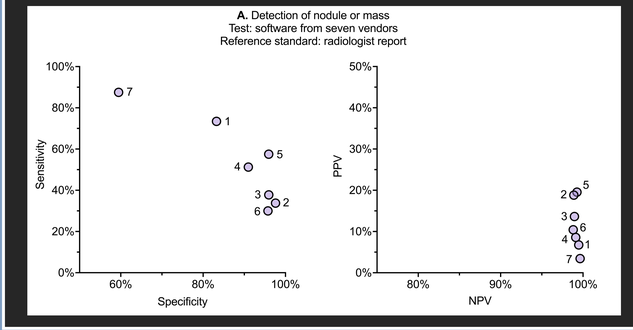
- 5235 radiographs from as many patients (median age 60 years, 53.4% female, 79.4% white)
- 85 cases of lung cancer (1.6% disease prevalence, median age 71 years, 58.8% female, 88.2% white)

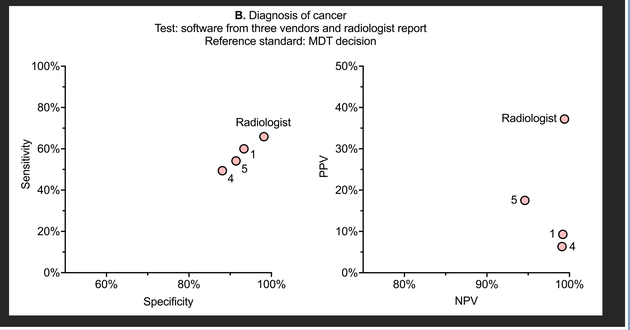
Software performance for nodule detection [A]

- Sensitivity and specificity varied across vendors (with some showing higher sensitivity and lower specificity, and vice versa)
- NPV was high (> 98%) and PPV was low (< 20%) across all vendors

Software performance for cancer diagnosis [B]

- Software from three vendors provided an overall decision for lung cancer diagnosis
- Sensitivity (49-55%) and specificity (88-97%) were similar across vendors but remained inferior to radiologists (66% and 98%)
- NPV was similar (>99%) to radiologists, but PPV was considerably lower (<18% compared with 37%)





Conclusions

- Diagnostic accuracy for varied across vendors, particularly for nodule or mass detection
- Software showed high NPV comparable to radiologists and could aid diagnostic pathway efficiency by excluding lung cancer
- However, software showed much lower PPV than radiologists, indicating a higher risk of false positives and over-investigation
- Caution must be exercised when selecting software from multiple options to ensure an appropriate balance of benefits and risks

This study was funded by the South Yorkshire Integrated Care System. Local ethics committee approval was obtained. [a] Listed in alphabetical order. [b] These represent results from preliminary analysis and are subject to change. [c] The vendors have been assigned numerical identifiers at random.

[1] NICE. Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment. 2023. [2] Maiter et al. Evaluating the performance of artificial intelligence software for lung nodule detection on chest radiographs in a retrospective real-world UK population. BMJ Open. 2023.