EVALUATING THE CLINICAL IMPACT OF AI-DRIVEN AUTONOMOUS CHEST RADIOGRAPH REPORTING

Dr .Farah Afzal, Sheikh Khalifa bin Zayed AJK Hospital Muzaffarabad[1], Dr. Mubashrah Aziz AFIRI Rawalpindi[2], Dr. Sadia Tahir PAF Hospital Shorkot[3] [1,2,3]Pakistan
ilmoakl@hotmail.com mubashratahir@gmail.com diatahir@hotmail.com

INTRODUCTION

Artificial intelligence is a rapidly emerging field with immense transformative potential. Automated interpretation of normal chest radiographs can reduce the workload of radiologists, but the performance of such an artificial intelligence (AI) tool compared to clinical radiology reports has not been established. Keeping in view the day by day increasing potential and improvements in artificial intelligence it is the need of the hour to research its role in imaging.

PURPOSE

The purpose of the study is to perform an external evaluation of a commercially available AI tool for

- ✓ number of chest radiographs autonomously reported
- ✓ sensitivity of AI detection of abnormal chest radiographs
- ✓ performance of AI compared with that of the clinical radiology reports.

MATERIAL AND METHODS

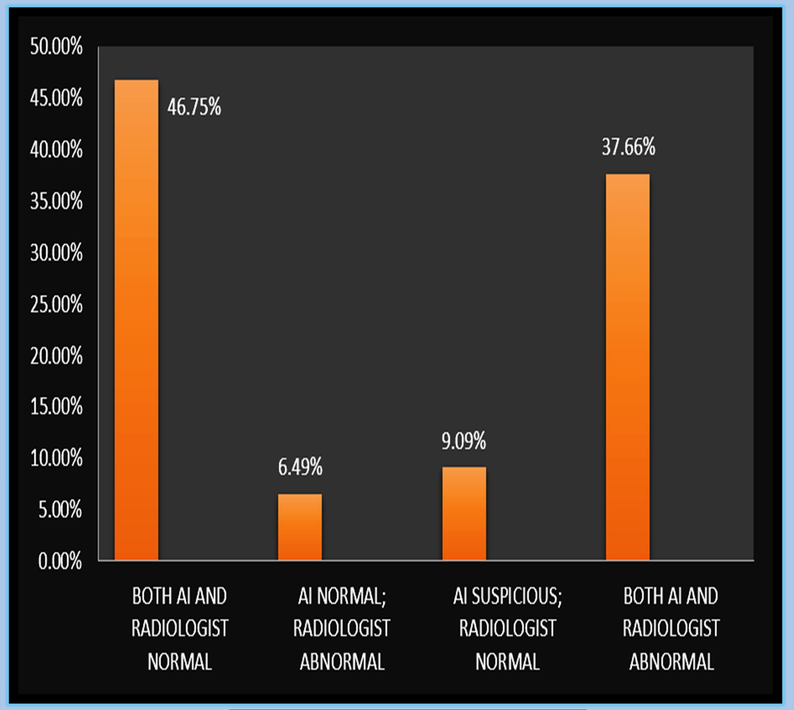
- ✓ In this retrospective study, consecutive chest radiographs from adult patients in SKBZ Hospital, Muzaffarabad Pakistan were evaluated by thoracic radiologists who labeled chest radiographs into the three categories i.e. normal, abnormal, and suspicious; according to a reference standard based on chest radiograph findings
- ✓ These radiographs included images from emergency department patients, in-hospital patients, and outpatients.
- ✓ Mc Nemar test was applied to compare the results of AI and radiologist findings

RESULTS

A total of 77 patients, median age 33 years were included in the analysis. Among these patients, 32 were female and 45 were male.

- ❑ Both AI and radiologists agreed on normal results in 36 cases (46.75%)
- ❑ Both AI and radiologists agreed on abnormal results in 29 cases (37.66%)
- ❑ In 5cases (6.49%), AI reported normal & radiologists reported abnormal
- ❑ In 7 cases (9.09%), AI diagnosed suspicious while radiologists diagnosed normal

The statistical values for McNemar's test were 0.083 (p-value of 0.711)



CONCLUSION

AI and radiologists reached an agreement on normal results in 46.75% of cases and on abnormal results in 37.66% of cases. The McNemar's test indicated no significant difference between AI and radiologist diagnoses (p = 0.711)