Comparison of chest radiograph interpretations by artificial intelligence vs radiologists in the diagnosis of pulmonary TB

Embargo: 0001H CEST Copenhagen local time Tuesday 18 April

2a. Tuberculosis and other mycobacterial infections (incl. epidemiology, clinical, diagnostics, antimycobacterial drugs, susceptibility testing, treatment & prevention)

Likely attendance
Onsite
Oral presentation
Z. Desita 1, T. Tadesse 1, A. Sifna 2, H. Fekadu 3, S. Bizuneh 3, S. Sridhar 4, C. Wejse 5, 2, 6, T. Schön 7, F. Rudolf 5, 2
1Department of Radiology, College of Medicine and Health Science, University of Gondar – Gondar (Ethiopia), 2Bandim Health Project - Bissau (Guinea-bissau), 3Department of Interna medicine, College of Medicine and Health Science, University of Gondar - Gondar (Ethiopia), 4Qure.ai – Mumbai (India), 5Aarhus University Hospital, Department of infectious diseases - Aarhus (Denmark), 6GloHAU, Center for Global Health, School of Public Health, Aarhus University - Aarhus (Denmark),
7Department of Microbiology and Infectious Medicine, Kalmar County Hospital - Kalmar (Sweden)

Background

Chest Xray (CXR) remains a diagnostic tool in smear negative preTB patients in low resouces, rural areas - but there is a lack of high-quality studies to assess its diagnostic accuracy as highlighted recently by WHO. Computer-aided detection (CAD) programs based on artificial intelligence (AI) algorithms have improved detection of microbiologically confirmed TB. CAD CXR may contribute to diagnosis of TB cases where microbiological methods fail.The aim of this study was to compare the performance of automated software (Qure) in chest X-ray (CXR) assessment for TB with that of two Ethiopian radiologists with different experience.

Methods

A retrospective study comparing performance of human and automated reading was conducted on traditional CXRs taken as part of a randomized trial. Two reference standards were applied; final TB diagnosis given on clinical or laboratory grounds and GenXpert confirmed TB. Performance was assessed by receiver operating characteristic analysis. Agreement between the readers was assessed by kappa scores. To improve applicability in low resource settings, the CXRs presented to the AI where pictures of CXRs taken by mobile phone.

Results

We included 498 CXRs in the final analysis. Of those, the less experienced radiologist found 50, the more experienced radiologist found 100 and the software found 83 to be indicative of TB. The overall AUC for the software was 0.84 for GenXpert confirmed cases. The less experienced radiologist assessments had a sensitivity of 62.5% and a specificity of 91.7% while the experienced radiologist's assessments were 75.0% sensitive and 82.0% specific for GenXpert confirmed cases. The agreement between the radiologists was moderate (k=0.45), as was the agreement between each radiologist and the software (k=0.36 and k=0.59) and the pooled assessment of the radiologists and the software (k=0.55).

Conclusions

CAD assisted TB detection could improve the diagnostic yield of TB in settings where well trained radiologists are scarce. A simple picture taken by a mobile phone is sufficient and thus makes it feasible in high-incidence low-resource settings.

Female, n (%)	250 (52)
Mean age (IQR)	38 (24 – 50)
Mean Bioscan (IQR)	4.4 (3 – 9)
HIV Infected, n (%)	28 (6)
Uninfected, n (%)	470 (94)
TB Yes, n (%)	57 (11)
GeniXpert, n (% of TB)	16 (28)
Clinical, n (% of TB)	41 (72)

Finding, n(%)	Radiologist 1	Radiologist 2	AI
Consolidation	76 (4)	53 (6)	69 (14)
Cavity	5 (1)	21 (4)	7 (1)
Fibrosis	45 (9)	66 (13)	39 (8)
Nodule	28 (5)	21 (4)	40 (10)
Blunted Costophrenic angle	19 (4)	11 (2)	10 (2)
Pleural effusion	14 (3)	6 (2)	11 (2)
Pneumothorax	0	0	5 (1)
Hilar lymphadenopathy	15 (3)	24 (5)	7 (1)
Tracheal shift	15 (3)	23 (4)	29 (6)
Cardiomegaly	5 (1)	64 (13)	15 (3)
Scoliosis	1 (0.1)	2 (0.4)	7 (1)
Abnormal classified*	89 (18)	73 (15)	101 (32)
Abnormal radiualified**	107 (21)	166 (55)	152 (27)
Abnormal TB***	84 (17)	59 (18)	110 (22)
rgc***	30 (10)	89 (30)	81 (16)

*as judged by interpreter

**classified: at analysis if one of the above found

***classified: at analysis if one of following found: consolidation, cavity, nodule, blunting, effusion, hilarymhadenopathy, tracheal shift

****as judged by interpreter
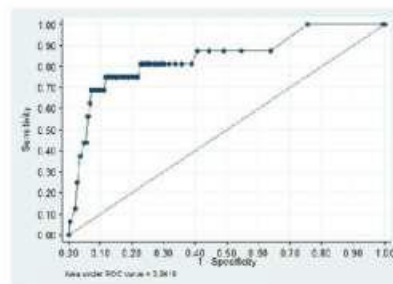
Figure 2 Qure only GX confirmed (*: Observer1, ▲: Observer2)

diagnosis

Conflicts of interest **Do you have any conflicts of interest to declare?** No