EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Research letter

Human reading *versus* computer automated reading of chest X-rays in a tuberculosis screening programme in Romania

Gerard de Vries, Dan Gainaru, Sytze Keizer, Beatrice Mahler, Ileana Radulescu, Marina Zamfirescu, Ibrahim Abubakar

Please cite this article as: de Vries G, Gainaru D, Keizer S, *et al*. Human reading *versus* computer automated reading of chest X-rays in a tuberculosis screening programme in Romania. *Eur Respir J* 2021; in press (https://doi.org/10.1183/13993003.04628-2020).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

©The authors 2021. For reproduction rights and permissions contact permissions@ersnet.org

Human reading versus computer automated reading of chest X-rays in a tuberculosis screening programme in Romania

Gerard de Vries[1 2], Dan Gainaru[3], Sytze Keizer[4], Beatrice Mahler[3], Ileana Radulescu[3], Marina Zamfirescu[5], Ibrahim Abubakar[6]

Affiliations of authors

1. KNCV Tuberculosis Foundation, The Hague, The Netherlands
2. National Institute for Public Health and the Environment, Bilthoven, The Netherlands (gerard.de.vries@rivm.nl)
3. Marius Nasta Institute of Pneumology, Bucharest, Romania (dgainaru@yahoo.com, mahlerbeatrice@yahoo.com, ileradulescu@yahoo.co.uk)
4. Independent consultant (sytze.keizer@gmail.com)
5. Audifon Medical Centre, Bucharest, Romania (jmzamfirescu@yahoo.fr)
6. Institute for Global Health, University College London, London, United Kingdom (i.abubakar@ucl.ac.uk)

Word count: 1216

One of the interventions in tuberculosis (TB) control is to screen people at high risk for TB with a chest x-ray (CXR).[1] CXRs in TB screening programmes are usually read by a radiographer or a pulmonologist specialized in TB. In recent years, computer-aided detection (CAD) software has become available for automated reading of CXRs and identifying people with presumptive TB [2, 3] and for TB screening.[4, 5] A systematic review published in 2016 concluded that the evidence of diagnostic accuracy of CAD was limited by the small number of studies, co-authored by owners of the only CAD software on the market at that time, and not generalisable to low TB and HIV settings.[6] The application of CAD software for TB detection has to our knowledge not been assessed in Europe.

From August 2018 till September 2019, TB screening was carried out in Romania in prisons, homeless persons, drug users and Roma population with a mobile digital X-ray unit equipped with CAD software (CAD4TB version 6, Delft Imaging, 's-Hertogenbosch, The Netherlands). The CAD4TB software analyses the unobscured lung fields of a posterior-anterior CXR for the presence of abnormalities and provides a score between 0 and 100.[7] The screening activity was part of the E-DETECT TB project that applies evidence-based interventions to ensure early diagnosis and treatment of TB in vulnerable populations in European Union countries.[8] A total of 5003 individuals were radiologically screened (all in posterior-anterior projection) in the Romanian project; 5000 had a valid CAD4TB score. All CXRs were read by a pulmonologist (DG), who had more than 20 years' experience in reading CXRs in the TB screening programme. The management of people with presumptive TB or other CXR abnormalities was based on Romanian TB guidelines.[9] Ten individuals were diagnosed with TB, all with positive *Mycobacterium tuberculosis* cultures or positive Xpert MTB/RIF (Cepheid, Sunnyvale, CA, USA), results. Detected TB prevalence was 200 per 100 000 persons screened (95% confidence interval (CI) 96-368 per 100 000). Nine had a CAD4TB score >60 and one had a CAD4TB score of 59.

We designed a study to compare the CAD4TB reading with human reading as a tool to rule in people with radiographical abnormalities who require further evaluation for active TB. We included all 258 CXRs with a CAD4TB score >60, and randomly selected 742 CXRs (out of 4742; 15.6%) from the remaining CXRs with a score ≤60. The selected CXRs, , except one that could not be accessed (total: 999), were read by two TB specialists (SK, IR) CXRs with discordant results between the two readers, were read by a third TB specialist (MZ). All three readers (two pulmonologists and one TB public health physician) had more than 20 years' experience in reading CXRs in TB screening programmes. Readers were blinded for age of the screened persons, and the CAD4TB score of the CXRs. The three readers were asked to classify the CXR findings according to a classification previously used [10] and modified for this study, i.e. i) normal; ii) highly suggestive for TB (cavities or extensive infiltrative disease, that warrant immediate action, i.e. airborne infection isolation of the patient and urgent sputum examination); iii) possibly suggestive for TB (including TB sequelae and other signs of previous TB). Action is to confirm or rule out active TB; immediate respiratory isolation not required, iv) CXR abnormality not suggestive for TB, including radiological findings suggestive for other pulmonary diseases (silicosis, pneumothorax), heart diseases (cardiomegaly), fractures, scoliosis, but also non-disease related CXR findings such as azygos lobe, a bullet, or a pacemaker. Individuals with a CXR highly suggestive for TB, e.g. cavities or extensive infiltrative disease, would need direct actions such as immediate sputum examination and/or respiratory isolation. Individuals with CXR abnormalities 'possibly suggestive for TB' would need additional investigation to confirm or rule out TB, e.g. by asking for a previous episode of TB treatment, comparison of the CXR with previous ones, sputum examination or otherwise. We combined the results of the CXRs highly and possibly suggestive for TB into one category, i.e. 'presumptive TB'.

Approval was provided by the ethical committee of Marius Nasta Institute of Pneumonology (Nr. 20276).

The results of the two readers were concordant for 798 (80%) of the CXRs (560 'normal', 159 'presumptive TB' and 79 'other CXR abnormalities'), The Kappa score was 0.62 for the three categories (moderate agreement). The 201 (20%) CXRs with discordant results were read by the third reader, who was blinded for the results of the first two readers. The identical result reported by two of the three readers was taken as an agreed composite results, i.e. 46 'normal', 32 'presumptive TB' and 102 'other CXR abnormalities'. In 21 CXRs, one of the readers classified the CXR as 'normal', one as 'presumptive TB' and the third as 'other CXR abnormalities'. These results were classified as 'other CXR abnormalities', since two out of three noticed an abnormality, and only one suggested that the CXR was suggestive for TB.

The CXRs of the ten TB patients were all classified by both readers as 'presumptive TB' and therefore not read by the third reader. Eight of these CXRs were classified by both readers as highly suspect for TB, one CXR was classified by one reader as highly suspect and by the other as possibly suggestive for TB, and one CXR was classified as possibly suggestive for TB by both readers. The figure shows a linear association between the CAD4TB scores and the human readers classifying the CXR as presumptive TB. All CXRs with a CAD4TB score >70 were scored abnormal, most of them as presumptive TB.

We tested the algorithm that only CXRs above a threshold would need human reading, and based on the decision of the clinician additional TB examinations. Based on the sample size, we estimated (by extrapolation) the number of people that would be ruled in by CAD4TB and human reading. A CAD4TB score >60 would rule in 5.2% (n=258) of the CXRs for human reading, who would identify 164 individuals with presumptive TB, including nine of ten patients diagnosed in the project (sensitivity 90%; 95% CI 60-98%).

This algorithm would have avoided human reading of 94.8% (n=4 742) of the CXRs, and the identification and subsequent examination of 178 individuals with presumptive TB, including one TB patient (CAD4TB score 59). If a cut-off of 50 was used, 16.3% (n=815) would be ruled-in for human reading, 275 individuals would be identified with presumptive TB, including all ten individuals who were diagnosed with TB in the project (sensitivity 100%). It would have avoided human reading of 83.7% (n=4 185) of the CXRs and (unnecessary) examination of 67 individuals with presumptive TB (none of them had TB).

We purposely asked readers in this study to rule in CXRs with a low suspicion for TB, in order to increase sensitivity. This has resulted in a lower specificity, possibly a lower Kappa score and overestimation of people with presumptive TB (total 6.8%). This high proportion was partly due to the high percentage (2.7%) of clients with self-reported previous TB.

We conclude that computer-aided detection software provides an opportunity to identify people with presumptive TB in screening, allowing a reduction in the proportion of CXRs that need human reading (to less than 20% in our project) and avoiding unnecessary TB examinations while maintaining high sensitivity.

Acknowledgement

The authors gratefully acknowledge the E-DETECT TB (709624) project which has received funding from the European Union's Health Programme (2014–2020). The views expressed here are the authors only and are their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the EU. IA acknowledges support from the UK National Institute for Health Research (SRF-2011-04-001) and (NF-SI-0616-10037).

Delft Imaging was a co-applicant on the EU grant that supported the project which required all partners to co-fund their contribution. The company was not involved in this study.

Support statement

The E-DETECT TB project has received funding from the European Commission Consumers, Health, Agriculture and Food Executive Agency (grant number: 709624). Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest

I. Abubakar reports grants from European Commission to undertake the project reported in this manuscript.

References

1. Systematic screening for active tuberculosis: principles and recommendations. WHO/HTM/TB/2013.04. Geneva: World Health Organization; 2013.

2. Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Zhiyun Xue null, Palaniappan K, Singh RK, Antani S, Thoma G, Yi-Xiang Wang null, Pu-Xuan Lu null, McDonald CJ. Automatic tuberculosis screening using chest radiographs. *IEEE Trans Med Imaging* 2014; 33: 233–245.

3. Melendez J, van Ginneken B, Maduskar P, Philipsen RHHM, Reither K, Breuninger M, Adetifa IMO, Maane R, Ayles H, Sánchez CI. A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-rays. *IEEE Trans Med Imaging* 2015; 34: 179–192.

4. Madhani F, Maniar RA, Burfat A, Ahmed M, Farooq S, Sabir A, Domki AK, Page-Shipp L, Khowaja S, Safdar N, Khan AJ, Khan PY. Automated chest radiography and mass systematic screening for tuberculosis. *Int J Tuberc Lung Dis* 2020; 24: 665–673.

5. Young C, Barker S, Ehrlich R, Kistnasamy B, Yassi A. Computer-aided detection for tuberculosis and silicosis in chest radiographs of gold miners of South Africa. *Int. J. Tuberc. Lung Dis.* 2020; 24: 444–451.

6. Pande T, Cohen C, Pai M, Ahmad Khan F. Computer-aided detection of pulmonary tuberculosis on digital chest radiographs: a systematic review. *The International Journal of Tuberculosis and Lung Disease* 2016; 20: 1226–1230.

7. Melendez J, Philipsen RHHM, Chanda-Kapata P, Sunkutu V, Kapata N, van Ginneken B. Automatic versus human reading of chest X-rays in the Zambia National Tuberculosis Prevalence Survey. *The International Journal of Tuberculosis and Lung Disease* 2017; 21: 880–886.

8. Abubakar I, Matteelli A, de Vries G, Zenner D, Cirillo DM, Lönnroth K, Popescu G, Barcellini L, Story A, Migliori GB. Towards tackling tuberculosis in vulnerable groups in the European Union: the E-DETECT TB consortium. *Eur. Respir. J.* 2018; 51.

9. Ghid metodologic de implementare a programului național de prevenire, supraveghere și control al tuberculozei. Marius Nasta Institute of Pneumology, Bucharest, Romania; 2015.

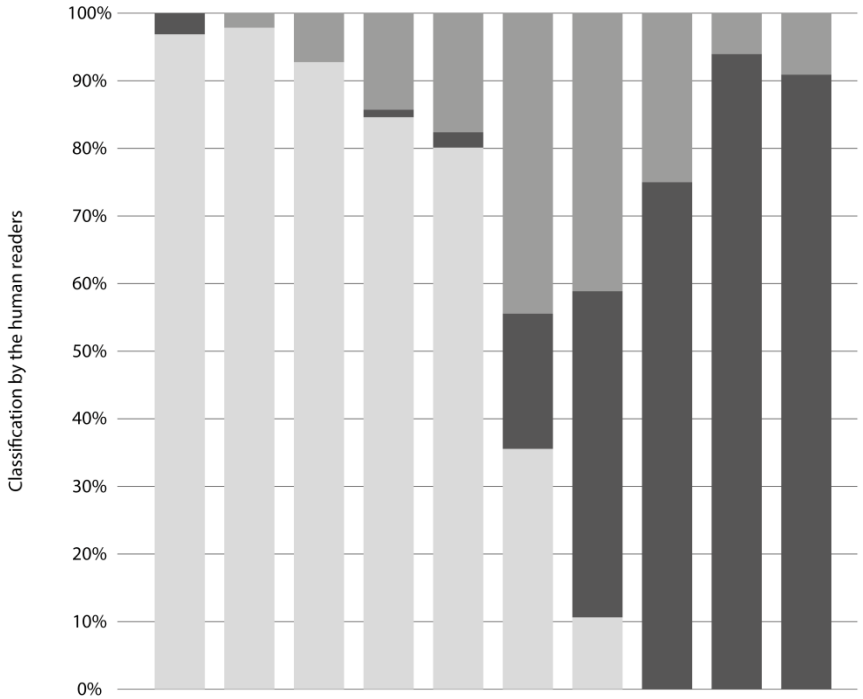
10. Breuninger M, Ginneken B van, Philipsen RHHM, Mhimbira F, Hella JJ, Lwilla F, Hombergh J van den, Ross A, Jugheli L, Wagner D, Reither K. Diagnostic Accuracy of Computer-Aided Detection of Pulmonary Tuberculosis in Chest Radiographs: A Validation Study from Sub-Saharan Africa. *PLOS ONE* 2014; 9: e106381.

Figure. Comparison of human reading results with computer-aided detection of chest X-rays in a sample of the chest X-rays of a screening project for active tuberculosis in Romania with extrapolation to the actual screened population.

CAD4TB = computer-aided detection for tuberculosis (Delft Imaging, ‘s-Hertogenbosch, the Netherlands)

* All chest X-rays (CXRs) with CAD4TB score >60, except one with a CAD4TB score of 64 that could not be accessed, were read, as well as 15.6% of CXRs with a score ≤60 (742 CXRs).

** Extrapolation was done for each CAD4TB category (of ten values) by multiplying the actual number of screened people in the E-DETECT TB project with CAD4TB results in these categories with the proportions of classification by the human readers in the respective categories in the study population..



Study population (n=999)*

	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
Normal	31	92	77	77	282	32	15	0	0	0
Presumptive tuberculosis	1	0	0	1	8	18	68	54	31	10
Other chest X-ray abnormalities	0	2	6	13	62	40	58	18	2	1

Extrapolation to the actual screened population (n=5000) in the E-DETECT TB project**

Total	**218**	**502**	**537**	**596**	**2332**	**557**	**142**	**72**	**33**	**11**
Normal	211	491	498	504	1868	198	15	0	0	0
Presumptive tuberculosis	7	0	0	7	53	111	69	54	31	10
Other chest X-ray abnormalities	0	11	39	85	411	248	58	18	2	1
			(1-50)			(51-60)		(61-100)		
Subtotals CAD4TB score categories			4185			557		258		