

ỨNG DỤNG BƯỚC ĐẦU TRÍ TUỆ NHÂN TẠO QURE.AI X-QUANG LỒNG NGỰC TRONG CHẨN ĐOÁN LAO PHỔI

Lê Hoàn¹, Nguyễn Thị Thu Thủy² và Lê Minh Hằng^{1,✉}

¹Trường Đại học Y Hà Nội

²Trường Đại học Y Dược Thái Bình

Bệnh lao phổi vẫn đang là gánh nặng bệnh tật trên toàn thế giới, đặc biệt ở các nước đang phát triển. Trí tuệ nhân tạo trong chẩn đoán lao phổi nhằm mục đích phát hiện sớm tổn thương, chẩn đoán chính xác bệnh. Nghiên cứu của chúng tôi nhằm đánh giá kết quả ứng dụng bước đầu trí tuệ nhân tạo Qure.AI X-quang lồng ngực trong chẩn đoán lao phổi. Nghiên cứu mô tả với 126 bệnh nhân nghi lao phổi được Qure.AI đọc X-quang lồng ngực và làm các xét nghiệm vi sinh hoặc mô bệnh học để chẩn đoán lao phổi và so sánh giữa kết quả chẩn đoán cuối cùng với phân tích AI cho kết quả tính đồng thuận ở mức tốt giữa 2 phương pháp. Độ nhạy của AI là 88,8%, độ đặc hiệu là 48,8%, giá trị dự đoán âm tính 71%. Diện tích dưới đường cong ROC là 77,1% với $p < 0,001$, KTC 95%: 0,69 - 0,86. Như vậy, hình ảnh X-quang lồng ngực được đọc bởi Qure.AI có giá trị trong chẩn đoán lao phổi với độ chính xác ở mức độ khá tốt với điểm cut off là 0,503. Kết quả cho thấy, AI có tiềm năng trở thành một công cụ cần thiết trong việc hỗ trợ chẩn đoán sớm lao phổi, sàng lọc lao phổi cộng đồng.

Từ khóa: Lao phổi, trí tuệ nhân tạo Qure.AI.

I. ĐẶT VẤN ĐỀ

Theo báo cáo của Tổ chức Y tế Thế giới năm 2021, trên thế giới có khoảng 1/3 dân số (2,2 tỉ người) đã nhiễm lao và con số đó sẽ tăng 1% mỗi năm (tương đương khoảng 65 triệu người).^{1,2} Bệnh lao ở Việt Nam có thể xếp vào loại trung bình cao ở khu vực Tây Thái Bình Dương, là khu vực có độ lưu hành lao trung bình trên thế giới. Tiêu chuẩn vàng để chẩn đoán bệnh lao là tìm thấy trực khuẩn lao trong bệnh phẩm lấy từ người bệnh. Tuy nhiên, trên thực tế, việc tìm ra trực khuẩn lao từ các phương pháp nuôi cấy cổ điển rất hạn chế, thời gian nuôi cấy dài làm chậm tiến hành điều trị do đó bệnh tiến triển nặng hơn và làm tăng nguy cơ lây bệnh cho cộng đồng. Như vậy, việc phát

hiện sớm, chính xác bệnh lao là một yêu cầu cấp thiết.

Cho đến nay, nhiều phương pháp mới đã được áp dụng nhằm chẩn đoán sớm bệnh lao phổi trong đó phải kể đến kĩ thuật ứng dụng trí tuệ nhân tạo (AI - artificial intelligence) để đọc phim X-quang lồng ngực. Trên thế giới, đã có nhiều bài báo quốc tế và công trình nghiên cứu nghiên cứu về vấn đề ứng dụng công nghệ AI đọc phim X-quang lồng ngực trong chẩn đoán lao phổi. Ở Việt Nam, ứng dụng công nghệ Qure.AI (một phần mềm trí tuệ nhân tạo của Ấn Độ) đọc phim X-quang lồng ngực mới được áp dụng khoảng 2 đến 3 năm trở lại đây tại một số bệnh viện tuyến trung ương. Hiện có rất ít nghiên cứu tại Việt Nam về vấn đề này. Vì vậy, chúng tôi làm nghiên cứu này với mục tiêu đánh giá kết quả ứng dụng bước đầu trí tuệ nhân tạo Qure.AI X-quang lồng ngực trong chẩn đoán lao phổi.

Tác giả liên hệ: Lê Minh Hằng

Trường Đại học Y Hà Nội

Email: hangle.hmu@gmail.com

Ngày nhận: 26/09/2023

Ngày được chấp nhận: 02/11/2023

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

1. Đối tượng

Bao gồm 126 bệnh nhân nghi lao phổi trên 18 tuổi đến khám hoặc điều trị nội trú tại Bệnh viện Đại học Y Hà Nội trong thời gian từ tháng 1 đến tháng 12/2022 và được sử dụng phần mềm trí tuệ nhân tạo Qure.AI để phân tích hình ảnh X-quang lồng ngực.

Tiêu chuẩn lựa chọn

Bệnh nhân nghi lao phổi theo Quyết định 3126/QĐ-BYT hướng dẫn chẩn đoán, điều trị, dự phòng lao ngày 23 tháng 05 năm 2018.⁴

- Lâm sàng:

+ Ho kéo dài (ho khan, ho có đờm, ho ra máu).

+ Gầy sút cân, kém ăn, mệt mỏi.

+ Sốt nhẹ về chiều.

+ Ra mồ hôi trộm ban đêm.

+ Đau ngực, đôi khi có khó thở.

- Cận lâm sàng:

+ X-quang ngực: hình ảnh gợi ý lao tiến triển như tổn thương nốt, thâm nhiễm chủ yếu xuất hiện ở đỉnh phổi, hình ảnh hang lao...

+ Bệnh nhân được làm các xét nghiệm vi khuẩn học (AFB, MGIT, GenXpert, PCR TB đờm/DPQ/DMP) hoặc xét nghiệm mô bệnh học (sinh thiết qua NSPQ, sinh thiết xuyên thành ngực, sinh thiết màng phổi).

Tiêu chuẩn loại trừ

- Bệnh nhân không được chụp phim X-quang lồng ngực.

- Bệnh nhân không được làm các xét nghiệm có giá trị chẩn đoán xác định lao phổi: xét nghiệm vi sinh hoặc xét nghiệm mô bệnh học.

2. Phương pháp

Thời gian và địa điểm nghiên cứu

- Thời gian: từ tháng 08/2022 đến 08/2023.

- Địa điểm: Bệnh viện Đại học Y Hà Nội.

Thiết kế nghiên cứu

Mô tả cắt ngang.

Cỡ mẫu và phương pháp chọn mẫu

- Cỡ mẫu n tính theo công thức:

$$n \geq Z_{1-\alpha/2}^2 \cdot \frac{p(1-p)}{d^2}$$

Thay $1 - \alpha = 95$, $p = 0,2$, $d = 0,07$ tính được $n \geq 126$.

- Phương pháp chọn mẫu: thuận tiện từ danh sách những bệnh nhân nghi lao phổi trên 18 tuổi đến khám và điều trị nội trú tại Bệnh viện Đại học Y Hà Nội trong thời gian từ tháng 1 đến tháng 12/2022.

Công cụ Qure.AI

- Qure.AI là giải pháp kiểm tra X-quang ngực được xây dựng bằng cách sử dụng học sâu xuất xứ từ Ấn Độ.

- Tiêu chí về hình ảnh X-quang để chẩn đoán lao phổi của Qure.AI: nốt dạng kê (đường kính 1 - 5mm, phân bố khá đều, thường tập trung vùng đỉnh phổi), tổn thương hang (hình sáng được giới hạn rõ rệt bởi một bờ cản quang liên tục, khép kín. Kích thước từ 2 - 4cm hoặc có thể rất to. Có thể một hoặc nhiều hang. Bên trong hang lao có thể có mức nước mức hơi), tổn thương thâm nhiễm (đám mờ nhạt ở dưới xương đòn, kích thước khác nhau, giới hạn không rõ, đôi khi tập trung thành đám tròn hoặc trên cơ sở một dải mờ nhạt có một số nốt nhỏ), tổn thương dạng nốt (bóng mờ có kích thước nhỏ, tròn hoặc bầu dục đường kính 3 - 10mm, thương gặp ở hạ đòn và đỉnh phổi 2 bên).

Quy trình nghiên cứu

- Từ danh sách các bệnh nhân nghi lao trên 18 tuổi đến khám và điều trị nội trú tại Bệnh viện Đại học Y Hà Nội lựa chọn các bệnh nhân thỏa mãn các tiêu chuẩn của nghiên cứu (tra cứu thông tin bệnh nhân trên hồ sơ bệnh án điện tử, các bệnh nhân có triệu chứng lâm sàng nghi lao, tổn thương phổi trên X-quang lồng ngực và được làm các xét nghiệm vi khuẩn học (AFB, MGIT, GenXpert, PCR TB đờm/DPQ/DMP) hoặc xét nghiệm mô bệnh học (sinh thiết qua

NSPQ/ sinh thiết xuyên thành ngực/ sinh thiết màng phổi). Đối với những bệnh nhân có xét nghiệm vi khuẩn và mô bệnh học âm tính với lao phải được khám lại tại Bệnh viện Đại học Y Hà Nội theo dõi tổn thương phổi và không được điều trị lao không bằng chứng mới được chọn vào nghiên cứu).

- Đọc phim X-quang lồng ngực bằng công nghệ Qure.AI.

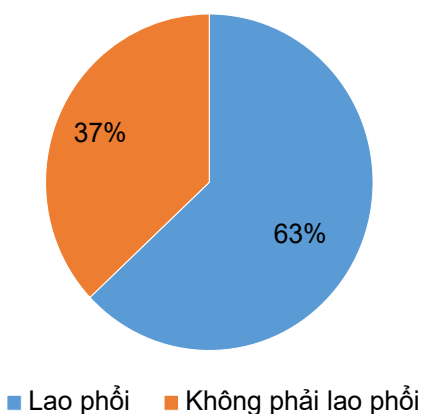
III. KẾT QUẢ

Phương pháp xử lý số liệu

Số liệu được nhập, xử lý, phân tích bằng phần mềm SPSS 20.0, sử dụng các thuật toán thống kê y học: tính giá trị trung bình, tính tỷ lệ %, tính độ nhạy, độ đặc hiệu, giá trị dự báo âm tính, giá trị dự báo dương tính...

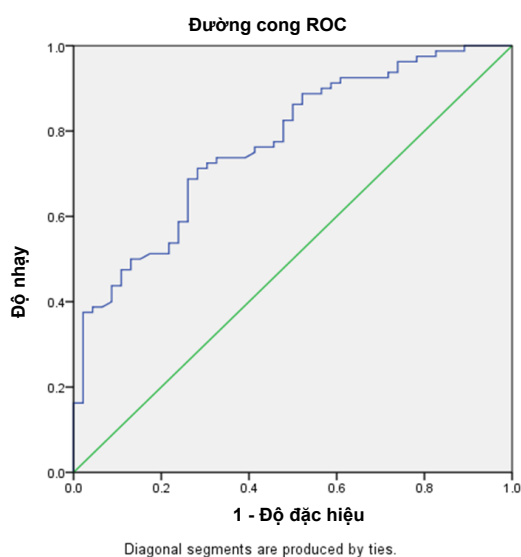
3. Đạo đức nghiên cứu

Nghiên cứu tuân thủ đầy đủ các nguyên tắc của nghiên cứu y học. Toàn bộ thông tin của người bệnh đều được bảo mật.



Biểu đồ 1. Kết quả chẩn đoán cuối cùng của các bệnh nhân nghiên cứu

Tổng số 126 bệnh nhân nghiên cứu, có 80 bệnh nhân chẩn đoán lao phổi chiếm tỷ lệ 63,5%; 46 bệnh nhân chẩn đoán không phải lao phổi chiếm 36,5%.



Biểu đồ 2. Đường cong ROC Kết quả X-quang lồng ngực phân tích bởi công nghệ Qure.AI

Diện tích dưới đường cong ROC là 77,1% với $p < 0,001$, KTC 95%: 0,69 - 0,86. Như vậy hình ảnh X-quang lồng ngực được đọc bởi

Qure.AI có giá trị trong chẩn đoán lao phổi với điểm cut off là 0,503.

Bảng 1. Kết quả X-quang lồng ngực phân tích bởi công nghệ Qure.AI

Kết quả đọc AI X-quang lồng ngực	Số lượng (n)	Tỷ lệ %
Hướng đến tổn thương lao ($\geq 0,503$)	95	75,4
Không hướng đến tổn thương lao ($< 0,503$)	31	24,6
Tổng	126	100,0

Trong tổng số 126 phim X-quang lồng ngực được đọc bởi công nghệ Qure.AI, có 95 phim X-quang được đọc hướng đến tổn thương lao

chiếm tỷ lệ 75,4%; 31 phim X-quang lồng ngực được đọc không hướng đến tổn thương lao chiếm tỷ lệ 24,6%.

Bảng 2. Mối liên quan giữa kết quả X-quang lồng ngực phân tích bởi công nghệ Qure.AI với kết quả chẩn đoán cuối cùng của các bệnh nhân nghiên cứu

Kết quả AI X-quang lồng ngực	Chẩn đoán xác định	
	Lao phổi	Không phải lao phổi
Hướng đến tổn thương lao	71	24
Không hướng đến tổn thương lao	9	22

Độ nhạy (Se): 88,8%. Độ đặc hiệu (Sp): 48,8%. Giá trị dự báo dương tính (PPV): 74,8%. Giá trị dự báo âm tính (NPV): 71,0%.

IV. BÀN LUẬN

Nghiên cứu của chúng tôi ghi nhận tỷ lệ bệnh nhân chẩn đoán xác định lao phổi chiếm 63,5%. Kết quả này gần tương tự với ghi nhận của Lê Thị Ba (2014) nghiên cứu trên 184 bệnh nhân nghi lao, trong đó có 112 bệnh nhân được chẩn đoán xác định lao phổi chiếm 60,8%.⁴ Kết quả của nghiên cứu của chúng tôi ghi nhận tỷ lệ mắc lao cao hơn nghiên cứu của Trịnh Việt Anh (2014) trên 253 bệnh nhân nghi lao phổi trong đó có 56 bệnh nhân chẩn đoán xác định là lao phổi chiếm 22,1%.⁵ Sự khác nhau này có thể do cách chọn bệnh nhân khác nhau dẫn đến kết quả khác nhau.

Nghiên cứu của chúng tôi gồm 126 bệnh

nhân nghi lao phổi, tất cả các bệnh nhân đều được chụp phim X-quang lồng ngực, và phần mềm trí tuệ nhân tạo Qure.AI phân tích kết quả của 126 phim X-quang lồng ngực này. Phần mềm Qure.AI phân tích kết quả X-quang lồng ngực bằng cách cho điểm theo thang điểm từ 0 đến 1, kết quả càng gần 1 thì khả năng tổn thương trên phim X-quang lồng ngực hướng tới do lao càng cao. Kết quả càng gần 0 thì khả năng tổn thương trên phim X-quang lồng ngực càng không hướng tới do lao.

Chạy đường cong ROC tìm được điểm cut-off là 0,503. Điều này có nghĩa là những phim X-quang lồng ngực có điểm $\geq 0,503$ thì được đọc kết quả hướng đến tổn thương lao, còn những phim X-quang lồng ngực có điểm $< 0,503$ được đọc kết quả không hướng đến tổn thương lao. Điểm cut-off này cũng gần giống như điểm cut-off mà phần mềm Qure.AI cài đặt

sẵn là 0,5. Diện tích dưới đường cong ROC là 77,1% với $p < 0,001$, KTC 95%: 0,69 - 0,86. Như vậy, hình ảnh X-quang lồng ngực được đọc bởi Qure.AI có giá trị trong chẩn đoán lao phổi với độ chính xác ở mức độ khá tốt với điểm cut off là 0,503. Kết quả này phù hợp với nghiên cứu của Zhi Zhen Qin và các cộng sự nghiên cứu về độ chính xác trong chẩn đoán lao phổi của 3 hệ thống trí tuệ nhân tạo trong đó có Qure.AI, kết quả diện tích dưới đường cong ROC của phần mềm Qure.AI là 94%, KTC 95%: 0,92 - 0,97.⁶ Nghiên cứu của Zhi Zhen Qin cũng phân tích độ chính xác của 2 hệ thống trí tuệ nhân tạo khác là CAD4TB và Lunit INSIGHT có kết quả diện tích dưới đường cong ROC tương tự như phần mềm Qure.AI lần lượt là 0,92 (KTC 95%: 0,90 - 0,95) và 0,94 (KTC 95%: 0,93 - 0,96).⁵ Một nghiên cứu khác của Madlen Nash và cộng sự thực hiện tại một bệnh viện ở Ấn Độ cũng đánh giá độ chính xác của phần mềm qXR của Qure.AI đọc X-quang lồng ngực cho kết quả AUC là 0,81 (KTC 95%: 0,78 - 0,84).⁶

Trong 126 phim X-quang lồng ngực có 95 phim X-quang được AI đọc hướng đến tổn thương lao chiếm 75,4%, 31 phim X-quang lồng ngực được AI đọc không hướng đến tổn thương lao chiếm 24,6%. So với kết quả chẩn đoán cuối cùng của các bệnh nhân nghiên cứu dựa vào kết quả vi sinh hoặc mô bệnh học (63,5% lao phổi và 36,5% không phải lao phổi) thì không có sự khác biệt quá lớn. Từ đó cho thấy, khả năng phân tích kết quả X-quang lồng ngực của phần mềm Qure.AI là khá tốt. Hứa hẹn sẽ là một công cụ hỗ trợ hiệu quả và đặc lực trong chẩn đoán sớm bệnh lao phổi và trong công tác sàng lọc lao phổi cộng đồng. Nghiên cứu của tác giả Lê Hoàn và các cộng sự về vai trò của ứng dụng trí tuệ nhân tạo VinDr trong dự báo nguy cơ ác tính của nốt phổi đơn độc cũng cho kết quả khả quan về khả năng dự báo nguy cơ ác tính của VinDr. Cụ thể là kết quả

phim chụp cắt lớp vi tính lồng ngực do VinDr phân tích 65,2% ác tính, 34,8% lành tính, kết quả mô bệnh học 78,3% ác tính; 22,7% lành tính. Qua đó cho thấy AI có tiềm năng trở thành một công cụ cần thiết trong việc dự đoán nguy cơ ác tính của nốt phổi đơn độc.⁷

Kết quả nghiên cứu ghi nhận độ nhạy của Qure.AI trong chẩn đoán lao phổi là 88,8%; độ đặc hiệu là 48,8%. Có thể do cỡ mẫu còn hạn chế nên nghiên cứu của chúng tôi chưa thể đưa ra được độ nhạy cũng như độ đặc hiệu ở mức tin cậy cho phương pháp AI mà chỉ đánh giá được mức độ đồng thuận trong dự đoán của AI về tổn thương phổi và tiêu chuẩn vàng để chẩn đoán lao phổi là kết quả vi sinh hoặc mô bệnh học. Tuy nhiên, đây là bước đầu để thấy được vai trò của AI trong việc hỗ trợ chẩn đoán lao phổi, cũng là bước đệm để nhóm nghiên cứu có thể tiến hành nghiên cứu với cỡ mẫu lớn hơn để đạt được các kết quả đáng tin cậy và nhiều ý nghĩa hơn. Tham khảo một số các nghiên cứu đã thực hiện trên thế giới, chúng tôi cũng nhận thấy vai trò đáng tin cậy của AI trong chẩn đoán lao phổi như nghiên cứu của Madlen Nash và cộng sự cho thấy độ nhạy và độ đặc hiệu của phần mềm Qure.AI trong chẩn đoán lao phổi lần lượt là 71% (95% CI: 66%, 76%) và 80% (95% CI: 77%, 83%).⁶ Nghiên cứu của Zhi Zhen Qin và các cộng sự cũng cho thấy phần mềm trí tuệ nhân tạo Qure.AI trong chẩn đoán lao phổi có độ nhạy $\geq 95\%$ và độ đặc hiệu $> 80\%$.⁵ Như vậy, chúng tôi nhận thấy có thể ứng dụng trí tuệ nhân tạo Qure.AI nói riêng và các phần mềm học sâu khác nói chung để giúp đỡ trong việc sàng lọc lao phổi tại cộng đồng, giảm bớt gánh nặng cho các bác sĩ chẩn đoán hình ảnh, đặc biệt tại những nơi ít có khả năng tiếp xúc với các phương pháp chẩn đoán xác định lao phổi.

V. KẾT LUẬN

Phân tích hình ảnh X-quang lồng ngực bởi

Qure.AI có giá trị trong chẩn đoán lao phổi với độ chính xác ở mức độ khá tốt với điểm cut off là 0,503. Độ nhạy của Qure.AI trong chẩn đoán lao phổi là 88,8%; độ đặc hiệu là 48,8%, giá trị dự báo âm tính là 71%.

Kết quả này cho thấy Qure. AI bước đầu thể hiện vai trò trong là phương pháp hỗ trợ đắc lực trong chẩn đoán sớm lao phổi và hỗ trợ cho kết quả vi sinh hoặc mô bệnh học. Tuy nhiên, cần tiếp tục nghiên cứu để cải thiện độ chính xác và ứng dụng thực tế của phương pháp AI trong chẩn đoán lao phổi.

TÀI LIỆU THAM KHẢO

1. WHO. Global Tuberculosis Programme. Global Tuberculosis Reports. Accessed June 24, 2022. <https://www.who.int/teams/global-tuberculosis-programme/tb-reports>
2. Hitzte KL. Preliminary report of the WHO seminar on the evaluation of tuberculosis prevention. *Bull Int Union Tuberc*. 1973;48(0):60-63.
3. Bộ Y tế. Quyết định 3126/QĐ-BYT hướng dẫn chẩn đoán, điều trị, dự phòng lao. Accessed June 23, 2022. <https://vnras.com/quyet-dinh-3126-qd-byt/>
4. Lê Thị Ba. Nghiên cứu đặc điểm lâm sàng, cận lâm sàng ở bệnh nhân nghi lao phổi tại Trung tâm Hô hấp, Bệnh viện Bạch Mai. Trường Đại học Y Hà Nội. 2017. Accessed September 1, 2023.
5. Trịnh Việt Anh. Đặc điểm lâm sàng, cận lâm sàng và xét nghiệm genxpert trong đờm ở những bệnh nhân nghi lao phổi tại trung tâm hô hấp Bệnh viện Bạch Mai. Luận văn Thạc sĩ. Trường Đại học Y Hà Nội. 2015. Accessed September 1, 2023.
6. Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep*. 2019;9(1):15000. doi:10.1038/s41598-019-51503-3
7. Nash M, Kadavigere R, Andrade J, et al. Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. *Sci Rep*. 2020;10:210. doi:10.1038/s41598-019-56589-3
8. Lê Hoàn, Lê Tuấn Linh, Đinh Thị Thanh Hồng, và cs. Vai trò của ứng dụng trí tuệ nhân tạo vindr trong dự báo nguy cơ ác tính của nốt phổi đơn độc. *Tạp chí Nghiên cứu Y học*. 2023;165(4):17-24. doi:10.52852/tcncyh.v165i4.1480

Summary

INITIAL APPLICATION OF QURE ARTIFICIAL INTELLIGENCE FOR CHEST X-RAY ANALYSIS IN DIAGNOSING PULMONARY TUBERCULOSIS

Tuberculosis is still a global health concern with yearly increased number of people diagnosed with pulmonary TB. Therefore, the issue of community-based TB screening and early diagnosis of pulmonary TB is an urgent need. Artificial intelligence was applied for the diagnosis of pulmonary tuberculosis to target early detection of lesions and accurate diagnosis. The purpose of our study was to initially apply Qure.AI for chest X-ray in the diagnosis of pulmonary tuberculosis.

The study described 126 patients suspected for pulmonary TB; Qure.AI read their chest Xrays and microbiological or histopathological tests were conducted to diagnose pulmonary TB. The final diagnosis were compared with AI. We found a good consensus between the 2 methods. AI sensitivity was 88.8%, specificity was 48.8%, predictive value was negative 71%. The area under the ROC curve was 77.1% with $p < 0.001$, 95%CI: 0.69 - 0.86. Thus, the chest X-ray images read by Qure.AI were valuable in diagnosing pulmonary tuberculosis with a fairly good level of accuracy with cut-off score of 0.503. Our research showed that AI had the potential to become an essential tool in supporting early diagnosis of pulmonary TB, and community lung TB screening.

Keywords: Tuberculosis, artificial intelligence Qure.AI.