

DOI: 10.58490/ctjump.2026i102.4724

ĐÁNH GIÁ KHẢ NĂNG PHÁT HIỆN TỔN THƯƠNG TRÊN X-QUANG NGỰC BẰNG CÔNG CỤ QURE.AI Ở BỆNH NHÂN VIÊM PHỔI CỘNG ĐỒNG

Ngô Phương Dũng, Dương Thị Thanh Vân, Huỳnh Trường Khánh,
Lâm Trường Hưng, Trần Gia Hân, Huỳnh Thanh Huy, Tu Thiện Thành, Bùi Quốc Nhân*

Trường Đại học Y Dược Cần Thơ

**Email: dungngophuong19@gmail.com*

Ngày nhận bài: 13/5/2026

Ngày phản biện: 29/7/2026

Ngày duyệt đăng: 25/8/2026

TÓM TẮT

Đặt vấn đề: Ứng dụng trí tuệ nhân tạo (Artificial Intelligence – AI) trong chẩn đoán hình ảnh ngày càng phổ biến và trở thành xu hướng của y học hiện đại. Viêm phổi là bệnh nhiễm trùng thường gặp với tỷ lệ tử vong cao. AI trong phân tích X-quang ngực giúp phát hiện sớm và nâng cao độ chính xác chẩn đoán. **Mục tiêu nghiên cứu:** Đánh giá giá trị của công cụ Qure.ai trong phân tích hình ảnh X-quang ngực ở bệnh nhân viêm phổi cộng đồng. **Đối tượng và phương pháp nghiên cứu:** Nghiên cứu mô tả cắt ngang thực hiện trên 311 bệnh nhân viêm phổi cộng đồng nhập viện tại Bệnh viện Đa khoa Trung ương Cần Thơ và Bệnh viện Trường Đại học Y Dược Cần Thơ từ tháng 06/2025 đến tháng 01/2026. Hình ảnh X-quang ngực được phân tích bằng công cụ Qure.ai và đối chiếu với tiêu chuẩn tham chiếu. Hệ số Kappa, độ nhạy, độ đặc hiệu và các giá trị chẩn đoán được tính toán bằng phần mềm SPSS 27.0. **Kết quả:** Tuổi trung bình của bệnh nhân là $71,3 \pm 13,4$ tuổi. Mức độ đồng thuận giữa Qure.ai và bác sĩ chẩn đoán hình ảnh trong phát hiện các tổn thương viêm phổi dao động từ trung bình đến cao. Tổn thương tràn dịch màng phổi có mức độ tương thích cao nhất ($\kappa = 0,708$). Trong nhóm bệnh nhân được chụp cắt lớp vi tính ngực làm tiêu chuẩn vàng, Qure.ai phát hiện tổn thương đông đặc/đám mờ với độ nhạy 91,7%. Đối với 60 ca có tổn thương trên CLVT ngực, công cụ Qure.ai có độ nhạy cao hơn bác sĩ (98,3% so với 91,7%), khác biệt không có ý nghĩa thống kê ($p = 0,125$). **Kết luận:** Qure.ai cho thấy tiềm năng hỗ trợ trong việc phát hiện các tổn thương viêm phổi trên X-quang ngực.

Từ khóa: Viêm phổi cộng đồng, X-quang ngực, trí tuệ nhân tạo, Qure.ai.

ABSTRACT

EVALUATION OF THE CAPABILITY OF QURE.AI TOOL IN DETECTING CHEST X-RAY LESIONS IN PATIENTS WITH COMMUNITY-ACQUIRED PNEUMONIA

Ngo Phuong Dung, Duong Thi Thanh Van, Huynh Truong Khanh,
Lam Truong Hung, Tran Gia Han, Huynh Thanh Huy, Tu Thien Thanh, Bui Quoc Nhan*

Can Tho University of Medicine and Pharmacy

Background: The application of Artificial Intelligence (AI) in diagnostic imaging is increasingly widespread and has become a major trend in modern medicine. Pneumonia is a common infectious disease associated with high mortality. The use of AI in analyzing chest X-rays helps in early detection and improves diagnostic accuracy. **Objectives:** To evaluate the value of the Qure.ai tool in analyzing chest X-ray images in patients with community-acquired pneumonia. **Materials and methods:** A cross-sectional descriptive study was conducted on 311 patients with community-acquired pneumonia admitted to Can Tho Central General Hospital and Can Tho University of Medicine and Pharmacy Hospital from June 2025 to January 2026. Chest X-ray

images were analyzed using the Qure.ai tool and compared with the reference standard. The kappa coefficient, sensitivity, specificity, and diagnostic values were calculated using SPSS 27.0 software.

Results: The mean age of the patients was 71.3 ± 13.4 years. The level of agreement between Qure.ai and radiologists in detecting pneumonia-related lesions ranged from moderate to high. Pleural effusion lesions showed the highest level of agreement ($\kappa = 0.708$). In the subgroup of patients who underwent chest computed tomography as the gold standard, Qure.ai detected consolidation/opacity lesions with a sensitivity of 91.7%. Among the 60 cases with confirmed lesions on chest CT, the Qure.ai tool exhibited higher sensitivity compared to radiologists (98.3% versus 91.7%), though the difference was not statistically significant ($p = 0.125$). **Conclusion:** Qure.ai demonstrates significant potential as a supportive tool for detecting pneumonic lesions on chest radiographs.

Keywords: Community-acquired pneumonia, Chest X-ray, Artificial intelligence, Qure.ai.

I. ĐẶT VẤN ĐỀ

Viêm phổi cộng đồng (VPCĐ) là một trong những bệnh lý nhiễm trùng hô hấp cấp tính phổ biến, gây gánh nặng lớn cho hệ thống y tế toàn cầu. X-quang ngực là phương tiện chẩn đoán hình ảnh cơ bản, được sử dụng rộng rãi để xác định tổn thương phổi trong VPCĐ. Tuy nhiên, việc chẩn đoán bệnh VPCĐ dựa vào X-quang ngực vẫn còn nhiều hạn chế. Việc đọc và phân tích hình ảnh X-quang ngực có sự khác nhau giữa các bác sĩ [1]. X-quang ngực là kỹ thuật chẩn đoán hình ảnh giúp sàng lọc nhưng vẫn chưa đặc hiệu cho chẩn đoán (dương tính giả hoặc âm tính giả) nên dễ bỏ sót hoặc quá mức tổn thương, vì vậy ảnh hưởng trực tiếp đến chẩn đoán và quyết định điều trị của bác sĩ lâm sàng.

Hiện nay, việc ứng dụng trí tuệ nhân tạo (Artificial Intelligence) trong y học là một cột mốc quan trọng trong sự phát triển khoa học công nghệ. Trong việc đọc X-quang ngực, sự phát triển các phần mềm, mô hình AI có sẵn giúp ích nhiều cho các bác sĩ lâm sàng, hỗ trợ đưa ra chẩn đoán bệnh một cách nhanh chóng và hiệu quả. Qure.ai là một công cụ AI ứng dụng trong chẩn đoán hình ảnh, với khả năng phát hiện nhiều dạng tổn thương giống với quy trình đọc phim X-quang ngực thông thường. Xuất phát từ thực tiễn trên, nghiên cứu này được thực hiện với mục tiêu: Đánh giá mức độ tương thích và giá trị chẩn đoán của công cụ này khi đối chiếu với kết quả đọc X-quang ngực của bác sĩ chẩn đoán hình ảnh, siêu âm lồng ngực và cắt lớp vi tính ngực.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Bệnh nhân (BN) ≥ 18 tuổi được chẩn đoán viêm phổi cộng đồng (VPCĐ) nhập viện tại Khoa Nội hô hấp Bệnh viện Đa khoa Trung ương Cần Thơ và Khoa Nội tổng hợp Bệnh viện Trường Đại học Y Dược Cần Thơ từ tháng 06/2025 đến tháng 01/2026.

- Tiêu chuẩn lựa chọn:

+ Bệnh nhân được chẩn đoán VPCĐ theo hướng dẫn của Hiệp hội Lồng ngực Anh (BTS) [2] khi có các triệu chứng sau: Các triệu chứng của nhiễm trùng đường hô hấp dưới cấp tính (ho và có ít nhất 1 triệu chứng của đường hô hấp dưới khác như khạc đờm, khó thở, đau ngực). Có ít nhất một trong các dấu hiệu toàn thân (sốt, vã mồ hôi, đau mỏi người và/hoặc nhiệt độ cơ thể tăng trên 38°C). Hình ảnh X-quang ngực: Tổn thương dạng thâm nhiễm mới xuất hiện. Các triệu chứng này xuất hiện ngoài bệnh viện hoặc trong vòng 48 giờ sau nhập viện.

+ Bệnh nhân có đầy đủ thông tin lâm sàng, cận lâm sàng cần thiết để phục vụ nghiên cứu. Bệnh nhân hoặc người thân đồng ý tham gia nghiên cứu.

- Tiêu chuẩn loại trừ: Hình ảnh X-quang không đạt tiêu chuẩn để đánh giá.

2.2. Phương pháp nghiên cứu

- **Thiết kế nghiên cứu:** Nghiên cứu mô tả cắt ngang.

- **Cỡ mẫu và chọn mẫu:**

$$n = \frac{Z^2_{(1-\frac{\alpha}{2})} \times p \times (1 - p)}{d^2}$$

Trong đó:

n: Cỡ mẫu

α : Mức ý nghĩa thống kê, chọn $\alpha = 0,05$ ứng với khoảng tin cậy 95%

Z: Hệ số tin cậy, $Z^2_{(1-\frac{\alpha}{2})} = 1,96^2$

Chọn d: Sai số mong muốn, $d = 0,03$

Chọn $p = 0,954$, là độ nhạy của công cụ AI trong phát hiện tổn thương viêm phổi trên X-quang ngực, dựa theo nghiên cứu của Becker và cộng sự (2022) [3]

Vậy cỡ mẫu tối thiểu cần thiết cho nghiên cứu là 188 mẫu.

- **Phương pháp chọn mẫu:** Chọn mẫu thuận tiện với bệnh nhân đã được chẩn đoán Viêm phổi cộng đồng.

- **Công cụ Qure.ai:** Qure.ai là một công cụ AI ứng dụng trong chẩn đoán hình ảnh, với khả năng phát hiện nhiều dạng tổn thương giống với quy trình đọc phim X-quang ngực thông thường. Các dạng tổn thương được mô tả trên Qure.ai:

Opacity – Đám mờ: Các vùng tăng đậm độ (trắng hơn bình thường) trong nhu mô phổi nhưng ranh giới không rõ.

Consolidation – Đông đặc: Tình trạng nhu mô phổi mất chứa khí và được thay thế bằng dịch, mủ hoặc các tế bào, biểu hiện bằng vùng mờ đồng nhất che lấp các mạch máu phổi.

Pleural Effusion – Tràn dịch màng phổi: Sự tích tụ dịch bất thường trong khoang màng phổi, làm mờ góc sườn hoành hoặc tạo đường cong Damoiseau.

Nodule – Nốt phổi: Tổn thương dạng khối mờ hình tròn hoặc bầu dục, có đường kính nhỏ (thường dưới 3 cm), nằm đơn độc hoặc rải rác trong nhu mô phổi.

Fibrosis – Xơ hóa: Các dải mờ dạng sợi hoặc lưới không đều.

Cardiomegaly – Bóng tim to: Tình trạng kích thước tim bất thường, được xác định khi chỉ số tim/ lồng ngực trên phim X-quang ngực thẳng lớn hơn 0,5.

- **Quy trình nghiên cứu:** Nhóm nghiên cứu đã thu thập được 567 hình ảnh X-quang ngực của 311 bệnh nhân, chia thành 3 phân nhóm để phân tích kết quả:

Phân nhóm 1: So sánh khả năng phát hiện tổn thương của công cụ Qure.ai và Kết luận của bác sĩ chẩn đoán hình ảnh khi đọc phim X-quang ngực ($n = 311$).

Phân nhóm 2: So sánh khả năng phát hiện tổn thương tràn dịch màng phổi của công cụ Qure.ai và Kết luận của bác sĩ chẩn đoán hình ảnh khi siêu âm lồng ngực ($n = 200$).

Phân nhóm 3: So sánh khả năng phát hiện tổn thương của Qure.ai và kết luận của bác sĩ chẩn đoán hình ảnh khi đọc CLVT ngực ($n = 72$) (sử dụng hình ảnh X-quang gần nhất với lúc chụp CLVT ngực).

- **Phương tiện nghiên cứu và phương pháp xử lý số liệu:** Hình ảnh X-quang ngực được phân tích bằng công cụ Qure.ai và đối chiếu với các tiêu chuẩn tham chiếu: Kết luận của bác sĩ chẩn đoán hình ảnh khi đọc phim X quang ngực, phim cắt lớp vi tính ngực và siêu âm lồng ngực. Các biến số so sánh được trình bày dưới dạng bảng chéo (crosstabs) với tần số và tỷ lệ phần trăm. Kết quả cắt lớp vi tính ngực, X-quang ngực và siêu âm lồng ngực được diễn giải độc lập bởi các bác sĩ chẩn đoán hình ảnh tại khoa Chẩn đoán hình ảnh của hai bệnh viện thực hiện nghiên cứu (có trên 5 năm kinh nghiệm), kết luận của bác sĩ chẩn

đoán hình ảnh là độc lập hoàn toàn so với kết quả thu được từ Qure.ai. Hệ số kappa của Cohen (Cohen's kappa coefficient - κ), độ nhạy, độ đặc hiệu và các giá trị chẩn đoán được tính toán bằng phần mềm SPSS 27.0.

- **Đạo đức trong nghiên cứu:** Nghiên cứu được Hội đồng Đạo đức trong nghiên cứu y sinh học của Trường Đại học Y Dược Cần Thơ thông qua với phiếu chấp thuận số 25.042.SV/PCT-HĐĐĐ ngày 12/05/2025.

III. KẾT QUẢ NGHIÊN CỨU

Tổng số 311 bệnh nhân nghiên cứu, độ tuổi trung bình $71,3 \pm 13,4$ với 567 hình ảnh X-quang thu thập được, kết quả như sau:

Bảng 1. Khả năng phát hiện tổn thương của công cụ Qure.ai và Bác sĩ chẩn đoán hình ảnh khi đọc phim X-quang ngực (n = 311)

Qure.ai		Kết luận X-quang ngực của Bác sĩ chẩn đoán hình ảnh		Tổng	p	κ	Khoảng tin cậy 95%
		Có n (%)	Không n (%)				
Tổn thương đồng đặc/ Đám mờ	Có n (%)	192 (93,7%)	59 (55,7%)	251	<0,001	0,424	0,319 - 0,530
	Không n (%)	13 (6,3%)	47 (44,3%)	60			
	Tổng	205	106	311			
Tràn dịch màng phổi	Có n (%)	73 (86,9%)	27 (11,9%)	100	<0,001	0,708	0,622 - 0,793
	Không n (%)	11 (13,1%)	200 (88,1%)	211			
	Tổng	84	227	311			
Bóng tim to	Có n (%)	12 (38,7%)	29 (10,4%)	41	<0,001	0,248	0,097 - 0,399
	Không n (%)	19 (61,3%)	251 (89,6%)	270			
	Tổng	31	280	311			
Xơ hóa	Có n (%)	26 (44,1%)	14 (5,6%)	40	<0,001	0,439	0,307 - 0,572
	Không n (%)	33 (55,9%)	238 (94,4%)	271			
	Tổng	59	252	311			
Nốt	Có n (%)	11 (50%)	61 (21,1%)	72	<0,001	0,141	0,029 - 0,253
	Không n (%)	11 (50%)	228 (78,9%)	239			
	Tổng	22	289	311			

Nhận xét: Trong các loại tổn thương, tràn dịch màng phổi là tổn thương có mức độ đồng thuận tốt nhất ($\kappa = 0,708$; $p < 0,001$). Tổn thương đồng đặc/ đám mờ có độ đồng thuận trung bình ($\kappa = 0,424$). Tổn thương xơ hóa cũng có mức độ đồng thuận trung bình ($\kappa = 0,439$). Tuy nhiên, tổn thương dạng nốt và bóng tim to có mức độ đồng thuận kém ($\kappa = 0,141$ và $0,248$)

Bảng 2. Khả năng phát hiện tổn thương của công cụ Qure.ai và Bác sĩ chẩn đoán hình ảnh khi siêu âm lồng ngực (n = 200)

Qure.ai		Kết luận siêu âm lồng ngực của bác sĩ chẩn đoán hình ảnh		Tổng	p	κ	Độ nhạy	Độ đặc hiệu
		Có n (%)	Không n (%)					
Tràn dịch màng phổi	Có n (%)	52 (59,1%)	11 (9,8%)	63	<0,001	0,508 (0,390 - 0,626)	59,1%	90,2%
	Không n (%)	36 (40,9%)	101 (90,2%)	137				
	Tổng	88	112	200				

Nhận xét: Khi đối chiếu với siêu âm lồng ngực trong phát hiện tổn thương tràn dịch màng phổi, công cụ Qure.ai cho thấy độ đồng thuận ở mức trung bình ($\kappa = 0,508$), với độ đặc hiệu (90,2%). Tuy nhiên, độ nhạy chỉ đạt 59,1%.

Bảng 3. Khả năng phát hiện tổn thương của công cụ Qure.ai và Bác sĩ chẩn đoán hình ảnh khi đọc CLVT ngực (n = 72)

Qure.ai		Kết luận CLVT ngực của bác sĩ chẩn đoán hình ảnh		Tổng	p	κ	Khoảng tin cậy 95%
		Có n (%)	Không n (%)				
Tổn thương đồng đặc/Đám mờ	Có n (%)	55 (91,7%)	5 (41,7%)	60	0,001	0,500	0,232 - 0,768
	Không n (%)	5 (8,3%)	7 (58,3%)	12			
	Tổng	60	12	72			
Tràn dịch màng phổi	Có n (%)	15 (50%)	9 (21,4%)	24	0,011	0,294	0,072 - 0,516
	Không n (%)	15 (50%)	33 (78,6%)	48			
	Tổng	30	42	72			
Xơ hóa	Có n (%)	2 (5,3%)	3 (8,8%)	5	0,553	-0,034	-0,147 - 0,080
	Không n (%)	36 (94,7%)	31 (91,2%)	67			
	Tổng	38	34	72			
Nốt	Có n (%)	4 (33,3%)	10 (16,7%)	14	0,183	0,156	-0,107 - 0,420
	Không n (%)	8 (66,7%)	50 (83,3%)	58			
	Tổng	12	60	72			

Nhận xét: Trong phân nhóm 72 bệnh nhân có kết quả chụp cắt lớp vi tính ngực, đồng thuận mức trung bình ghi nhận ở tổn thương đồng đặc/đám mờ ($\kappa = 0,500$). Sự đồng thuận thấp đối với tràn dịch màng phổi ($\kappa = 0,294$) và nốt ($\kappa = 0,156$).

Bảng 4. Giá trị chẩn đoán của Qure.ai khi lấy CLVT ngực làm tiêu chuẩn vàng (n = 72)

	Độ nhạy	Độ đặc hiệu	Giá trị tiên đoán dương (PPV)	Giá trị tiên đoán âm (NPV)	Độ chính xác
Tổn thương đông đặc/ Đám mờ	91,7%	58,3%	91,7%	58,3%	86,1%
Tràn dịch màng phổi	50%	78,6%	62,5%	68,8%	66,7%
Xơ hóa	5,3%	91,2%	40%	46,3%	45,8%
Nốt	33,3%	83,3%	28,5%	86,2%	75%

Nhận xét: Đối với tổn thương đặc trưng của viêm phổi là đông đặc/đám mờ, AI đạt độ nhạy 91,7% và giá trị tiên đoán dương là 91,7%. Đối với các tổn thương mạn tính như xơ hóa hay nốt phổi, AI có độ đặc hiệu lần lượt là 91,2% và 83,3% nhưng độ nhạy chỉ 5,3% và 33,3%.

Bảng 5. So sánh hiệu năng phát hiện tổn thương giữa công cụ Qure.ai và Bác sĩ chẩn đoán hình ảnh khi đọc X-quang ngực (Đối chiếu tiêu chuẩn vàng CLVT ngực, n = 72)

	Công cụ Qure.ai	Bác sĩ chẩn đoán hình ảnh khi đọc X-quang ngực	p (McNemar Test)
Độ nhạy	59/60 (98,3%)	55/60 (91,7%)	0,125
Độ đặc hiệu	5/12 (41,7%)	7/12 (58,3%)	0,500

Nhận xét: Đối với 60 ca có tổn thương trên CLVT ngực, công cụ Qure.ai có độ nhạy cao hơn bác sĩ (98,3% so với 91,7%), khác biệt không có ý nghĩa thống kê ($p = 0,125$). Ngược lại, ở 12 ca âm tính, Bác sĩ có độ đặc hiệu cao hơn Qure.ai (58,3% so với 41,7%), khác biệt cũng không có ý nghĩa thống kê ($p = 0,500$).

Bảng 6. So sánh hiệu năng phát hiện tổn thương tràn dịch màng phổi giữa công cụ Qure.ai và Siêu âm lồng ngực (Đối chiếu với tiêu chuẩn vàng là CLVT ngực, n = 72)

	Công cụ Qure.ai	Siêu âm lồng ngực	p (McNemar test)
Độ nhạy	15/30 (50,0%)	27/30 (90,0%)	0,002
Độ đặc hiệu	33/42 (78,6%)	37/42 (88,1%)	0,219

Nhận xét: Trong 30 ca tràn dịch màng phổi theo CLVT ngực, siêu âm lồng ngực có độ nhạy cao hơn rõ rệt so với Qure.ai (90,0% so với 50,0%), sự khác biệt có ý nghĩa thống kê ($p = 0,002$). Đối với 42 ca không tràn dịch, siêu âm lồng ngực đạt độ đặc hiệu cao hơn Qure.ai (88,1% so với 78,6%) nhưng khác biệt không có ý nghĩa thống kê ($p = 0,219$).

IV. BÀN LUẬN

4.1. Sự đồng thuận giữa AI và bác sĩ chẩn đoán hình ảnh

Nghiên cứu này trên 311 bệnh nhân viêm phổi cộng đồng (VPCĐ) cho thấy mức độ đồng thuận giữa Qure.ai và bác sĩ chẩn đoán hình ảnh có sự biến thiên tùy thuộc vào loại tổn thương. Cụ thể, tổn thương tràn dịch màng phổi có mức độ đồng thuận cao nhất ($\kappa = 0,708$, $p < 0,001$), phản ánh khả năng nhận diện hình ảnh dịch màng phổi khá rõ ràng và khách quan của cả AI lẫn mắt thường. Đối với tổn thương đặc trưng của viêm phổi là đông đặc/đám mờ, mức độ đồng thuận đạt mức trung bình ($\kappa = 0,424$). Đặc biệt, tổn thương dạng nốt và bóng tim to có sự đồng thuận rất thấp (lần lượt = 0,141 và 0,248). Sự đồng thuận ở mức trung bình đến thấp ở một số tổn thương có thể được giải thích do các thuật toán học

sâu (Deep Learning) thường được thiết lập ngưỡng cắt (cut-off) thiên về tối ưu hóa độ nhạy để đóng vai trò như một công cụ sàng lọc, dẫn đến việc AI có xu hướng "nhảy quá mức" và ghi nhận nhiều tổn thương hơn so với đánh giá lâm sàng của bác sĩ. Điều này hoàn toàn phù hợp với ghi nhận trong nghiên cứu của Kvak năm 2023 [4] khi so sánh AI với bác sĩ, trong đó AI có xu hướng đánh giá nhiều hình ảnh bình thường thành bất thường do thiết lập ngưỡng nhạy cao.

4.2. Giá trị chẩn đoán của Qure.ai so với tiêu chuẩn vàng (CLVT ngực và Siêu âm)

Khi đối chiếu với CLVT ngực làm tiêu chuẩn vàng trên 72 bệnh nhân, Qure.ai thể hiện hiệu năng xuất sắc trong việc phát hiện tổn thương đông đặc/đám mờ với độ nhạy lên tới 91,7% và giá trị tiên đoán dương (PPV) là 91,7%, trong khi độ đặc hiệu đạt 58,3%. Kết quả này có tính tương đồng rất cao với nghiên cứu của Becker năm 2022 [3] trên bệnh nhân cấp cứu, trong đó AI phát hiện tổn thương viêm phổi với độ nhạy 95,4%, độ đặc hiệu 66,0% và diện tích dưới đường cong (AUC) đạt 0,923. Tương tự, nghiên cứu của Kvak năm 2023 [4] cũng báo cáo độ nhạy của AI trong phát hiện bất thường trên X-quang là 0,925 và độ đặc hiệu là 0,644. Bên cạnh đó, các nghiên cứu trong nước sử dụng phần mềm Qure.ai trên bệnh nhân nghi lao phổi của tác giả Lê Hoàn năm 2023 [5] cũng ghi nhận độ nhạy cao của Qure.ai (88,8%) và độ đặc hiệu ở mức trung bình (48,8%). Như vậy, có thể thấy đặc tính chung của Qure.ai là có độ nhạy rất cao nhưng độ đặc hiệu chỉ ở mức trung bình - khá. Trong bối cảnh bệnh lý nhiễm trùng hô hấp cấp tính như VPCĐ, một công cụ có độ nhạy > 90% như Qure.ai là cực kỳ giá trị để tầm soát, giúp bác sĩ cấp cứu không bỏ sót các tổn thương nguy hiểm [3]. Đối với tràn dịch màng phổi, khi so sánh với siêu âm lồng ngực (n=200), Qure.ai có độ đặc hiệu rất cao (90,2%) nhưng độ nhạy chỉ đạt 59,1%. Điều này cho thấy khi Qure.ai báo cáo có tràn dịch thì khả năng chính xác là rất cao, nhưng phần mềm có thể bỏ sót các trường hợp tràn dịch lượng ít mà siêu âm có thể dễ dàng phát hiện. Theo nghiên cứu của Al Nufaei năm 2025 [6], siêu âm lồng ngực có độ nhạy (86-98%) cao hơn hẳn X-quang (70-85%) trong phát hiện tràn dịch màng phổi và đông đặc nhu mô. Mặc dù CT scan là tiêu chuẩn vàng với độ nhạy cao nhất (> 95%), nhưng việc sử dụng thường bị hạn chế do chi phí cao và nguy cơ phơi nhiễm bức xạ [7].

4.3. Hiệu năng chẩn đoán giữa các kỹ thuật dưới sự đối chiếu của tiêu chuẩn vàng

Khi đối chiếu trực tiếp với tiêu chuẩn vàng CLVT ngực, kết quả nghiên cứu cho thấy sự khác biệt đặc trưng về hiệu năng giữa các kỹ thuật. Đối với phát hiện các tổn thương trên viêm phổi (Bảng 5), công cụ Qure.ai đạt độ nhạy sàng lọc rất cao là 98,3% (59/60 ca), vượt trội hơn so với kết quả đọc của Bác sĩ chẩn đoán hình ảnh (91,7%), dù sự khác biệt chưa có ý nghĩa thống kê với $p = 0,125$. Ngược lại, Bác sĩ có ưu thế hơn về độ đặc hiệu với 58,3% so với 41,7% của AI ($p = 0,500$), điều này phù hợp với các ghi nhận của Becker [3] và Wu [7] về xu hướng tối ưu hóa độ nhạy sàng lọc của các thuật toán học sâu nhằm hạn chế bỏ sót tổn thương nguy hiểm. Đối với tổn thương tràn dịch màng phổi (Bảng 6), siêu âm lồng ngực thể hiện năng lực vượt trội rõ rệt so với Qure.ai về độ nhạy phát hiện dịch với 90,0% so với 50,0%, sự khác biệt có ý nghĩa thống kê rất rõ ràng với $p = 0,002$. Kết quả thực nghiệm này tương đồng với tổng quan y văn của Al Nufaei (2025) [6], khẳng định giá trị cao của siêu âm lồng ngực đối với các tổn thương khoang màng phổi lượng ít hoặc bị che lấp. Tuy nhiên, về độ đặc hiệu trong phát hiện tràn dịch, hai kỹ thuật đều đạt mức cao và tương đương nhau với 88,1% ở siêu âm lồng ngực và 78,6% ở AI ($p = 0,219$). Sự phân định hiệu năng này minh chứng rằng công cụ AI đóng một vai trò nhất định trong việc tầm soát

diện rộng tổn thương nhu mô phổi, trong khi siêu âm lồng ngực là phương tiện hỗ trợ đắc lực và không thể thay thế đối với các tổn thương màng phổi chuyên biệt.

V. KẾT LUẬN

Qua nghiên cứu trên 311 bệnh nhân viêm phổi, công cụ Qure.ai thể hiện mức độ đồng thuận từ trung bình đến tốt với các bác sĩ chẩn đoán hình ảnh trong việc phát hiện các tổn thương chính, đạt giá trị cao nhất đối với tràn dịch màng phổi ($\kappa = 0,708$). Tuy nhiên, sự đồng thuận còn thấp với các tổn thương dạng nốt ($\kappa = 0,141$) do xu hướng phát hiện nhạy quá mức của thuật toán. Khi đối chiếu với tiêu chuẩn vàng CLVT ngực, Qure.ai phát hiện tốt đối với tổn thương đặc trưng đông đặc/đám mờ với độ nhạy và giá trị tiên đoán dương đều đạt 91,7%. Trong phân tích so sánh các kỹ thuật, CLVT ngực là kết quả tham chiếu, công cụ Qure.ai đạt độ nhạy phát hiện các tổn thương viêm phổi ở mức 98,3% so với Bác sĩ chẩn đoán hình ảnh (91,7%, $p = 0,125$) dù có độ đặc hiệu thấp hơn (41,7% so với 58,3%). Ngược lại, đối với tổn thương tràn dịch màng phổi, siêu âm lồng ngực thể hiện năng lực phát hiện vượt trội rõ rệt so với công cụ AI với độ nhạy lần lượt là 90,0% và 50,0% ($p = 0,002$). Qure.ai cho thấy tiềm năng hỗ trợ trong việc phát hiện các tổn thương viêm phổi trên X-quang ngực. Việc ứng dụng công cụ này giúp tối ưu hóa quy trình chẩn đoán sớm viêm phổi cộng đồng, đặc biệt hiệu quả tại các cơ sở y tế đang đối mặt với tình trạng quá tải hoặc thiếu hụt nhân lực chuyên khoa.

TÀI LIỆU THAM KHẢO

1. Hoàng Thụy Đoan Trâm, Lê Trung Thi, Châu Hữu Hầu. Đánh giá hệ thống thang điểm TSS và Brixia trong X-quang ngực ở bệnh nhân mắc bệnh COVID-19. *Tạp chí Y học Việt Nam*. 2022. 510(1), 219-223, doi: 10.51298/vmj.v510i1.1935.
2. Nottinghamshire Area Prescribing Committee. Community-acquired Pneumonia (CAP) V2. 2025.
3. Becker J., Decker J.A., Römmele C., Kahn M., Messmann H., *et al.* Artificial Intelligence-Based Detection of Pneumonia in Chest Radiographs. *Diagnostics*. 2022. 12, 1465, doi: 10.3390/diagnostics12061465.
4. Kvak D., Chromcová A., Biroš M., Hrubý R., Kvaková K., *et al.* Chest X-ray Abnormality Detection by Using Artificial Intelligence: A Single-Site Retrospective Study of Deep Learning Model Performance. *Biomedinformatics*. 2023. 3, 82-101, doi: 10.3390/biomedinformatics3010006.
5. Lê Hoàn, Nguyễn Thị Thu Thủy, Lê Minh Hằng. Ứng dụng bước đầu trí tuệ nhân tạo Qure.AI X-quang lồng ngực trong chẩn đoán lao phổi. *Tạp chí Nghiên cứu Y học*. 2023. 171(10), 306-312, doi: 10.52852/tcncyh.v171i10.2066.
6. Al Nufaiei Z.F., Alshamrani K.M. Comparing Ultrasound, Chest X-Ray, and CT Scan for Pneumonia Detection. *Medical Devices: Evidence and Research*. 2025. 18, 149-159, doi: 10.2147/MDER.S501714.
7. Wu J.T., Wong K.C.L., Gur Y., Ansari N., Karagyris A., *et al.* Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. *JAMA Network Open*. 2020, 3(10), e2022779, doi: 10.1001/jamanetworkopen.2020.22779.