

KẾT QUẢ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG HỖ TRỢ PHÁT HIỆN NỐT/KHỐI PHỔI TRÊN PHIM X-QUANG NGỰC THẲNG

Bùi Trọng Dương¹, Phùng Anh Tuấn^{1,2}*

Tóm tắt

Mục tiêu: Đánh giá kết quả ứng dụng phần mềm trí tuệ nhân tạo (artificial intelligence - AI) DrAid trong phát hiện tổn thương dạng nốt/khối ở phổi trên phim X-quang ngực thẳng. **Phương pháp nghiên cứu:** Nghiên cứu tiến cứu, mô tả cắt ngang trên 637 bệnh nhân (BN) được chụp phim X-quang, cắt lớp vi tính (CLVT) lồng ngực và điều trị tại Bệnh viện Trung ương Quân đội 108 từ tháng 10/2024 - 3/2025. Kết quả chẩn đoán của phần mềm DrAid, của bác sỹ chẩn đoán hình ảnh (CĐHA) và của bác sỹ CĐHA sau khi tham khảo kết quả của phần mềm DrAid được đối chiếu với kết quả chụp CLVT. Đánh giá thông qua tính độ nhạy (Se), độ đặc hiệu (Sp) và diện tích dưới đường cong (AUROC). **Kết quả:** DrAid có Se, AUROC cao hơn nhưng Sp lại thấp hơn so với kết quả của bác sỹ CĐHA. Khi có sự hỗ trợ của DrAid, chỉ số Se, Sp và AUROC trong phát hiện nốt/khối phổi đều tăng. **Kết luận:** DrAid là công cụ hiệu quả trong việc hỗ trợ nâng cao chất lượng tầm soát các nốt/khối phổi trên phim chụp X-quang ngực thẳng.

Từ khóa: Trí tuệ nhân tạo; X-quang ngực thẳng; Nốt/khối phổi.

EFFECTIVENESS OF ARTIFICIAL INTELLIGENCE IN DETECTING PULMONARY NODULES/MASSES ON CHEST X-RAY

Abstract

Objectives: To evaluate the efficacy of the DrAid artificial intelligence (AI) software in detecting pulmonary nodules/masses on chest X-rays. **Methods:** A prospective, cross-sectional descriptive study was conducted on 637 patients who underwent posteroanterior chest X-rays and chest computed tomography (CT) scans during their treatment at 108 Military Central Hospital from October 2024 to March 2025.

¹Học viện Quân y

²Bệnh viện Quân y 103, Học viện Quân y

*Tác giả liên hệ: Phùng Anh Tuấn (phunganhtuanbv103@gmail.com)

Ngày nhận bài: 16/6/2025

Ngày được chấp nhận đăng: 23/7/2025

<http://doi.org/10.56535/jmpm.v50i7.1399>

The results of AI, the radiologists' results, and the radiologists' results after consulting AI were compared with the results of CT scans. The sensitivity (Se), specificity (Sp), and area under the receiver operating characteristic curve (AUROC) of analyses were assessed. **Results:** The DrAid demonstrated higher Se and AUROC compared to radiologists, but lower Sp. When assisted by AI, radiologists exhibited improvements in Se, Sp, and AUROC. **Conclusion:** DrAid is an effective tool for enhancing the quality of screening for pulmonary nodules/masses on chest X-rays.

Keywords: Artificial intelligence; Posteroanterior chest X-rays; Pulmonary nodule/mass.

ĐẶT VẤN ĐỀ

Ung thư phổi là loại ung thư phổ biến thứ hai ở cả nam giới và nữ giới, tuy nhiên lại là nguyên nhân ung thư gây tử vong phổ biến nhất, nhiều hơn cả tổng số ung thư vú, trực tràng và tiền liệt tuyến cộng lại [1]. Tầm soát ung thư và các tổn thương dạng nốt, khối ở phổi hiện tại vẫn chủ yếu sử dụng hình ảnh X-quang ngực, với các ưu điểm về chi phí, tiện lợi và tính cơ động. Tuy nhiên, do sự chồng lấp của tổn thương và các cấu trúc xung quanh, hạn chế về độ phân giải hình ảnh, khiến cho Se và Sp của X-quang tương đối thấp. Khoảng 90% các trường hợp ung thư phổi bị bỏ sót do sai sót trong chẩn đoán trên hình ảnh X-quang phổi [2]. Phim chụp X-quang hầu như không thể phát hiện được các nốt phổi < 5mm, chỉ có Se khoảng 50% với nốt 6 - 10mm và thậm chí có thể bỏ sót các khối đường kính lên tới 35mm [3].

Ngày nay, AI đang là xu hướng tiềm năng để nâng cao hiệu quả của X-quang trong tầm soát ung thư phổi. Nhiều nghiên cứu đã được tiến hành chứng minh AI có khả năng tăng hiệu quả phát hiện các tổn thương dạng nốt/khối trên phim X-quang ngực [4 - 6]. Tuy nhiên, một nghiên cứu gần đây chỉ ra đối với các nốt nhỏ, Se của AI lại kém hơn so với bác sĩ CDHA [7]. Điều này làm tăng nghi ngại về hiệu quả hỗ trợ của AI, vốn được kỳ vọng để giảm thiểu tỷ lệ bỏ sót các nốt nhỏ trên phim X-quang phổi. Tại Việt Nam, phần mềm DrAid của VinBrain đã được bước đầu ứng dụng ở một vài bệnh viện để đánh giá kết quả phim chụp X-quang phổi, tuy nhiên hiệu quả thu được còn chưa rõ. Vì vậy, chúng tôi tiến hành nghiên cứu nhằm: *Đánh giá hiệu quả của phần mềm AI DrAid (gọi chung là AI) trong phát hiện các nốt/khối tổn thương phổi trên phim X-quang ngực thẳng.*

ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1. Đối tượng nghiên cứu

Gồm 637 bệnh nhân (BN) khám và điều trị tại Bệnh viện Trung ương Quân đội 108 từ tháng 10/2024 - 3/2025.

* *Tiêu chuẩn lựa chọn*: BN được chụp phim X-quang ngực thẳng và CLVT lồng ngực, dữ liệu lưu dưới định dạng DICOM; BN có thời gian chụp CLVT sau chụp X-quang ≤ 24 giờ đối với những tổn thương cấp tính và 72 giờ đối với những trường hợp khác; BN có hồ sơ bệnh án đầy đủ.

* *Tiêu chuẩn loại trừ*: Tổn thương trên phim CLVT có kích thước $< 6\text{mm}$; tổn thương trên phim CLVT nằm sau bóng tim, dưới vòm hoành và được xác định tương ứng với những vùng mờ trên phim X-quang ngực thẳng; BN từ chối tham gia nghiên cứu.

2. Phương pháp nghiên cứu

* *Thiết kế nghiên cứu*: Nghiên cứu mô tả cắt ngang, số liệu lấy theo hình thức tiền cứu.

* *Phương tiện sử dụng*:

AI được sử dụng trong nghiên cứu là phần mềm DrAid do VinBrain tiến hành, đã được FDA chấp thuận, đã thương mại hóa tại một số cơ sở y tế tại Việt Nam.

Chụp X-quang ngực thẳng và chụp CLVT lồng ngực được thực hiện theo quy trình đang áp dụng tại Khoa X-quang Chẩn đoán, Trung tâm CDHA, Bệnh viện

Trung ương Quân đội 108 trên máy X-quang kỹ thuật số Brivo DR-F của GE Medical Systems sản xuất năm 2021 và máy CLVT 32 dãy Access CT, hãng Philips, Hà Lan sản xuất năm 2020.

* *Các bước tiến hành nghiên cứu*:

Đưa dữ liệu phim chụp X-quang ngực thẳng vào phần mềm DrAid dưới định dạng chuẩn DICOM. Phần mềm tự động đưa ra đánh giá.

Bác sỹ chuyên ngành CDHA với 5 năm kinh nghiệm đánh giá phim X-quang ngực thẳng độc lập. Sau đó, bác sỹ được cung cấp kết quả của AI để tham khảo và đưa ra đánh giá lại.

Các kết quả của AI, của bác sỹ CDHA và của bác sỹ sau kết hợp với AI được đối chiếu với kết quả chụp CLVT lồng ngực do trường khoa xác định.

* *Biến số nghiên cứu*:

Nội dung đánh giá: Có hay không có tổn thương dạng nốt/khối, vị trí, số lượng tổn thương. Nốt là những hình mờ có kích thước $< 3\text{cm}$, có đậm độ đặc, bán đặc hoặc kính mờ. Khối là những hình mờ có kích thước $\geq 3\text{cm}$, có độ đậm đặc. Kích thước tổn thương là kích thước lớn nhất đo trên hình tái tạo coronal, cửa sổ nhu mô phim chụp CLVT.

* *Xử lý số liệu*: Các số liệu được phân tích bằng phần mềm SPSS 26.0 và Medcalc 23.2.1. Các kết quả phân tích của AI, bác sỹ và bác sỹ với hỗ trợ của AI được đánh giá Se, Sp, giá trị tiên đoán âm, giá trị tiên đoán dương, độ chính xác và AUROC. Kích thước trung

bình các nốt/khối được so sánh bằng kiểm định Mann-Whitney U. Độ tương đồng của các mô hình chẩn đoán được đánh giá bằng kiểm định Mc Nemar và Cochran's Q test đối với Se, Sp và Delong method với AUC.

3. Đạo đức nghiên cứu

Nghiên cứu đã được Hội đồng Đạo đức trong nghiên cứu Y sinh học của

Học viện Quân y thông qua theo Quyết định giao đề tài số 3487/QĐ-HVQY ngày 23/8/2024. Nghiên cứu được tiến hành độc lập với quá trình điều trị lâm sàng. Số liệu nghiên cứu được Bệnh viện Trung ương Quân đội 108 cho phép sử dụng và công bố. Nhóm tác giả cam kết không có xung đột về lợi ích trong nghiên cứu.

KẾT QUẢ NGHIÊN CỨU

Bảng 1. Đặc điểm tổn thương nốt/khối trên CLVT (n = 637).

Trường phổi	Số lượng, n (%)	Kích thước mm (median, min, max)
Chung	149 (23,39)	16 (6 - 135)
Trái	95 (14,91)	15 (7 - 135)
Phải	94 (14,76)	16,5 (6 - 80)
p	> 0,05	> 0,05

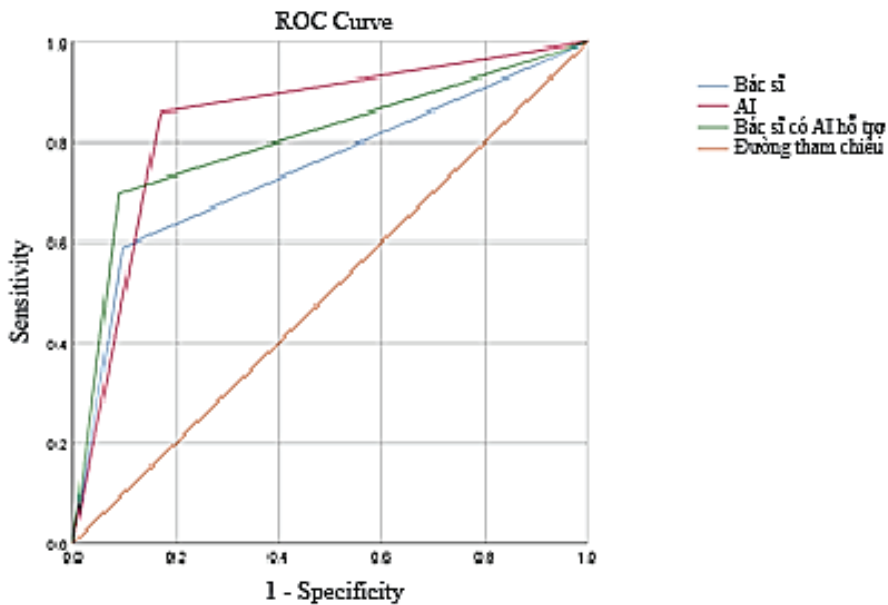
(p: Mann-Whitney U test)

Tổn thương dạng nốt/khối được phát hiện ở 149 BN (23,39%). Không có sự khác biệt về phân bố số lượng cũng như kích thước tổn thương ở hai bên.

Bảng 2. Mối liên quan giữa kết quả của các phương thức chẩn đoán với tiêu chuẩn đối chiếu (n = 637).

Phương thức		Kết quả CLVT	
		Không có tổn thương	Có tổn thương
Bác sỹ	Không có tổn thương	440	61
	Có tổn thương	48	88
AI	Không có tổn thương	405	21
	Có tổn thương	83	128
Bác sỹ + AI	Không có tổn thương	444	45
	Có tổn thương	44	104

AI phát hiện được nhiều nốt/khối nhất. Bác sỹ với sự hỗ trợ của AI phát hiện được nhiều phim âm tính nhất. Se, Sp của bác sỹ, AI và bác sỹ với sự hỗ trợ của AI lần lượt là: 59,06% và 90,16%; 85,91% và 83%; 69,8% và 91%.



Biểu đồ 1. Đường cong ROC biểu thị giá trị của các mô hình chẩn đoán.

AI có Se cao hơn, Sp thấp hơn so với bác sỹ độc độc lập. Se, Sp và AUC đều tăng lên khi bác sỹ có thêm sự hỗ trợ của AI.

Bảng 3. So sánh giá trị chẩn đoán trên phim X-quang của các mô hình.

Tiêu chí	Bác sỹ	AI	Bác sỹ + AI	p
Se	59,06%* (55,24 - 62,88)	85,91%* (83,21 - 88,61)	69,80%* (66,23 - 73,37)	< 0,001
Sp	90,16% (87,85 - 92,47)	83%* (80,08 - 85,92)	91% (88,78 - 93,22)	< 0,001
AUC	0,746** (0,695 - 0,797)	0,884** (0,807 - 0,882)	0,804** (0,757 - 0,850)	0,008

(p: Sự khác biệt chung giữa 3 phương pháp chẩn đoán (Cochran's Q test);

*: Mô hình khác biệt có ý nghĩa với các mô hình còn lại (Mc Nemar test, $p < 0,05$);

** : Mô hình khác biệt có ý nghĩa với các mô hình còn lại (Delong method, $p < 0,05$))

Việc kết hợp với AI giúp tăng AUROC của bác sỹ từ 0,746 lên 0,804, đồng thời tăng Sp của AI từ 83% lên 91%, khác biệt có ý nghĩa thống kê.

Bảng 4. Đối chiếu những trường hợp bác sỹ sau khi tham khảo AI vẫn không đồng ý với AI (n = 113).

CLVT	AI		Tổng
	Có tổn thương	Không có tổn thương	
Có tổn thương	29	5	34
Không có tổn thương	59	20	79
Tổng	88	25	113

Trong số 113 trường hợp không có đồng thuận, bác sỹ đúng trong 64 trường hợp (56,64%) còn AI đúng trong 49 trường hợp (43,36%).

BÀN LUẬN

Trong nghiên cứu này, có 149 BN (23,39%) được xác nhận có tổn thương dạng nốt/khối trên chụp CLVT (*Bảng 1*). Tỷ lệ này cao hơn một số nghiên cứu tương tự [6, 8], tuy nhiên, điều này phụ thuộc vào những BN được lựa chọn và phương pháp xác định tiêu chuẩn đối chiếu. Trong nghiên cứu này, chúng tôi chọn ngẫu nhiên dữ liệu từ phim chụp tại phòng khám bệnh của bệnh viện và xây dựng tiêu chuẩn đối chiếu dựa trên CLVT nên tỷ lệ tổn thương sẽ cao hơn. Ngoài ra, không có sự khác biệt có ý nghĩa thống kê về sự xuất hiện của tổn thương ở trường phổi trái và phải.

Kích thước các nốt phổi trong nghiên cứu trung bình là 16mm, dao động từ 6 - 135mm (*Bảng 1*). Các nghiên cứu chỉ ra hình ảnh X-quang hầu như không thể phát hiện được các nốt phổi < 5mm; với nốt 6 - 10mm, Se khoảng 50% [3] nên các tổn thương phát hiện được sẽ tập trung nhiều trong khoảng ≥ 10 mm.

Trong nghiên cứu này, chúng tôi chỉ tập trung đánh giá những nốt tổn thương có kích thước ≥ 6 mm trên hình CLVT. Những tổn thương < 6mm được coi là không phát hiện được trên hình ảnh X-quang và bị loại khỏi nghiên cứu.

Trong nghiên cứu này, AI cho thấy Se cao hơn nhưng Sp thấp hơn so với bác sỹ đọc độc lập, sự khác biệt có ý nghĩa thống kê với $p < 0,05$. Kết quả này phù hợp với các nghiên cứu tương tự [8 - 10]. Tuy nhiên, về tổng thể, AUC của chẩn đoán AI cao hơn so với bác sỹ (0,884 so với 0,746, sự khác biệt có ý nghĩa thống kê với $p < 0,05$) và độ chính xác cao hơn (83,67% so với 82,89%). Khi so sánh giữa có hoặc không có sự hỗ trợ của AI, toàn bộ các tiêu chí về Se, Sp và AUC đều tăng lên khi bác sỹ có thêm sự hỗ trợ của AI (lần lượt là 59,06%; 90,16% và 0,746 so với 69,8%; 91% và 0,804). Điều này một lần nữa chứng minh giá trị của AI trong việc nâng cao hiệu quả trong chẩn đoán X-quang ngực.

Tuy nhiên, khi so sánh kết quả chẩn đoán giữa AI và bác sỹ có hỗ trợ bởi AI, điều bất ngờ là Se và AUC của kết quả do AI tạo ra lại cao hơn kết luận của bác sỹ sau khi tham khảo chính kết quả đó (lần lượt là 95,91% - 0,884 so với 69,8% - 0,804, *Bảng 3*), sự khác biệt có ý nghĩa thống kê với $p < 0,05$. Điều này liên quan tới vấn đề về quyết định của bác sỹ sau khi tham khảo gợi ý của AI.

Kết quả nghiên cứu của chúng tôi cũng tương tự một vài nghiên cứu khác. Tác giả Farouk S [4] nghiên cứu trên 150 BN chụp X-quang phổi và CLVT từ tháng 5/2021 - 7/2022 cho thấy có tất cả 216 nốt được phát hiện trên CLVT, 188 nốt phát hiện bởi AI, 170 nốt phát hiện bởi bác sỹ và 181 nốt được phát hiện sau khi bác sỹ tham khảo thêm AI, các nốt có kích thước 10mm (5 - 89mm). Đánh giá theo số lượng BN, hình ảnh CLVT phát hiện nốt trên 99 BN (66%), AI phát hiện được nốt ở 92 BN (61,3%). Việc kết hợp bác sỹ và AI giúp phát hiện 93 nốt ở BN (62%). Đối với đánh giá theo số BN có hoặc không có nốt, AUROC của AI là 0,890 (0,830 - 0,951), của kết hợp bác sỹ và AI là 0,866 (0,799 - 0,933) và của bác sỹ là 0,816 (0,740 - 0,893). So sánh hiệu quả chẩn đoán của AI với bác sỹ, của AI với bác sỹ kết hợp AI và bác sỹ độc lập với bác sỹ kết hợp AI, giá trị p lần lượt là 0,037; 0,437 và 0,023. Tác giả Bennani và CS [9] nghiên cứu trên 500 BN được chụp X-quang và CLVT trong vòng 72 giờ

cho thấy có 241 BN có ít nhất 1 tổn thương trên phim X-quang, tổng cộng có 522 tổn thương. Đối với tổn thương dạng nốt và khối ở trung thất, rốn phổi, AI giúp tăng Se lần lượt 12% và 5,9%, tăng Sp lần lượt 2,1% và 3,7%.

Bảng 4 trình bày kết quả của 113 BN mà bác sỹ không đồng ý với quan điểm của AI và vẫn giữ ý kiến mình. Không đồng ý với 88 BN được AI xác định có tổn thương, bác sỹ bỏ sót 29 BN có tổn thương thật nhưng cũng giúp tránh phát hiện nhầm 59 BN không có tổn thương. Không đồng ý với 25 BN được AI xác định không có tổn thương, bác sỹ giúp phát hiện thêm 5 BN có tổn thương nhưng cũng nhầm 20 BN không có tổn thương thật. Bằng việc không đồng ý với AI, bác sỹ đã chẩn đoán đúng 64 BN (5 có tổn thương và 59 không có tổn thương), bỏ sót 24 BN (21,2%), nhưng giúp tránh phát hiện nhầm 39 BN (34,5%). Kết quả này cho thấy tỷ lệ dương tính giả trong các gợi ý của AI tương đối cao, tạo ra nhiều cho các bác sỹ trong quá trình sử dụng. Chính vì vậy, để phát huy tối đa hiệu quả hỗ trợ của AI, bác sỹ cần phải hiểu rõ về ưu và nhược điểm của công cụ mình đang sử dụng để vừa tránh bỏ sót tổn thương, đồng thời không bị quá phụ thuộc vào gợi ý của AI.

KẾT LUẬN

Nghiên cứu của chúng tôi cho thấy phần mềm DrAid có Se và giá trị chẩn đoán cao trong việc phát hiện các tổn

thương dạng nốt/khối phổi trên phim X-quang ngực, mặc dù Sp còn tương đối thấp. Với sự trợ giúp của phần mềm DrAid, bác sỹ CDHA nâng cao hiệu quả chẩn đoán các nốt/khối phổi cả về Se, Sp và AUROC.

TÀI LIỆU THAM KHẢO

1. Kratzer TB, Bandi P, Freedman ND, Smith RA, Travis WD, Jemal A, et al. Lung cancer statistics, 2023. *Cancer*. 2024; 130(8):1330-1348.

2. Del Ciello A, Franchi P, Contegiacomo A, Cicchetti G, Bonomo L, Larici AR. Missed lung cancer: When, where, and why? *Diagnostic and interventional radiology (Ankara, Turkey)*. 2017; 23(2):118-126.

3. Diederich S, Das M. Solitary pulmonary nodule: Detection and management. *Cancer imaging: The official publication of the International Cancer Imaging Society*. 2006; 6(Spec No A):42-46.

4. Farouk S, Osman AM, Awadallah SM, Abdelrahman AS. The added value of using artificial intelligence in adult chest X-rays for nodules and masses detection in daily radiology practice. *Egyptian Journal of Radiology and Nuclear Medicine*. 2023; 54(1):142.

5. Miró Catalina Q, Vidal-Alaball J, Fuster-Casanovas A, Escalé-Besa A, Ruiz Comellas A, Solé-Casals J. Real-world testing of an artificial intelligence algorithm for the analysis of chest

X-rays in primary care settings. *Scientific Reports*. 2024; 14(1):5199.

6. Niehoff JH, Kalaitzidis J, Kroeger JR, Schoenbeck D, Borggreffe J, Michael AE. Evaluation of the clinical performance of an AI-based application for the automated analysis of chest X-rays. *Scientific Reports*. 2023; 13(1):3680.

7. Guo L, Zhou C, Xu J, Huang C, Yu Y, Lu G. Deep learning for chest X-ray diagnosis: Competition between radiologists with or without artificial intelligence assistance. *Journal of Imaging Informatics in Medicine*. 2024; 37(3):922-934.

8. Ahn JS, Ebrahimian S, McDermott S, Lee S, Naccarato L, Di Capua JF, et al. Association of artificial intelligence-aided chest radiograph interpretation with reader performance and efficiency. *JAMA Netw Open*. 2022; 5(8):e2229289.

9. Bennani S, Regnard N-E, Ventre J, Lassalle L, Nguyen T, Ducarouge A, et al. Using AI to improve radiologist performance in detection of abnormalities on chest radiographs. *Radiology*. 2023; 309(3):e230860.

10. Robert D, Sathyamurthy S, Singh AK, Matta SA, Tadeipalli M, Tanamala S, et al. Effect of artificial intelligence as a second reader on the lung nodule detection and localization accuracy of radiologists and non-radiology physicians in chest radiographs: A multicenter reader study. *Academic Radiology*. 2025; 32(3):1706-1717.