

không tiếp xúc trực tiếp với nguồn nước, nguy cơ vẫn tồn tại do vắt có thể bám vào kết mạc trong quá trình đi rừng.

Về xử trí, mặc dù nhiều tác giả khuyến cáo dùng dung dịch ưu trương hoặc nhỏ thuốc tê trực tiếp lên vắt để làm chúng co rút, trong ca bệnh của chúng tôi, vắt vẫn bám rất chặt dù đã nhỏ Tetracain nhiều lần. Điều này gợi ý rằng mức độ đáp ứng có thể khác nhau tùy loài vắt hoặc tình trạng sinh lý của ký sinh trùng. Do đó, việc chuẩn hóa phác đồ điều trị vẫn còn là thách thức.

Ngoài nguy cơ xuất huyết và nhiễm trùng cấp tính, cần lưu ý các biến chứng muộn như sẹo giác mạc, viêm mạn tính hoặc thậm chí giảm thị lực. Do đó, việc theo dõi sau thủ thuật là cần thiết, ngay cả khi bệnh nhân ổn định sau 3 ngày.

Vắt hút nhãn cầu rất hiếm gặp. Lực hút rất chặt của vắt vào ký chủ khiến quá trình lấy vắt rất khó khăn, đưa đến nguy cơ sót dị vật là nguồn nhiễm trùng có thể gây nên các biến chứng nghiêm trọng. Hiện nay, vẫn chưa có khuyến cáo điều trị chính thức cho các trường hợp vắt hút nhãn cầu, do đó, việc bảo hộ đúng cách khi đi rừng là cực kỳ quan trọng.

IV. KẾT LUẬN

Trường hợp này góp phần nhấn mạnh tầm quan trọng của giáo dục cộng đồng về bảo vệ mắt khi tham gia hoạt động tại rừng núi. Việc sử dụng kính bảo hộ nên được khuyến khích như một biện pháp phòng ngừa đơn giản nhưng hiệu quả. Đây cũng là báo cáo có ý nghĩa học thuật

khi bổ sung dữ liệu hiếm gặp từ Việt Nam vào y văn quốc tế về bệnh cảnh này.

TÀI LIỆU THAM KHẢO

1. **Ralp Buchsbaum MB, John Pearse,** Vicki Pearse Animals Without Backbone Chicago: The University of Chicago Press 1987:312 - 317
2. **Sawyer R.** Neurobiology of the Leech New York: Cold Spring Harbor Laboratory 1981:7-26
3. **Daryabari SH, Hassanpour SR, Morshedi M.** Complications of leech therapy directly on the eyes. Oman J Ophthalmol. Sep-Dec 2021;14(3): 205-206. doi:10.4103/ojo.ojo_129_21
4. **Rubi Dey DO SDM.** Ocular Lecch Infestation Archives of Iranian Medicine September 2010;13 (5)
5. **Yoshikazu Ito TN, Mutsuko Ohara.** Ocular infestation by a juvenile leech, Myxobdella sinanensis in Japan. American journal of Ophthalmology Case Reports March 2022 25
6. **Partyka C, Fogg T.** Leech on the eye: a novel use for hypertonic saline. Emerg Med Australas. Feb 2009;21(1): 84-5. doi:10.1111/j.1742-6723.2009.01156.x
7. **Alcelik T, Cekic O, Totan Y.** Ocular leech infestation in a child. Am J Ophthalmol. Jul 1997;124(1): 110-2. doi:10.1016/s0002-9394(14) 71655-1
8. **Lewis G, Coombes A.** Adult ocular leech infestation. Eye (Lond). Mar 2006;20(3):391-2. doi:10.1038/sj.eye.6701862
9. **Auw-Haedrich C, Keim A, Kist M.** Conjunctival infestation of a child with Theromyzon tessulatum. Br J Ophthalmol. Sep 1998;82(9): 1093-4. doi:10.1136/bjo.82.9.e1090
10. **Li WW, Shen T, Jiang J.** Ocular leech infestation initially misdiagnosed as conjunctival pigmented nevus. Int J Ophthalmol. 2013;6(4): 557-8. doi:10.3980/j.issn.2222-3959.2013.04.28

CƠ SỞ DỮ LIỆU GEN VÀ GIẢI TRÌNH TỰ THỂ HỆ MỚI PHÁT HIỆN ĐỘT BIẾN ĐÁP ỨNG THUỐC UNG THƯ PHỔI TẠI VIỆT NAM

Vũ Thị Huyền¹, Nguyễn Thị Trang¹, Hoàng Công Minh¹,
Lê Tú Linh^{1,2}, Vũ Thị Lan Anh³, Lê Thị Minh Phương^{1,4}

TÓM TẮT

Mục tiêu: Xây dựng cơ sở dữ liệu gen ung thư phổi phù hợp cho bệnh nhân Việt Nam, ứng dụng trí tuệ nhân tạo (AI) và giải trình tự gen thế hệ mới (NGS) nhằm phát hiện các biến thể liên quan đến điều

trị. **Phương pháp nghiên cứu:** Nghiên cứu mô tả cắt ngang trên dữ liệu Y văn PubMed và 21 bệnh nhân ung thư phổi không tế bào nhỏ giai đoạn tiến triển (2024-2025). Kỹ thuật NGS được áp dụng trên mẫu mô ung thư, đồng thời sử dụng AI và xử lý ngôn ngữ tự nhiên (NLP) để sàng lọc dữ liệu từ các cơ sở quốc tế (ClinVar, COSMIC, OncoKB...). **Kết quả:** Hệ thống đã sàng lọc >400.000 tài liệu, trích xuất >9 triệu thực thể và xác định 97 gen liên quan đến sinh bệnh, tiên lượng và đáp ứng điều trị, trong đó có EGFR, KRAS, ALK, BRAF, TP53. Ở 21 bệnh nhân được giải trình tự, phát hiện 663 biến thể; 12 biến thể ở 11 bệnh nhân được xác định là gây bệnh và có thể gây bệnh liên quan điều trị đích, phổ biến nhất là EGFR del19 (66,7%) và EGFR L858R (16,7%). Ngoài ra, ghi nhận 1 trường hợp đồng biến thể del19 và T790M liên quan kháng thuốc TKI. Tần số alen biến thể (VAF) dao động

¹Trường Đại học Y Hà Nội

²Bệnh viện Phổi Trung Ương

³Bệnh viện Đại học Y Hà Nội

⁴Trường Đại học Y Dược, Đại học Quốc Gia Hà Nội

Chịu trách nhiệm chính: Lê Thị Minh Phương

Email: minhphuong88.hmu@gmail.com

Ngày nhận bài: 17.9.2025

Ngày phản biện khoa học: 22.10.2025

Ngày duyệt bài: 26.11.2025

5-66%, phản ánh tính dị hợp của khối u. **Kết luận:** Ứng dụng AI và NGS cho thấy hiệu quả trong khai thác dữ liệu gen ung thư phổi, bước đầu tạo nền tảng cho cơ sở dữ liệu đặc thù người Việt.

Từ khóa: Ung thư phổi, trí tuệ nhân tạo, giải trình tự gen thế hệ mới, EGFR, cơ sở dữ liệu gen.

SUMMARY

APPLICATION OF GENETIC DATABASES AND NGS TO IDENTIFY DRUG-RESPONSIVE MUTATIONS IN LUNG CANCER IN VIETNAM

Objective: To establish a lung cancer genetic database tailored for Vietnamese patients by applying artificial intelligence (AI) and next-generation sequencing (NGS) for the detection of therapy-related variants. **Methods:** A cross-sectional descriptive study was conducted using PubMed literature data and samples from 21 patients with advanced non-small cell lung cancer (2024–2025). NGS was performed on tumor tissue samples, while AI and natural language processing (NLP) were employed to filter and extract information from international databases (ClinVar, COSMIC, OncoKB, etc.). **Results:** The system screened over 400,000 publications, extracted more than 9 million entities, and identified 97 genes associated with pathogenesis, prognosis, and therapeutic response, including EGFR, KRAS, ALK, BRAF, and TP53. In 21 sequenced patients, 663 variants were detected; 12 variants in 11 patients were classified as pathogenic or likely pathogenic with therapeutic relevance. The most frequent were EGFR del19 (66.7%) and EGFR L858R (16.7%). Additionally, one case carried concurrent del19 and T790M variants, associated with TKI resistance. Variant allele frequencies (VAF) ranged from 5-66%, reflecting tumor heterogeneity. **Conclusion:** The integration of AI and NGS proved effective in mining genetic data for lung cancer and provided an initial foundation for a population-specific database in Vietnamese patients.

Keywords: Lung cancer, artificial intelligence, next-generation sequencing, EGFR, genetic database.

I. ĐẶT VẤN ĐỀ

Ung thư phổi là nguyên nhân gây tử vong hàng đầu trên toàn cầu. Theo GLOBOCAN 2022, tại Việt Nam ghi nhận khoảng 24.000 ca mắc mới và 22.000 ca tử vong mỗi năm, lần lượt đứng thứ 2 và 3 trong các loại ung thư.¹ Mặc dù đã có nhiều tiến bộ, tiên lượng vẫn xấu, đặc biệt ở giai đoạn muộn, do đó phát hiện sớm và điều trị cá thể hóa dựa trên dữ liệu di truyền là chìa khóa cải thiện kết quả. Gần đây, giải trình tự gen thế hệ mới (NGS) cho phép phát hiện hàng loạt biến thể gen liên quan đến bệnh sinh, tiên lượng và đáp ứng điều trị. Các cơ sở dữ liệu như PubMed, ClinVar, COSMIC, OncoKB, CIVIC... cung cấp nguồn thông tin phong phú nhưng còn hạn chế về tính thống nhất, chi phí và chưa phản ánh đặc điểm người Việt Nam. Xuất phát từ thực tế này, chúng tôi xây dựng cơ sở dữ liệu gen ung

thư phù hợp với Việt Nam, ứng dụng trí tuệ nhân tạo (AI) và xử lý ngôn ngữ tự nhiên (NLP) để sàng lọc, phân loại, chiết xuất dữ liệu, đồng thời tiến hành giải trình tự NGS ở bệnh nhân ung thư phổi nhằm phục vụ chẩn đoán và lựa chọn liệu pháp điều trị.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu. Các y văn được công bố chính thức trên kho dữ liệu Y sinh và 21 người bệnh được chẩn đoán ung thư phổi được giải trình tự gen thế hệ mới.

2.2. Phương pháp nghiên cứu. Nghiên cứu mô tả cắt ngang

* **Cỡ mẫu:** 21 bệnh nhân được chẩn đoán ung thư phổi theo tiêu chuẩn lựa chọn: Đã được chẩn đoán ung thư phổi không tế bào nhỏ (NSCLC) giai đoạn tiến triển theo "Hướng dẫn chẩn đoán và điều trị ung thư phổi không tế bào nhỏ" của Bộ Y Tế năm 2018. - Đã có kết quả xét nghiệm đột biến gen ung thư phổi bằng các kỹ thuật sinh học phân tử. - Có mẫu mô ung thư được lưu trữ tại bệnh viện. - Có đầy đủ thông tin bệnh sử, tiền sử, các triệu chứng lâm sàng và cận lâm sàng theo mẫu bệnh án thống nhất.

* **Thời gian nghiên cứu:** từ năm 2024-2025

* **Địa điểm:** Trung tâm Di truyền Lâm sàng - Bệnh viện Đại học Y Hà Nội, Bộ môn Y Sinh học Di truyền trường Đại học Y Hà Nội, Viện Công nghệ ADN và phân tích di truyền.

Quy trình AI xác định tài liệu liên quan

Xây dựng module thu thập, phân tích và lọc các thực thể được đề cập đến trong các bài báo khoa học có liên quan đến ung thư phổi và gen, gồm các thành phần sau:

- Thu thập dữ liệu: Data crawler tương tác, truy vấn và lấy dữ liệu từ các bài báo trên PubMed.

- Nhận dạng thực thể (Named Entity Recognition- NER): Trích xuất các thực thể liên quan đến lĩnh vực y sinh trong ung thư phổi, đào tạo trên các mô hình học sâu PubmedBert-CRF.

- Trích xuất mối quan hệ (Relation Extraction- RE): Xác định và trích xuất các mối quan hệ giữa các thực thể y sinh trong văn bản liên quan đến ung thư, bao gồm quan hệ giữa gen, biến thể, ung thư và thuốc điều trị.

- Trích xuất chủng tộc và quốc gia: kỹ thuật NLP (natural language processing) phân tích và nhận dạng các đặc điểm liên quan đến, tập trung vào việc xác định các chủng tộc và quốc gia của Châu Á và Việt Nam.

- Ánh xạ và chuẩn hóa dữ liệu (Data Mapping and Normalization- DMN): Chuẩn hóa dữ liệu từ các bài báo khác nhau thành một định dạng thống nhất.

- Quản lý cơ sở dữ liệu: sử dụng MySQL.

Kỹ thuật NGS

- Mẫu sinh thiết mô được thu thập từ người bệnh ung thư phổi
- Tách chiết DNA bằng QIAamp DNA Mini Kit (QIAGEN).
 - DNA sau tách chiết được kiểm tra chất lượng và định lượng bằng Qubit Fluorometer 4.
 - Chuẩn bị thư viện bằng BGI DNA Library Prep Kit, bao gồm các bước cắt đoạn DNA, gắn adapter, làm giàu và khuếch đại trên panel probe dựa trên các biến thể gen đã ghi nhận từ module AI.
 - Thực hiện giải trình tự trên hệ thống MGISEQ-2000.
 - Dữ liệu NGS thô được xử lý bằng BGI Bioinformatics Pipeline, bao gồm lọc read chất lượng thấp, căn chỉnh với bộ gen tham chiếu GRCh38, phân tích và lọc biến thể bằng GATK, cuối cùng chú thích trên SnpEff nhằm chọn các biến thể liên quan ung thư phổi
 - Kết quả được so sánh với các cơ sở dữ liệu ClinVar, COSMIC và dbSNP để xác định các biến

thể có ý nghĩa.

2.3. Phân tích số liệu. Số liệu được thu thập và quản lý bằng phần mềm Excel, phân tích bằng SPSS 27.

2.4. Đạo đức nghiên cứu. Nghiên cứu là sản phẩm của đề tài cơ sở “Ứng dụng công nghệ giải trình tự gen thế hệ mới trong phát hiện đột biến gen liên quan đến sự đáp ứng của thuốc điều trị ung thư phổi tại Việt Nam” thuộc đề tài cấp Nhà nước KC4.0, mã số KC-4.0-45/19-25; đề tài được thông qua hồ sơ đạo đức số 912/GCN-HĐĐĐNCYSH-ĐHYHN.

III. KẾT QUẢ NGHIÊN CỨU

3.1. Cơ sở dữ liệu các gen liên quan đến ung thư phổi tại Việt Nam. Qua thu thập xử lý và nhận dạng các thực thể thu được hơn 400.000 tài liệu và hơn 9.000.000 thực thể xuất hiện trong các tài liệu, kết quả xác định được 6944 tài liệu có liên quan đến ung thư phổi, trong đó ghi nhận 97 gen có liên quan đến cơ chế bệnh sinh, triệu chứng hoặc đáp ứng điều trị của ung thư phổi.

Bảng 1. Một số gen trong cơ sở dữ liệu gen liên quan tới ung thư phổi đã thiết lập

STT	Tên gen	Vị trí	Chức năng
1	AKT1	14q32.33	Hoạt hóa qua PI3K, thúc đẩy tăng sinh tế bào, chống apoptosis
2	ALK	2p23.2	Mã hóa receptor tyrosine kinase, tăng sống sót, điều hòa biệt hóa và tăng trưởng tế bào
3	BCL2	18q21.33	Kéo dài sự sống của tế bào, ngay cả khi có tín hiệu stress hoặc tổn thương DNA.
5	BRAF	7q34	Mã hóa serine/threonine kinase thuộc họ RAF, vai trò trung tâm trong con đường RAS/RAF/MEK/ERK, điều hòa phiên mã gen liên quan đến tăng trưởng tế bào, đột biến gây tăng sinh tế bào mất kiểm soát
6	BRCA1	17q21	Trực tiếp tham gia sửa chữa DNA đứt gãy kép bằng tái tổ hợp tương đồng, đột biến làm tế bào mất khả năng sửa chữa chính xác tổn thương DNA, tăng nguy cơ hình thành ung thư
7	CCND1	11q13.3	Thành phần của hệ thống kiểm soát chu kỳ tế bào
8	EGFR	7p11.2	Thụ thể yếu tố tăng trưởng biểu mô, vai trò chính trong điều hòa biệt hóa và tăng sinh tế bào
9	ERBB2	17q12	Kích hoạt các con đường tín hiệu MAPK, PI3K/AKT, dẫn đến tăng sinh tế bào, biệt hóa và sống sót
10	ERCC1	19q13.32	Tham gia quá trình sửa chữa cắt bỏ nucleotide (NER)
11	FGF1	5q31.3	Yếu tố tăng trưởng họ FGF, điều hòa tăng sinh, biệt hóa và di chuyển tế bào
12	FLT1	13q12	Mã hóa thụ thể tyrosine kinase của yếu tố tăng trưởng nội mô mạch máu, điều hòa tân sinh mạch
13	HRAS	11p15.5	Mã hóa GTPase thuộc họ RAS, vai trò công tắc trong điều hòa tăng trưởng và biệt hóa tế bào
14	JAK2	9p24.1	Vai trò trong dẫn truyền tín hiệu của nhiều cytokine và hormone tăng trưởng thông qua con đường JAK/STAT
15	KRAS	12p12.1	Mã hóa GTPase thuộc họ RAS, vai trò công tắc trong điều hòa tăng trưởng và biệt hóa tế bào
16	MAP2K1	15q22.31	Điều hòa biểu hiện gen, tăng trưởng và biệt hóa tế bào thông qua con đường RAS/RAF/MEK/ERK
17	MET	7q31	Hoạt hóa các con đường điều hòa tăng trưởng tế bào qua gắn với HGF
18	MTOR	1p36.13	Vai trò trung tâm của con đường mTOR, điều hòa dịch mã, tăng trưởng tế bào

19	PIK3CA	3q26.32	Vai trò trung tâm trong con đường tín hiệu PI3K/AKT/mTOR, điều hòa trao đổi chất và tăng trưởng tế bào
20	RAD50	5q31	Tham gia sửa chữa DNA đứt kép, tương tác với BRCA1 và các protein sửa chữa khác trong giám sát và duy trì ổn định hệ gen
21	RET	10q11.21	Điều hòa tăng trưởng, biệt hóa và sống còn của tế bào thông qua các con đường tín hiệu RAS/MAPK và PI3K/AKT
22	ROS1	6q22.1	Điều hòa tăng trưởng, biệt hóa và sống sót của tế bào thông qua hoạt hóa các con đường PI3K/AKT, JAK/STAT và MAPK/ERK
23	TP53	17p13.1	Kiểm soát chu kì tế bào, ngừng tăng sinh tế bào khi phát hiện tổn thương DNA, điều hòa apoptosis, hoạt hóa các con đường sửa chữa DNA

Bảng 2. Một số biến thể trong cơ sở dữ liệu gen đã thiết lập

STT	Gen	Biến thể	Điều trị	Đáp ứng/Kháng
1	ALK	L1196M	Crizotinib	Kháng
		C1156F	Crizotinib	Kháng
		G1269A	Ceritinib	Kháng
		V1180L	Alectinib	Kháng
2	ARAF	S214C	Sorafenib	Đáp ứng
3	BRAF	V600E	Osimertinib	Kháng
		V600E	Dabrafenib + Trametinib	Đáp ứng
4	EGFR	L858R	Osimertinib	Đáp ứng
		T790M	Osimertinib	Đáp ứng
		G719X	Gefitinib	Đáp ứng
		L861Q	Osimertinib	Đáp ứng
		C797S	Osimertinib	Kháng
5	DDR3	S768R	Erlotinib	Đáp ứng
6	FGFR3	S249C	Fexagratinib	Đáp ứng
7	KRAS	G12C	Selumetinib	Đáp ứng
		G12C	Gefitinib	Kháng
		G12V	Selumetinib	Đáp ứng
		G12V	Gefitinib	Kháng
8	KIT	D816G	Crizotinib	Kháng
9	MET	D1246N	Capmatinib	Kháng
		D1246V	Savolitinib	Kháng
		Y1248H	Capmatinib	Kháng
10	PIK3CA	E454K	Gefitinib	Kháng
11	POLD1	C284Y	Pembrolizumab	Đáp ứng
12	SMO	P641A	Vismodegib	Đáp ứng

3.2. Kết quả giải trình tự NGS ở người bệnh ung thư phổi không tế bào nhỏ. Đã thực hiện giải trình tự gen thế hệ mới cho 21 bệnh nhân được chẩn đoán xác định ung thư phổi không tế bào nhỏ, kết quả ghi nhận 12 biến thể của 11 bệnh nhân. Các biến thể này đều nằm trên gen EGFR thuộc các gen nằm trong CSDL đã thiết lập liên quan đến điều trị đích với tỷ lệ và tần xuất như ở bảng sau:

Biến thể	Số lượng (n)	Tỷ lệ (%)
Phát hiện biến thể	11	52,38
EGFR del 19	8	66,7
EGFR L858R	2	16,7
EGFR insC	1	8,3
EGFR T790M	1	8,3
Không phát hiện biến thể	10	47,62

Nhận xét: biến thể gen EGFR del19 và EGFR L858R là 2 biến thể phổ biến nhất, chiếm 83,4% tổng số biến thể phát hiện có ý nghĩa lâm sàng. Biến thể EGFR insC và EGFR T790M ít gặp hơn (8,3%). Có 1 bệnh nhân mang đồng thời 2 biến thể EGFR del19 và EGFR T790M. 10 bệnh nhân chưa phát hiện biến thể có ý nghĩa lâm sàng.

Tần xuất các biến thể (VAF):

	VAF (%)
Nhỏ nhất	5
Lớn nhất	66
Trung bình	22,58±16,92

Nhận xét: Tần xuất xuất hiện các biến thể trong từng mẫu có sự giao động lớn từ 5 - 66%, trung bình 22,58±16,92%.

IV. BÀN LUẬN

Nghiên cứu của chúng tôi cho thấy việc ứng

dụng trí tuệ nhân tạo (AI) kết hợp với công nghệ xử lý ngôn ngữ tự nhiên (NLP) là một hướng đi tiềm năng và mang lại nhiều giá trị trong việc xây dựng cơ sở dữ liệu gen liên quan đến ung thư phổi. Trong nhiều thập kỷ qua, việc khai thác dữ liệu y sinh thường phụ thuộc chủ yếu vào các phương pháp tìm kiếm thủ công và phân tích truyền thống. Cách tiếp cận này không chỉ tốn nhiều thời gian, nguồn lực nhân lực mà còn dễ dẫn đến nguy cơ bỏ sót những thông tin quan trọng, đặc biệt trong bối cảnh số lượng công bố khoa học ngày càng tăng nhanh. Thực tế cho thấy, mỗi năm có hàng chục nghìn bài báo mới liên quan đến ung thư phổi và các dữ liệu di truyền học được cập nhật liên tục trên các cơ sở dữ liệu quốc tế. Điều này đặt ra thách thức lớn đối với các nhà nghiên cứu và bác sĩ lâm sàng trong việc chọn lọc, hệ thống hóa và ứng dụng thông tin một cách hiệu quả.

Trong bối cảnh đó, việc áp dụng AI và NLP có thể xem là một giải pháp đột phá. Công nghệ này giúp tăng tốc độ xử lý, đảm bảo độ chính xác và đặc biệt là nâng cao tính hệ thống trong tổng hợp dữ liệu. Không chỉ dừng lại ở việc rút trích thông tin từ các bài báo khoa học, AI còn có khả năng phân loại, nhận diện mối liên quan giữa các biến thể gen và cơ chế bệnh sinh, đồng thời gợi ý các liên kết với tiên lượng cũng như đáp ứng điều trị.

Trong bối cảnh khối lượng dữ liệu khoa học ngày càng gia tăng, đặc biệt là trên các cơ sở dữ liệu quốc tế như PubMed, ClinVar, COSMIC, OncoKB, việc ứng dụng AI giúp tăng tốc độ, độ chính xác và tính hệ thống của việc tổng hợp thông tin.^{2,3} Kết quả nghiên cứu đã cho phép sàng lọc hơn 400.000 tài liệu và trích xuất hơn 9 triệu thực thể, từ đó xác định được 97 gen có liên quan đến cơ chế bệnh sinh, tiên lượng và đáp ứng điều trị của ung thư phổi. Đáng chú ý, nhiều gen quan trọng được hệ thống nhận diện, bao gồm EGFR, KRAS, ALK, BRAF, TP53.... Các gen này vốn đã được ghi nhận trong nhiều nghiên cứu trước đây là đóng vai trò trung tâm trong sinh ung thư phổi, đồng thời là đích tác động của các liệu pháp điều trị nhắm trúng đích hiện nay. Ví dụ, các biến thể trên gen EGFR như del19 hay L858R có liên quan mật thiết đến hiệu quả điều trị bằng thuốc ức chế tyrosine kinase (TKI), trong khi đột biến T790M lại là nguyên nhân phổ biến dẫn đến kháng thuốc. Sự xác nhận lại các gen và biến thể này trong nghiên cứu không chỉ củng cố giá trị của hệ thống AI-NLP mà còn chứng minh tính khả thi của việc ứng dụng công nghệ vào thực hành lâm sàng.

Qua thực hiện giải trình tự thế hệ mới trên

mẫu mô của 21 bệnh nhân ung thư phổi, chúng tôi ghi nhận tổng cộng 663 biến thể thuộc các gen nằm trong cơ sở dữ liệu đã xây dựng. Qua quá trình lọc và phân loại biến thể, phát hiện 12 biến thể của 11 bệnh nhân liên quan đến điều trị đích được xác định là gây bệnh hoặc có thể gây bệnh phù hợp với lâm sàng của bệnh nhân. Các biến thể phổ biến nhất bao gồm EGFR del 19 (chiếm 66,7%) và EGFR L858R (chiếm 16,7%). Biến thể EGFR del19 gây ra bởi tình trạng loại bỏ một phần exon thứ 19 của gen EGFR trong khi EGFR L858R gây ra bởi một đột biến điểm trong miền tyrosine kinase của protein EGFR. Kết quả của các biến đổi này đều dẫn đến protein EGFR hoạt động quá mức, gây ra sự phát triển không kiểm soát của tế bào ung thư.

Kết quả này tương đồng với nghiên cứu của Marc Ladanyi và cộng sự năm 2008 hoặc Nguyễn Thị Thủy và cộng sự năm 2024 cho thấy EGFR del19 và L858R là 2 biến thể phổ biến nhất ở bệnh nhân ung thư phổi không tế bào nhỏ, chiếm khoảng 90% tổng số các biến thể di truyền có liên quan đến độ nhạy cảm với EGFR-TKI.^{4,5}

Bên cạnh đó, nghiên cứu cũng phát hiện 1 trường hợp mang đồng thời 2 biến thể trên gen EGFR bao gồm del19 và T790M. Nếu như EGFR del19 dự báo bệnh nhân nhạy cảm với thuốc điều trị đích TKI thì ngược lại, EGFR T790M lại gây ra tình trạng kháng thuốc TKI thế hệ I.

Tần số alen biến thể (VAF) đo tỷ lệ phần trăm trình tự DNA trong một mẫu mang biến thể di truyền cụ thể tại một vị trí cụ thể. Nó được tính bằng số lần đọc có biến thể chia cho tổng số lần đọc tại vị trí đó. VAF cung cấp thông tin chi tiết về tính không đồng nhất của khối u, giúp xác định các đột biến điều khiển và có thể đánh giá nguồn gốc của biến thể, phân biệt giữa các kiểu gen dòng mầm, thể soma và khảm. Trong nghiên cứu của chúng tôi, VAF có sự biến thiên khá lớn từ 5-66% tùy từng mẫu. Những mẫu có VAF cao thường là đột biến xảy ra ở giai đoạn đầu phát triển của khối u và xuất hiện ở tỷ lệ lớn các tế bào ung thư và có nhiều khả năng là mục tiêu tốt cho liệu pháp điều trị.

Một hạn chế quan trọng của nghiên cứu là số lượng mẫu giải trình tự còn nhỏ, chưa đủ để phản ánh đầy đủ sự đa dạng di truyền trong cộng đồng. Ngoài ra, việc chưa xác định được biến thể gây bệnh rõ ràng cho thấy cần mở rộng cỡ mẫu, kết hợp thêm các nghiên cứu chức năng (functional study) cũng như phát triển cơ sở dữ liệu gen đặc thù cho người Việt Nam. Những yếu tố này sẽ tạo nền tảng cho cá thể hóa điều trị, góp phần cải thiện hiệu quả điều trị và tiên lượng bệnh nhân ung thư phổi.

V. KẾT LUẬN

Nghiên cứu của chúng tôi đã bước đầu cho thấy tiềm năng ứng dụng AI trong khai thác dữ liệu y sinh và sự cần thiết của việc xây dựng cơ sở dữ liệu gen riêng cho bệnh nhân Việt Nam. Trong tương lai, việc kết hợp AI, NGS và các nghiên cứu đa trung tâm với cỡ mẫu lớn hơn được kỳ vọng sẽ mở ra hướng đi mới trong chẩn đoán, điều trị và quản lý ung thư phổi theo hướng y học chính xác và cá thể hóa.

TÀI LIỆU THAM KHẢO

1. **Bray F, Laversanne M, Sung H, et al.** Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. n/a(n/a). doi:10.3322/caac.21834
2. **Alsharif F.** Artificial Intelligence in Oncology: Applications, Challenges and Future Frontiers. *Int J Pharm Investigation*. 2024;14(3):647-656. doi:10.5530/ijpi.14.3.76
3. **Lee JH, Hwang EJ, Kim H, Park CM.** A narrative review of deep learning applications in lung cancer research: from screening to prognostication. *Transl Lung Cancer Res*. 2022;11(6):1217-1229. doi:10.21037/tlcr-21-1012
4. **Shi Y, Li J, Zhang S, et al.** Molecular Epidemiology of EGFR Mutations in Asian Patients with Advanced Non-Small-Cell Lung Cancer of Adenocarcinoma Histology - Mainland China Subset Analysis of the PIONEER study. *PLoS One*. 2015;10(11):e0143515. doi:10.1371/journal.pone.0143515
5. **Shi H, Seegobin K, Heng F, et al.** Genomic landscape of lung adenocarcinomas in different races. *Front Oncol*. 2022;12:946625. doi:10.3389/fonc.2022.946625

KHẢO SÁT TÌNH TRẠNG TIÊM VẮC XIN HPV CỦA SINH VIÊN NỮ NGÀNH PHỤC HỒI CHỨC NĂNG TRƯỜNG ĐẠI HỌC KỸ THUẬT Y TẾ HẢI DƯƠNG

Lê Văn Thâm¹, Lê Thúy Hương¹

TÓM TẮT

Mục tiêu: Mô tả tình trạng tiêm vắc xin HPV của sinh viên nữ ngành Phục hồi chức năng trường Đại học kỹ thuật y tế Hải Dương. **Phương pháp:** Mô tả cắt ngang. **Kết quả nghiên cứu:** Tỷ lệ sinh viên nữ ngành Phục hồi chức năng trường Đại học kỹ thuật y tế Hải Dương tiêm vắc xin HPV còn chưa cao (37,1%), trong đó chỉ có 16,1% tiêm đủ 3 mũi; Lý do sinh viên tiêm vắc xin HPV nhiều nhất là tư nhân thức được tầm quan trọng của việc tiêm vaccine (56,9%), Tiếp đó là các sinh viên được gia đình khuyến khích, đưa đi tiêm (32,6%); Lý do sinh viên chưa tiêm vắc xin HPV do giá thành cao (53,5%), không biết chỗ nào để tiêm (12,5%), không có thời gian nên không tiêm (11,8%), không biết về vaccine HPV (8,3%) và lo ngại về tác dụng phụ của vaccine (5,5%). **Kết luận:** Tỷ lệ sinh viên nữ ngành Phục hồi chức năng trường Đại học kỹ thuật y tế Hải Dương tiêm vắc xin HPV còn chưa cao (37,1%), trong đó chỉ có 16,1% tiêm đủ 3 mũi.

SUMMARY

ASSESSMENT OF HPV VACCINATION STATUS AMONG FEMALE STUDENTS IN THE REHABILITATION FACULTY AT HAI DUONG MEDICAL TECHNICAL UNIVERSITY

Objective: To assess the HPV vaccination status among female students in the Faculty of Rehabilitation

at Hai Duong Medical Technical University. **Methods:** A cross-sectional study was conducted. **Results:** The HPV vaccination rate among female rehabilitation students remained relatively low at 37.1%, with only 16.1% having completed the full three-dose schedule. The most common reason for receiving the vaccine was personal awareness of its importance (56.9%), followed by family encouragement and support (32.6%). Among those unvaccinated, the primary barriers were high vaccine cost (53.5%), lack of information about vaccination sites (12.5%), limited time (11.8%), lack of awareness about the HPV vaccine (8.3%), and concerns about potential side effects (5.5%). **Conclusion:** The overall HPV vaccination coverage among female rehabilitation students at Hai Duong Medical Technical University remains suboptimal (37.1%), with a particularly low proportion completing the recommended three-dose schedule (16.1%).

I. ĐẶT VẤN ĐỀ

Tổ chức y tế thế giới (WHO) ước tính tỉ lệ nhiễm HPV trên toàn thế giới ở phụ nữ tính đến năm 2017 là 11,7%. Theo thống kê hơn 95% trường hợp ung thư cổ tử cung là do HPV lây truyền qua đường tình dục. Đây là loại ung thư phổ biến thứ hai ở phụ nữ trên toàn cầu, với ước tính có khoảng 604.000 ca mắc mới và 342.000 ca tử vong vào năm 2020 [1]. Khoảng 90% ca mắc mới và tử vong trên toàn thế giới vào năm 2020 xảy ra ở các quốc gia có thu nhập thấp và trung bình [2].

Năm 2012, ước tính tại Việt Nam có 5.146 phụ nữ mắc ung thư cổ tử cung, tỷ lệ mắc mới

¹Trường Đại học Kỹ thuật Y tế Hải Dương

Chịu trách nhiệm chính: Lê Văn Thâm

Email: themlv2003@gmail.com

Ngày nhận bài: 22.9.2025

Ngày phản biện khoa học: 23.10.2025

Ngày duyệt bài: 26.11.2025