BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC PHAN CHÂU TRINH

ĐỀ TÀI
NGHIÊN CỨU KHOA HỌC

TÊN ĐỀ TÀI:

NGHIÊN CỨU SỰ HÀI LÒNG CỦA BỆNH NHÂN ĐỐI VỚI CHATGPT VÀ BÁC SĨ LÂM SÀNG TRONG TƯ VẤN SỨC KHỎE TẠI CÁC BỆNH VIỆN TÂM TRÍ NĂM 2024

LĨNH VỰC CỦA ĐỀ TÀI: **Công nghệ thông tin trong Y học**

CHỦ NHIỆM ĐỀ TÀI: **BS.CKII LÊ XUÂN TÚY**

THÀNH VIÊN NHÓM NGHIÊN CỨU:

1. **CN NGUYỄN NGỌC THẢO**
2. **BS VÕ THANH LONG**
3. **CN NGUYỄN THỊ HỒNG GIAN**
4. **ThS. BS LÊ THỊ MINH TÂM**

QUẢNG NAM – 2024

DANH MỤC CHỮ VIẾT TẮT

AI	Trí tuệ nhân tạo (Artifical Intelligence)
AMBOSS	Thư viện điện tử
COVID 19	Corona virus 2019
CGYT	Chuyên gia y tế
Chat GPT	Chat Generative Pre-training Transformer
EQIP	Công cụ đảm bảo thông tin chất lượng cho bệnh nhân
InstructGPT	Mô hình ngôn ngữ kiểu GPT được tinh chỉnh bằng phản hồi của con người
NBME	Hội đồng giám định y tế quốc gia (National Board of Medical Examiner)
OpenAI	Phòng thí nghiệm nghiên cứu về trí tuệ nhân tạo
RLHF	Học tăng cường từ phản hồi của con người
SL	Số lượng
T2DM	Đái tháo đường típ 2
TGAI	Văn bản do AI tạo ra
USMLE	Kỳ thi cấp phép y tế Hoa Kỳ (United States Medical Licensing Examination)
WHO	Tổ chức Y tế Thế giới (World Health Organization)

MỤC LỤC

ĐẶT VẤN ĐỀ……..…….1
Chương 1: TỔNG QUAN TÀI LIỆU.....2
1.1. Các khái niệm……...….....2
1.2. Tình hình nghiên cứu trong và ngoài nước…….3
Chương 2: ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU.....7
2.1. Địa điểm, đối tượng và thời gian nghiên cứu….....7
2.2. Phương pháp nghiên cứu…….....7
2.3. Phương pháp thu thập số liệu……...……...……8
2.4. Biến số nghiên cứu và phương pháp đo lường các biến số…….....9
2.5. Phân tích và xử lý số liệu…….....……..9
2.6. Các sai số có thể gặp và biện pháp khống chế sai số ….....……10
2.7. Đạo đức nghiên cứu……10
Chương 3: KẾT QUẢ NGHIÊN CỨU.....11
Chương 4: BÀN LUẬN.....……..……15
KẾT LUẬN
KHUYẾN NGHỊ
TÀI LIỆU THAM KHẢO
PHỤ LỤC

DANH MỤC BẢNG

Bảng 2.1. Tên biến số và định nghĩa biến số…….....…...….....9
Bảng 3.1. Mức độ hài lòng của bệnh nhân đối với cách tư vấn sức khỏe của Bác sĩ lâm sàng và ChatGPT 4.0…….....…...….....15
Bảng 3.2. Tỷ lệ mức độ hài lòng của bệnh nhân về tư vấn sức khỏe của Bác sĩ lâm sàng theo các đặc điểm chung của đối tượng nghiên cứu……..……16
Bảng 3.3. Tỷ lệ mức độ hài lòng của bệnh nhân về tư vấn sức khỏe của ChatGPT theo các đặc điểm chung của đối tượng nghiên cứu……..…….17
Bảng 3.4. Phân tích hồi quy logistic một số yếu tố liên quan đến mức độ hài lòng của bệnh nhân với tư vấn sức khỏe của Bác sĩ lâm sàng…….....18
Bảng 3.5. Phân tích hồi quy logistic một số yếu tố liên quan đến mức độ hài lòng của bệnh nhân với tư vấn sức khỏe của ChatGPT 4.0…….....19

DANH MỤC BIỂU ĐỒ

Biểu đồ 3.1. Phân bố nhóm tuổi của bệnh nhân tham gia nghiên cứu……11
Biểu đồ 3.2. Phân bố giới tính của bệnh nhân tham gia nghiên cứu……12
Biểu đồ 3.3. Phân bố trình độ học vấn của bệnh nhân tham gia nghiên cứu……12
Biểu đồ 3.4. Phân bố tình trạng hôn nhân của bệnh nhân tham gia nghiên cứu…13
Biểu đồ 3.5. Phân bố nơi sống của bệnh nhân tham gia nghiên cứu……...14
Biểu đồ 3.6. Phân bố nghề nghiệp của bệnh nhân tham gia nghiên cứu……14

ĐẶT VẤN ĐỀ

Chat-Generative Pre-Trained Transformer (ChatGPT) là 1 chatbot do OpenAI phát triển dựa trên mô hình Transformer của Google và sử dụng trí tuệ nhân tạo giúp tạo nên các cuộc trò chuyện tự động và trả lời các câu hỏi về nhiều chủ đề và lĩnh vực khác nhau. ChatGPT đã đạt 100 triệu người dùng ngay trong 2 tháng đầu tiên ra mắt [1]. ChatGPT được ghi nhận là khởi đầu cho sự bùng nổ AI, dẫn đến sự phát triển nhanh chóng và chưa từng có trong lĩnh vực trí tuệ nhân tạo [2].

ChatGPT đã đạt hoặc gần ngưỡng vượt qua Kì thi Cấp phép Y tế Hoa Kỳ (USMLE) mà không cần bất kỳ khóa đào tạo chuyên môn nào, cho thấy tiềm năng của nó đối với giáo dục y tế và hỗ trợ đưa ra quyết định lâm sàng [3].Trong một giới hạn nào đó, ChatGPT có thể hỗ trợ các bác sĩ đưa ra các quyết định kịp thời, sáng suốt, cải thiện chất lượng và hiệu quả trong việc chăm sóc sức khỏe [4].

Đại dịch Covid 19 đã đẩy nhanh việc áp dụng dịch vụ chăm sóc sức khỏe ảo, đồng thời số lượng tin nhắn điện tử của bệnh nhân tăng gấp 1,6 lần, mỗi tin nhắn cộng thêm 2,3 phút làm việc trong hồ sơ sức khỏe điện tử và nhiều thời gian làm việc ngoài giờ hơn. Lượng tin nhắn bổ sung dẫn đến tình trạng kiệt sức gia tăng ở các bác sĩ lâm sàng và cũng khiến nhiều khả năng tin nhắn của bệnh nhân sẽ không được trả lời hoặc nhận được những phản hồi không hữu ích [5]. Dịch COVID-19 bùng phát, tại Việt Nam một số tỉnh thành thực hiện giãn cách xã hội, nhiều người dân có tâm lý e ngại đến khám, chữa bệnh ở các cơ sở y tế. Trong khi ở nhà thì mắc phải hàng loạt bệnh thông thường như: Tăng huyết áp, tim mạch, dạ dày, tiểu đường, phụ khoa, xương khớp, hô hấp… lại khó khăn không biết xoay sở thế nào khi gặp vấn đề. Trong khi đó, ChatGPT là công cụ hỗ trợ đầy tiềm năng để giải quyết gánh nặng nêu trên. Hiện nay, chưa có nhiều nghiên cứu so sánh hiệu quả về tư vấn sức khỏe của ChatGPT so với bác sĩ lâm sàng. Chính vì những lý do trên, chúng tôi thực hiện nghiên cứu đề tài: **"Nghiên cứu sự hài lòng của bệnh nhân đối với ChatGPT và bác sĩ lâm sàng trong tư vấn sức khỏe tại các Bệnh viện Tâm Trí năm 2024"** với mục tiêu: *Đánh giá mức độ hài lòng của bệnh nhân sau phản hồi tư vấn sức khỏe từ ChatGPT 4.0 và bác sĩ lâm sàng.*

Chương 1

TỔNG QUAN TÀI LIỆU

1.1. Các khái niệm

1.1.1. Chat GPT

ChatGPT, hay Chat Generative Pre-trained Transformer, là một chatbot được phát triển bởi OpenAI, được ra mắt vào ngày 30 tháng 11 năm 2022. Dựa trên mô hình ngôn ngữ lớn, Chat GPT cho phép người dùng điều chỉnh và hướng dẫn cuộc trò chuyện theo chiều dài, định dạng, phong cách, mức độ chi tiết và ngôn ngữ mong muốn. ChatGPT sử dụng các mô hình GPT cơ bản như GPT-3.5 và GPT-4, được tinh chỉnh đặc biệt cho ứng dụng hội thoại. Quá trình tinh chỉnh này sử dụng cả học có giám sát và học tăng cường, trong đó đặc biệt bao gồm học tăng cường từ phản hồi của con người (RLHF). Các phương pháp này tận dụng sự hỗ trợ của người huấn luyện để cải thiện hiệu suất của mô hình [6].

1.1.2. Sức khỏe

Theo Tổ chức Y tế Thế giới (WHO), Sức khỏe là trạng thái thoải mái toàn diện về thể chất, tinh thần và xã hội chứ không phải chỉ bao gồm có tình trạng không có bệnh hay thương tật [7].

1.1.3. Tư vấn sức khỏe

Tư vấn sức khỏe là một phần trong hệ thống các chương trình y tế cộng đồng được xây dựng để nâng cao sức khỏe cộng đồng và giảm thiểu các nguy cơ về thương tổn tâm thần trong cộng đồng Người tư vấn cần đưa ra các thông tin quan trọng, chính xác để đối tượng có thể tự đánh giá, thấy rõ được vấn đề của họ và họ có thể tự suy nghĩ những vấn đề mà họ phải đương đầu, cuối cùng giúp họ đưa ra các quyết định đúng đắn để có thể giải quyết vấn đề của họ một cách tốt nhất [8].

1.1.4. Hài lòng

Theo Oliver (1985), sự hài lòng là phản ứng của người tiêu dùng đối với việc được đáp ứng những mong muốn. Nghĩa là sự thỏa mãn chính là sự hài lòng của người tiêu dùng trong việc tiêu dùng sản phẩm hoặc dịch vụ do nó đáp ứng những mong muốn của họ, bao gồm cả mức độ đáp ứng trên mức mong muốn và dưới mức mong muốn [9].

1.2. Tình hình nghiên cứu trong và ngoài nước

Trong nghiên cứu được thực hiện bởi John W. Ayers và các cộng sự vào năm 2023, mục tiêu chính là so sánh chất lượng và mức độ đồng cảm trong các phản hồi đến từ chatbot AI và bác sĩ, dựa trên các câu hỏi y tế đặt ra trên diễn đàn truyền thông xã hội. Câu hỏi nghiên cứu tập trung vào việc liệu trợ lý chatbot AI có khả năng cung cấp các câu trả lời đáp ứng được chất lượng và sự đồng cảm tương đương với những gì bác sĩ cung cấp hay không. Kết quả từ nghiên cứu này cho thấy rằng phản hồi của chatbot không chỉ nhận được sự ưu ái hơn từ người dùng mà còn được đánh giá cao hơn một cách đáng kể về mặt chất lượng và sự đồng cảm so với phản hồi của bác sĩ [10]. Tuy nhiên, nghiên cứu này cũng chỉ ra rằng, mặc dù có những kết quả ban đầu hứa hẹn, việc sử dụng chatbot AI trong môi trường lâm sàng cần được nghiên cứu thêm để hiểu rõ hơn về hiệu quả và tiềm năng thực sự của chúng. Bất chấp những hạn chế hiện tại và sự thận trọng trong việc áp dụng công nghệ mới, việc tích hợp trợ lý AI trong quy trình giao tiếp với bệnh nhân được kỳ vọng sẽ cải thiện kết quả điều trị cho cả bác sĩ lẫn bệnh nhân [10].

Nghiên cứu do Aidan Gilson và các cộng sự (2023) tập trung vào việc đánh giá hiệu suất của ChatGPT trong việc trả lời các câu hỏi từ kỳ thi Step 1 và Step 2 của Kỳ thi cấp phép y tế Hoa Kỳ (USMLE). Nghiên cứu cũng nhằm phân tích khả năng hiểu và xử lý thông tin của ChatGPT để người dùng có thể dễ dàng tiếp nhận. Kết quả cho thấy ChatGPT không chỉ vượt trội hơn InstructGPT với mức chênh lệch trung bình là 8,15% trên tất cả các tập dữ liệu, mà còn cho thấy GPT-3 hoạt động chỉ ngang bằng với cơ hội ngẫu nhiên. Đáng chú ý là hiệu suất của ChatGPT giảm đáng kể khi đối mặt với các câu hỏi có độ khó cao hơn trong bộ dữ liệu AMBOSS-Step1, với mức độ ý nghĩa thống kê ($p = 0.01$) [4]. Nghiên cứu cũng nhấn mạnh rằng lý do hợp lý cho việc ChatGPT chọn câu trả lời xuất hiện trong 100% các kết quả đầu ra từ bộ dữ liệu của NBME (Hội đồng giám định y khoa Quốc gia). Điều này chứng tỏ ChatGPT không chỉ cung cấp câu trả lời chính xác mà còn đưa ra bối cảnh logic và thông tin chi tiết cho hầu hết các câu hỏi, gợi mở khả năng của nó như một công cụ giáo dục y tế tương tác hữu ích. Kết luận của nghiên cứu chỉ ra rằng ChatGPT đã đạt được điểm số tương đương với điểm đậu của một sinh viên y khoa năm thứ ba, đánh dấu sự cải tiến đáng kể trong mô hình xử lý ngôn ngữ tự nhiên cho các nhiệm vụ trả lời câu hỏi y tế. Sự kết hợp của những bằng chứng

này cung cấp một lập luận thuyết phục về tiềm năng của ChatGPT trong việc hỗ trợ và nâng cao quá trình học y khoa [4].

Nghiên cứu do Kurt R. Herzer và các cộng sự thực hiện vào năm 2021 đã đưa ra cái nhìn sâu sắc về chăm sóc ảo, mô tả đó là quá trình tương tác từ xa giữa bệnh nhân và bác sĩ liên quan đến việc chẩn đoán, đánh giá và quản lý bệnh. Các tương tác này được thực hiện thông qua sự kết hợp của văn bản, âm thanh, và video, có thể là đồng bộ hoặc không đồng bộ. Nghiên cứu này nhấn mạnh rằng chăm sóc ảo mở ra tương lai mà trong đó một tỷ lệ đáng kể các dịch vụ chăm sóc sức khỏe có thể được cung cấp từ xa qua mạng. Quá trình này được hỗ trợ bởi số lượng ngày càng tăng và sự đa dạng của các thiết bị đeo, thiết bị y tế từ xa, và ứng dụng di động, tất cả đều tích hợp với hồ sơ sức khỏe điện tử. Trong ngắn hạn, chăm sóc ảo có thể đem lại sự chính xác và hiệu quả cao. Tuy nhiên, nghiên cứu cũng nhấn mạnh rằng trách nhiệm cơ bản của bác sĩ đối với bệnh nhân vẫn không thay đổi, ngay cả khi hình thức cung cấp dịch vụ có thể biến đổi. Để chăm sóc ảo thực sự hiệu quả và công bằng, các hệ thống y tế, chương trình y tế, và các công ty công nghệ y tế cần phải chứng minh rằng phương thức này đóng góp một cách có hiệu quả và công bằng vào hệ thống chăm sóc sức khỏe tại Hoa Kỳ [11].

Nghiên cứu của Carlos A Hernandez và cộng sự năm 2023 tập trung vào việc đánh giá độ chính xác của các phản hồi văn bản do AI (TGAI) tạo ra, cụ thể là ChatGPT, liên quan đến câu hỏi về bệnh đái tháo đường loại 2 (T2DM). Mục đích là để kiểm tra tính nhất quán của TGAI trong việc cung cấp tư vấn và giáo dục cho bệnh nhân về các biến chứng của T2DM thông qua việc giải quyết các câu hỏi phổ biến liên quan đến bệnh [12]. Trong tổng số 110 câu hỏi T2DM ban đầu, 40 câu hỏi đã bị loại bỏ do không phù hợp. Kết quả cho thấy, 98,5% phản hồi của AI được đánh giá là phù hợp, từ đó khẳng định độ tin cậy của nó so với các công cụ tìm kiếm trực tuyến truyền thống. Tuy nhiên, tỷ lệ 1,5% phản hồi không phù hợp nhấn mạnh tầm quan trọng của việc cải tiến AI liên tục và tuân thủ nghiêm ngặt các quy trình y tế. Nghiên cứu kết luận rằng TGAI có khả năng cung cấp thông tin y tế có chất lượng và độ tin cậy cao, với 98,5% câu trả lời đáp ứng tiêu chuẩn chăm sóc. Những kết quả này mở ra cơ hội tích hợp AI như một công cụ bổ sung hứa hẹn, nhằm nâng cao kết quả và chất lượng tư vấn, giáo dục cho bệnh nhân [12].

Nghiên cứu của Hana L. Haver và cộng sự (2023) tập trung vào đánh giá hiệu quả của ChatGPT trong việc cung cấp thông tin khuyến nghị về sàng lọc và phòng ngừa ung

thư vú. Kết quả cho thấy ChatGPT đã đưa ra câu trả lời phù hợp với 88% (22 trên 25) câu hỏi trong cả hai bối cảnh được nghiên cứu. Tuy nhiên, một câu hỏi (4%) được đánh giá là không phù hợp và hai câu hỏi (8%) được đánh giá là không đáng tin cậy. Đáng chú ý, phản hồi không phù hợp đề cập đến việc lên lịch chụp X-quang tuyến vú liên quan đến tiêm chủng ngừa COVID-19. Nghiên cứu này nhấn mạnh tiềm năng của ChatGPT trong việc tự động hóa việc cung cấp thông tin chăm sóc sức khỏe chính xác liên quan đến phòng ngừa và sàng lọc ung thư vú. Tuy nhiên, sự cần thiết của sự giám sát bác sĩ trong việc sử dụng công cụ này được nhấn mạnh, nhất là với tỷ lệ phản hồi không phù hợp và không đáng tin cậy được ghi nhận, phản ánh những hạn chế đã được cảnh báo trước đó của ChatGPT, đặc biệt là trong bối cảnh chụp X quang. Nghiên cứu khuyến khích việc nghiên cứu thêm về tiềm năng của mô hình ngôn ngữ lớn trong việc cải thiện việc cung cấp giáo dục và tư vấn chăm sóc sức khỏe [13]. Ngoài ra, nghiên cứu khác của cùng nhóm tác giả đánh giá hiệu quả của ChatGPT trong việc trả lời các câu hỏi cơ bản về phòng ngừa và sàng lọc ung thư phổi, dựa trên đánh giá của các bác sĩ X quang tim lồng ngực được hội đồng chứng nhận. Kết quả từ nghiên cứu này cho thấy 84% phản hồi của ChatGPT là phù hợp, trong khi 8% không phù hợp và 8% không nhất quán. Dù những phát hiện này cung cấp hỗ trợ bổ sung cho tiềm năng của AI trong việc cung cấp tài liệu giáo dục liên quan cho bệnh nhân, vẫn còn tồn tại những lỗ hổng trong kiến thức về chủ đề X quang, điều đó nhấn mạnh tầm quan trọng của sự giám sát của bác sĩ [14].

Nghiên cứu do Harriet Louise Walker và cộng sự thực hiện vào năm 2023 nhằm đánh giá độ tin cậy của thông tin y tế được cung cấp bởi ChatGPT liên quan đến năm bệnh lý gan-tụy-mật. Đánh giá này được thực hiện dựa trên công cụ đảm bảo thông tin chất lượng cho bệnh nhân (EQIP). Theo kết quả, điểm EQIP trung bình cho tất cả các tình trạng bệnh lý được nghiên cứu là 16. Sự đồng thuận giữa các khuyến nghị hướng dẫn chăm sóc sức khỏe và phản hồi của ChatGPT đạt mức 60%. Đáng chú ý là tính nhất quán nội bộ trong các câu trả lời do ChatGPT cung cấp đạt 100%. Nghiên cứu này kết luận rằng ChatGPT cung cấp thông tin y tế có chất lượng tương đối tương đương với thông tin có sẵn trên Internet. Tuy chất lượng của thông tin từ ChatGPT hiện tại còn có những hạn chế, nhưng các mô hình ngôn ngữ lớn đang dần chứng minh tiềm năng trở thành tiêu chuẩn trong tương lai. Điều này sẽ giúp bệnh nhân và các chuyên gia chăm sóc sức khỏe có thể thu thập thông tin y tế một cách hiệu quả hơn [15].

Nghiên cứu do Greeni Maheshwari thực hiện vào năm 2023 nhằm tìm hiểu các yếu tố ảnh hưởng đến ý định chấp nhận và sử dụng ChatGPT của sinh viên trong giáo dục đại học tại Việt Nam. Kết quả phân tích chỉ ra rằng các yếu tố như đánh giá cá nhân về tính dễ sử dụng, hữu ích, cá nhân hóa, và tương tác của ChatGPT đều có ảnh hưởng đáng kể đến thái độ và hành vi của sinh viên đối với công cụ này. Sinh viên có xu hướng chấp nhận và sử dụng ChatGPT nhiều hơn khi họ cảm nhận được sự thuận tiện và thân thiện với người dùng của công cụ. Tuy nhiên, nghiên cứu cũng chỉ ra rằng cảm nhận về độ đáng tin cậy và thông minh của ChatGPT không ảnh hưởng trực tiếp đến ý định sử dụng của sinh viên. Thay vào đó, đánh giá về tính hữu ích của ChatGPT tác động gián tiếp đến ý định sử dụng thông qua yếu tố cá nhân hóa (khả năng ChatGPT được cá nhân hóa theo nhu cầu của sinh viên) và tính tương tác (các tính năng tương tác mà công cụ này cung cấp). Điều này cho thấy sinh viên không coi độ thông minh của ChatGPT là yếu tố quan trọng trong việc quyết định sử dụng công cụ này. Những hiểu biết sâu sắc thu được từ nghiên cứu của Greeni Maheshwari cung cấp thông tin quý giá cho các tổ chức giáo dục trong việc đưa ra quyết định liên quan đến việc triển khai ChatGPT và các công cụ AI tạo sinh khác trong môi trường giáo dục, nhằm tối ưu hóa việc chấp nhận và sử dụng công nghệ trong ngành giáo dục [16].

Chương 2
ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Địa điểm, đối tượng và thời gian nghiên cứu

2.1.1. Địa điểm nghiên cứu

Nghiên cứu được tiến hành tại Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Bệnh viện Đa khoa Tâm Trí Quảng Nam.

2.1.2. Đối tượng nghiên cứu

Bệnh nhân nội trú tại Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Bệnh viện Đa khoa Tâm Trí Quảng Nam.

Tiêu chuẩn lựa chọn

- Bệnh nhân nội trú từ 18 tuổi trở lên tại 02 bệnh viện: Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Bệnh viện Đa khoa Tâm Trí Quảng Nam trong khoảng thời gian nghiên cứu.
- Bệnh nhân nội trú được phỏng vấn ngay trong vòng 24 giờ sau khi nhập viện.
- Bệnh nhân đủ năng lực hành vi, tình trạng sức khỏe cho phép đảm bảo trả lời phỏng vấn.
- Bệnh nhân tự nguyện tham gia và đồng ý thực hiện quy trình nghiên cứu mới được chọn.

Tiêu chuẩn loại trừ

- Bệnh nhân đang có tình trạng cấp cứu.
- Bệnh nhân có những rối loạn về giao tiếp, tâm thần, sa sút trí tuệ.
- Bệnh nhân không đồng ý tham gia nghiên cứu.

2.1.3. Thời gian nghiên cứu

Nghiên cứu được tiến hành từ 02/2024 đến 04/2024. Thời gian xử lý số liệu và viết báo cáo từ 04/2024 đến tháng 06/2024.

2.2. Phương pháp nghiên cứu

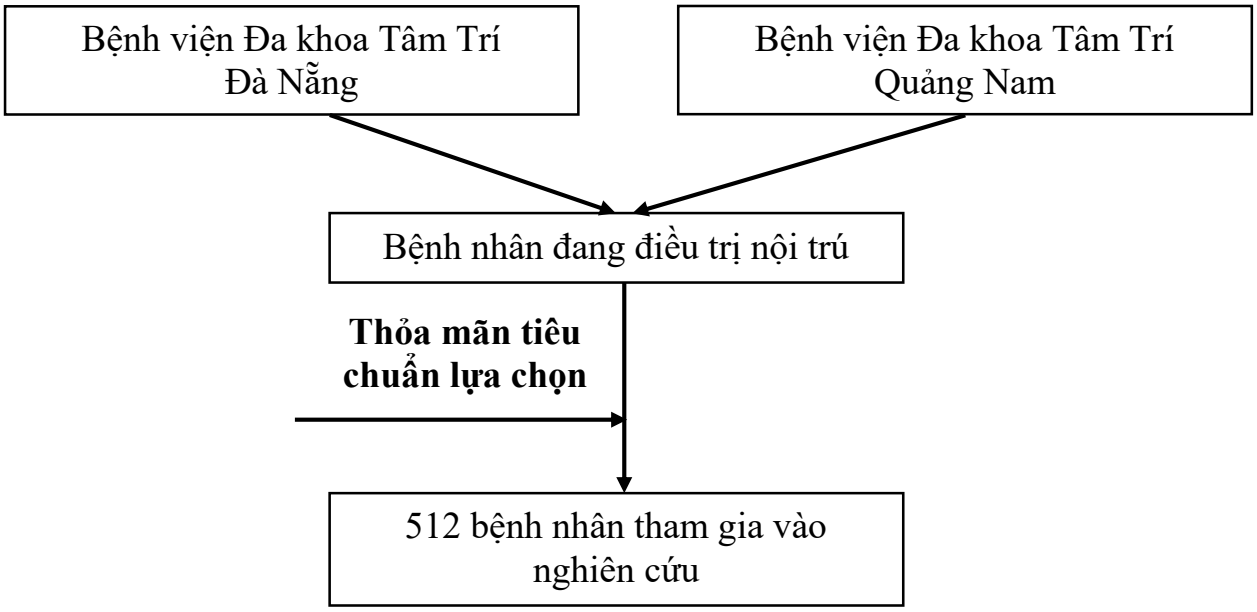
2.2.1. Thiết kế nghiên cứu

Nghiên cứu mô tả cắt ngang có so sánh.

2.2.2. Cỡ mẫu và cách chọn mẫu

- Cỡ mẫu: 512 bệnh nhân đang điều trị nội trú tại Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Bệnh viện Đa khoa Tâm Trí Quảng Nam.

- Phương pháp chọn mẫu: chọn mẫu thuận tiện tất cả các bệnh nhân nội trú đủ tiêu chuẩn lựa chọn và không có tiêu chuẩn loại trừ.



Sơ đồ nghiên cứu và chọn mẫu

2.3. Phương pháp thu thập số liệu

Các biến số về nhân khẩu học của bệnh nhân như tuổi, giới, nơi sống, trình độ học vấn, nghề nghiệp, tình trạng hôn nhân đều được thu thập thông qua hỏi đáp trực tiếp với bệnh nhân.

Câu hỏi của bệnh nhân được thu thập trong vòng 24 giờ kể từ lúc người bệnh nhập viện. Bệnh nhân đặt 1 câu hỏi liên quan đến tình trạng sức khỏe mà họ cần tư vấn, đảm bảo phản ánh chính xác nhu cầu thông tin của bệnh nhân, bác sĩ sau đó sẽ trả lời những câu hỏi này bằng văn bản cũng trong vòng 24 giờ sau khi tiếp nhận câu hỏi của bệnh nhân, mỗi câu trả lời được soạn thảo kỹ lưỡng nhằm cung cấp thông tin chính xác và hữu ích cho bệnh nhân. Cùng những câu hỏi đó, điều tra viên sẽ nhập vào Chatbot, sử dụng kỹ thuật Instruction Prompt để cải thiện độ chính xác và hiệu quả câu trả lời của ChatGPT 4.0, sử dụng các lệnh rõ ràng, chi tiết để hướng dẫn ChatGPT trả lời, nhắc nhở ChatGPT có thể kết hợp bất kỳ nguyên tắc y tế nào vào câu trả lời nếu phù hợp. Sau đó cả hai câu trả lời của ChatGPT 4.0 và bác sĩ lâm sàng được bệnh nhân đánh giá mức độ hài lòng mà không biết nguồn gốc câu trả lời.

2.4. Biến số nghiên cứu và phương pháp đo lường các biến số

Bảng 2.1. Tên biến số và định nghĩa biến số

STT	Tên biến	Định nghĩa biến
1	Tuổi	Tuổi được tính theo năm sinh dương lịch
2	Giới	1. Nam 2. Nữ
3	Nơi sống	1. Thành thị 2. Nông thôn
4	Nghề nghiệp	1. Thất nghiệp/già 2. Có việc làm
5	Trình độ học vấn	1. Trung học phổ phông và thấp hơn 2. Trên trung học phổ thông
6	Tình trạng hôn nhân	1. Góa/ly dị/ly thân/chưa kết hôn 2. Đã kết hôn
7	Mức độ hài lòng của bệnh nhân	1. Rất không hài lòng 2. Không hài lòng 3. Chấp nhận được 4. Hài lòng 5. Rất hài lòng.

Tiêu chí đánh giá mức độ hài lòng của bệnh nhân:

Mức độ hài lòng được tính theo thang điểm Likert, với mỗi câu trả lời có 5 lựa chọn cho điểm từ 1 đến 5 điểm. Nhóm hài lòng có điểm Likert ≥ 4 điểm, nhóm không hài lòng có điểm Likert < 4 [10].

2.5. Phân tích và xử lý số liệu

- Nhập và quản lý số liệu bằng phần mềm Epidata.
- Phân tích số liệu bằng phần mềm SPSS 22.0

Những thông tin về đặc điểm chung của bệnh nhân, tỷ lệ mức độ hài lòng của bệnh nhân với cách tư vấn sức khỏe của bác sĩ lâm sàng và chatGPT được tính ra tần số và tỷ lệ (thống kê mô tả).

Sử dụng mô hình hồi quy logistic đơn biến và mô hình hồi quy logistic đa biến để xác định độ mạnh của các yếu tố liên quan đến mức độ hài lòng của bệnh nhân đối với câu trả lời của bác sĩ lâm sàng và chat GPT 4.0. Kết luận có mối liên quan khi giá trị p < 0,05. Quá trình lựa chọn các biến trong mô hình hồi quy logistic đa biến có chủ đích, bắt đầu bằng một phân tích đơn biến của từng biến. Các biến độc lập được lựa chọn dựa trên phân tích dơn biến. Những biến có giá trị p < 0,25 từ phân tích đơn biến sẽ được đưa vào mô hình hồi quy đa biến để kiểm tra mối liên quan với biến phụ thuộc. Các biến này bao gồm các yếu tố nhân khẩu học như tuổi, giới tính, nơi sống, nghề nghiệp, trình độ học vấn, tình trạng hôn nhân và mức độ hài lòng của bệnh nhân [17]. Trước khi chạy mô hình quy quy, kiểm tra các giả định cần thiết bao gồm kiểm tra tính tuyến tính giữa các biến độc lập và biến phụ thuộc, kiểm tra đa cộng tuyến giữa các biến độc lập bằng hệ số phóng đại phương sai (VIF), Kết quả cho thấy mối quan hệ tuyến tính giữa các biến độc lập và biến phụ thuộc được thỏa mãn và không có hiện tượng đa cộng tuyến.

2.6. Các sai số có thể gặp và biện pháp khống chế sai số

Sai số có thể gặp là: Sai số do điều tra viên, sai số do đối tượng nghiên cứu khai thông tin không chính xác, sai số do nhập và xử lý số liệu.

Cách khắc phục: Phỏng vấn thử để kiểm tra tính phù hợp của bộ câu hỏi, kiểm tra số liệu sau mỗi ngày, tuân thủ theo đúng quy trình nhập và xử lý số liệu.

2.7. Đạo đức nghiên cứu

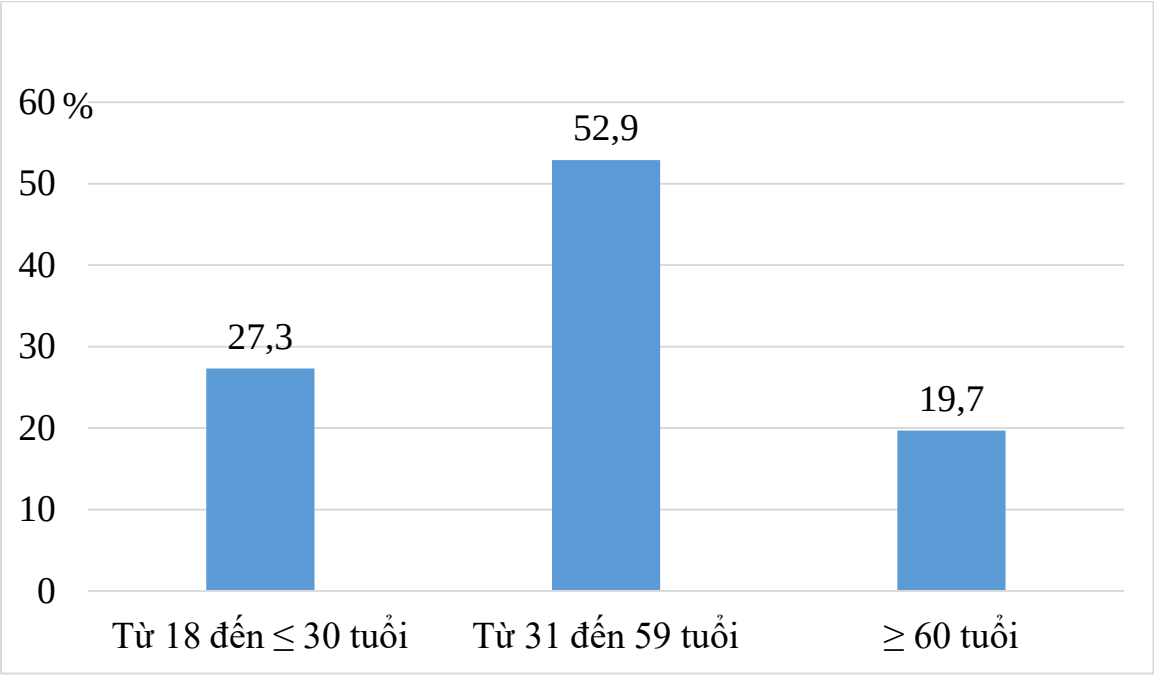
Nghiên cứu được tiến hành sau khi được Hội đồng khoa học - Trường Đại học Phan Châu Trinh thông qua; sự đồng ý của Ban lãnh đạo Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Bệnh viện Đa khoa Tâm Trí Quảng Nam

Đảm bảo quyền tự nguyện tham gia nghiên cứu của bệnh nhân. Những bệnh nhân mời tham gia nghiên cứu được điều tra viên giải thích rõ ràng về mục đích, ý nghĩa, các thông tin sẽ thu thập của cuộc điều tra và bệnh nhân có quyền lựa chọn có đồng ý tham gia vào nghiên cứu hay không.

Bệnh nhân có quyền từ chối không tham gia, có quyền không trả lời, có thể yêu cầu dừng và hủy kết quả phỏng vấn bất cứ khi nào. Các thông tin của bệnh nhân được đảm bảo hoàn toàn bí mật, chỉ phục vụ cho mục đích nghiên cứu, ngoài ra không có mục đích nào khác. Ngoài ra, bệnh nhân tham gia nghiên cứu không phải chi trả thêm bất kỳ một khoản chi phí nào khác.

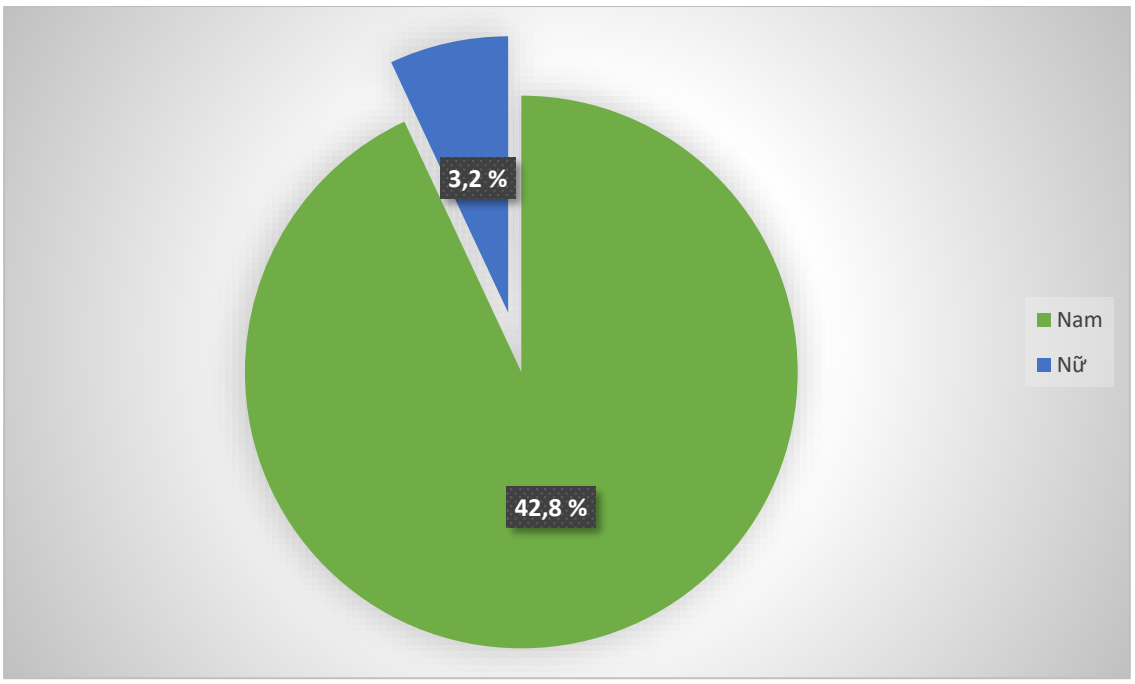
Chương 3
KẾT QUẢ NGHIÊN CỨU

Trong khoảng thời gian từ 02/2024 đến 04/2024, chúng tôi thu thập được 512 bệnh nhân điều trị nội trú tại Bệnh viện Tâm Trí Đà Nẵng và Bệnh viện Tâm Trí Quảng Nam. Tất cả các bệnh nhân đều tự nguyện tham gia vào nghiên cứu và đồng ý cung cấp thông tin.



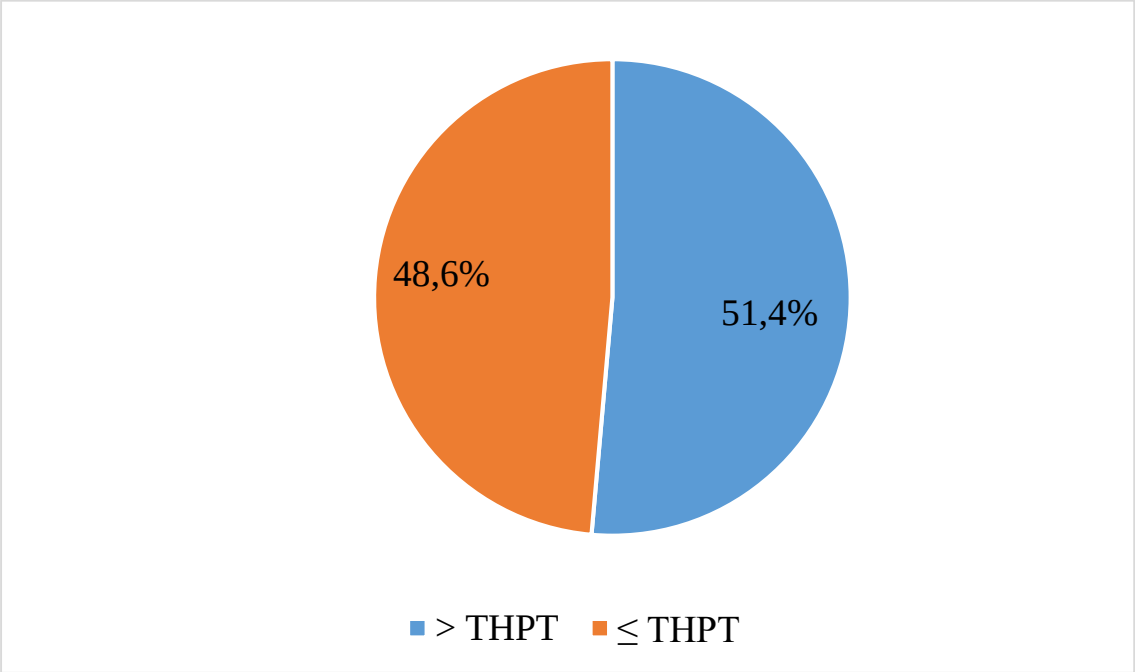
Biểu đồ 3.1. Phân bố nhóm tuổi của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.1 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo độ tuổi cho thấy nhóm từ 31 đến 59 tuổi có tỷ lệ cao nhất, chiếm 52,9%; tiếp theo là nhóm tuổi từ 18 đến 30 chiếm 27,3%; cuối cùng từ 60 tuổi trở lên là 19,7%.



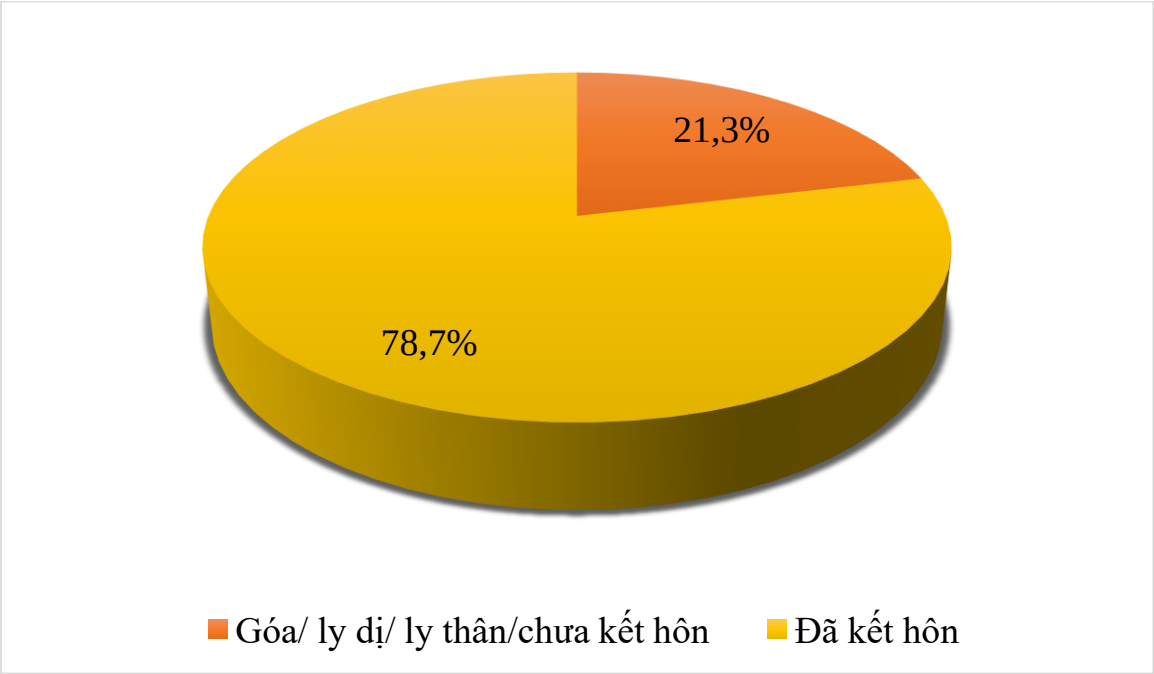
Biểu đồ 3.2. Phân bố giới tính của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.2 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo giới tính, bệnh nhân nam chiếm 42,8% và bệnh nhân nữ chiếm 57,3%.



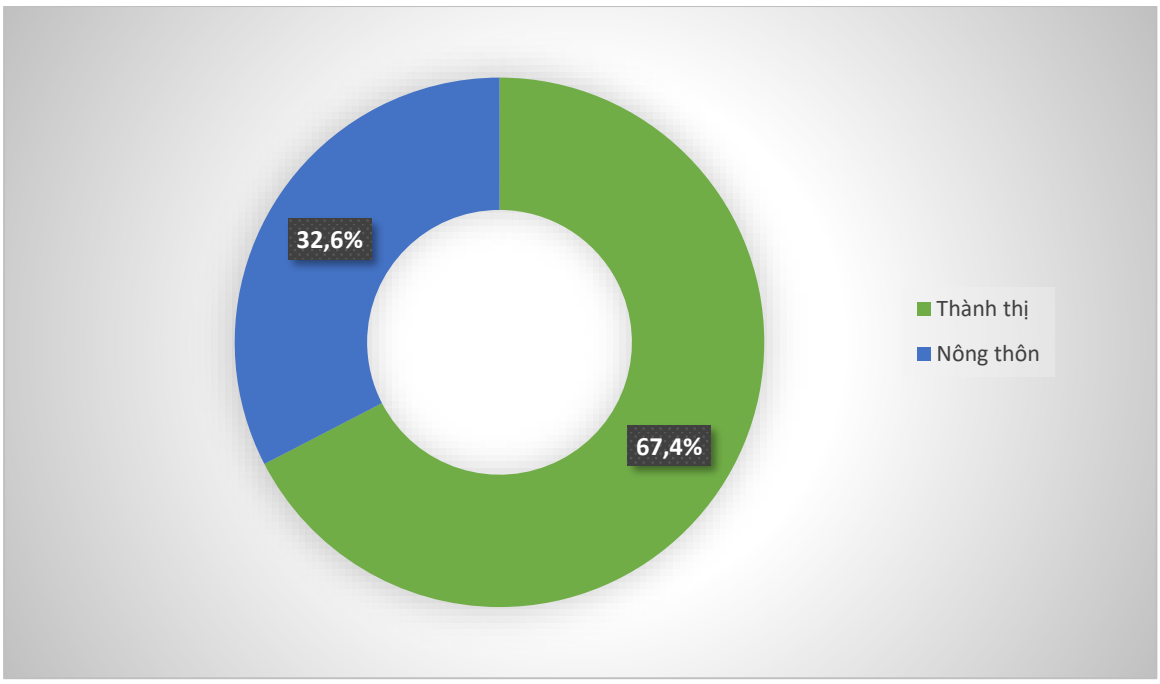
Biểu đồ 3.3. Phân bố trình độ học vấn của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.3 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo trình độ học vấn, bệnh nhân có trình độ >THPT chiếm 51,4% và trình độ ≤ THPT là 48,6%.



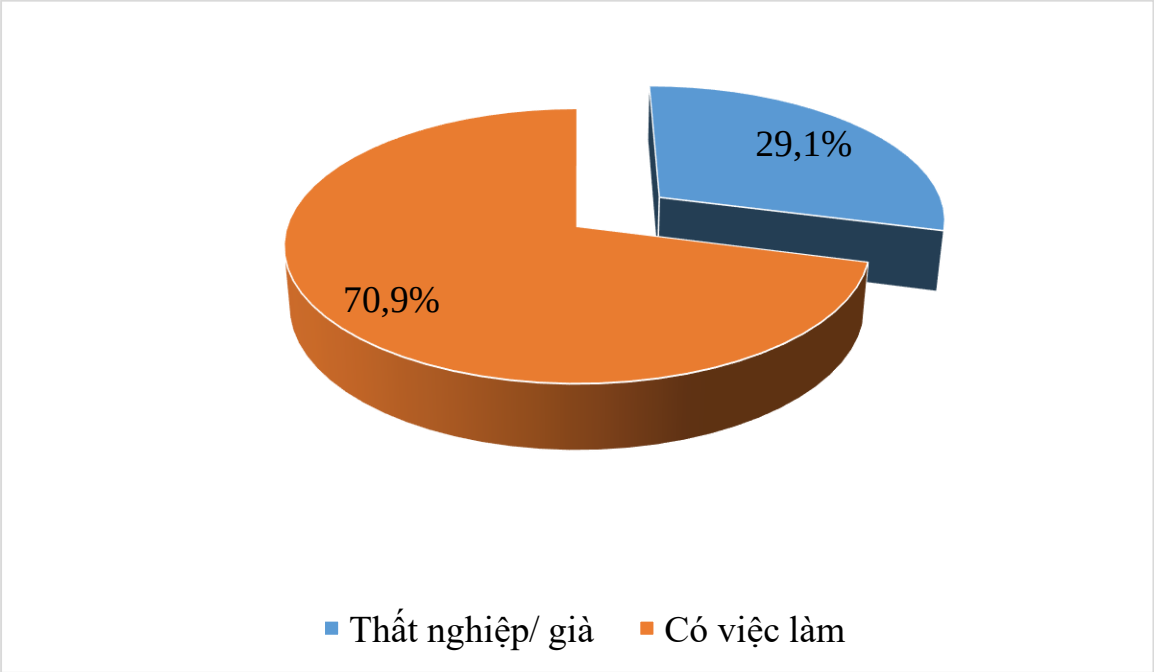
Biểu đồ 3.4. Phân bố tình trạng hôn nhân của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.4 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo tình trạng hôn nhân, bệnh nhân đã kết hôn chiếm tỷ lệ cao là 78,7% và Góa/ ly dị/ ly thân/ chưa kết hôn chiếm 21,3%.



Biểu đồ 3.5. Phân bố nơi sống của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.5 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo nơi sống, bệnh nhân sống ở khu vực thành thị chiếm 67,4% và sống ở nông thôn chiếm 32,6%.



Biểu đồ 3.6. Phân bố nghề nghiệp của bệnh nhân tham gia nghiên cứu (n=512)

Kết quả biểu đồ 3.5 cho thấy: Trong số 512 bệnh nhân tham gia nghiên cứu, phân bố theo nghề nghiệp, bệnh nhân có việc làm chiếm 70,9% và thất nghiệp/ già chiếm 29,1%.

Bảng 3.1. Mức độ hài lòng của bệnh nhân đối với cách tư vấn sức khỏe của Bác sĩ lâm sàng và ChatGPT 4.0 (n = 512)

Mức độ hài lòng	**Bác sĩ lâm sàng**		**ChatGPT 4.0**	
	SL	**%**	**SL**	**%**
Rất không hài lòng	0	0,0	0	0,0
Không hài lòng	9	1,8	3	0,6
Chấp nhận được	163	31,8	113	22,1
Hài lòng	248	48,4	260	50,8
Rất hài lòng	92	18,0	136	26,6
Điểm trung bình ± Độ lệch chuẩn	**3,83 ± 0,73**		**4,03 ± 0,71**	

Kết quả bảng 3.1 cho thấy: Điểm trung bình chung mức độ hài lòng của bệnh nhân đối với bác sĩ lâm sàng là 3,83 ± 0,73 điểm, thấp hơn so với số điểm 4,03 ± 0,71 điểm của chatGPT 4.0. Tỷ lệ bệnh nhân không hài lòng với cách tư vấn sức khỏe của bác sĩ lâm sàng cao hơn so với chatGPT 4.0 (1,8% so với 0,6%). Có 31,8% bệnh nhân bày tỏ mức độ chấp nhận được với cách tư vấn của bác sĩ lâm sàng, trong khi tỷ lệ này ở chatGPT 4.0 thấp hơn, ở mức 22,1%. Tỷ lệ bệnh nhân có mức độ hài lòng về tư vấn sức khỏe của bác sĩ lâm sàng và chatGPT 4.0 khá tương đương, trong đó bác sĩ lâm sàng chiếm 48,4% và chatGPT là 50,8%. Về mức độ rất hài lòng của bệnh nhân với cách tư vấn sức khỏe, chatGPT vượt trội hơn với 26,6% so với 18,0% của bác sĩ lâm sàng. Tỷ lệ bệnh nhân rất hài lòng với bác sĩ lâm sàng và chatGPT 4.0 lần lượt là 18,0% và 26,6%.

Bảng 3.2. Tỷ lệ mức độ hài lòng của bệnh nhân về tư vấn sức khỏe của Bác sĩ lâm sàng theo các đặc điểm chung của đối tượng nghiên cứu (n = 512)

Biến số		**Rất không hài lòng**		**Không hài lòng**		**Chấp nhận được**		**Hài lòng**		**Rất hài lòng**	
		SL	**%**	**SL**	**%**	**SL**	**%**	**SL**	**%**	**SL**	**%**
Nhóm tuổi	Từ 18 đến ≤ 30 tuổi	0	0,0	4	2,9	48	34,3	70	50,0	18	12,9
	Từ 31 đến 59 tuổi	0	0,0	5	1,8	85	31,4	129	47,6	52	19,2
	≥ 60 tuổi	0	0,0	0	0,0	30	29,7	49	38,5	22	21,8
Giới tính	Nam	0	0,0	8	3,7	93	42,5	73	33,3	45	20,5
	Nữ	0	0,0	1	0,3	70	23,9	175	59,7	47	16,0
Trình độ học vấn	> THPT	0	0,0	5	1,9	100	38,0	110	41,8	48	18,3
	≤ THPT	0	0,0	4	1,6	63	25,4	138	55,6	43	17,3
Tình trạng hôn nhân	Góa/ ly dị/ ly thân/chưa kết hôn	0	0,0	4	3,7	38	34,9	45	41,3	22	20,2
	Đã kết hôn	0	0,0	5	1,2	163	31,8	248	48,4	92	18,0
Nơi sống	Thành thị	0	0,0	7	2,0	112	32,5	167	48,4	59	17,1
	Nông thôn	0	0,0	2	1,2	51	30,5	81	48,5	33	19,8
Nghề nghiệp	Thất nghiệp/ già	0	0,0	1	0,7	31	20,8	81	54,4	36	24,2
	Có việc làm	0	0,0	8	2,2	132	36,4	167	46,0	56	15,4

Kết quả bảng 3.2 cho thấy: Nhìn chung bệnh nhân có tỷ lệ hài lòng khá cao về tư vấn sức khỏe của bác sĩ lâm sàng. Trong đó, tỷ lệ hài lòng cao nhất ở bệnh nhân từ 18 đến ≤ 30 tuổi là 50,0%; rất hài long là 12,9% và chỉ có 2,9% không hài lòng. Có 3,7% bệnh nhân nam không hài lòng; 42,5% hài lòng và 20,5% rất hài lòng. Theo trình độ học vấn, tỷ lệ hài lòng của bệnh nhân >THPT là 41,8% và rất hài lòng (18,3%), nhóm trình độ ≤ THPT có tỷ lệ lần lượt là 43,4% hài lòng và 17,3% rất hài lòng. Bệnh nhân đã kết hôn có tỷ lệ hài lòng và rất hài lòng cao lần lượt là 48,4% và 18,0%, tỷ lệ không hài lòng rất thấp (1,2%). Về nghề nghiệp, tỷ lệ hài lòng ở nhóm có việc làm là 46,0% và rất hài lòng là 15,4%, chỉ có 2,2% không hài lòng.

Bảng 3.3. Tỷ lệ mức độ hài lòng của bệnh nhân về tư vấn sức khỏe của chatGPT theo các đặc điểm chung của đối tượng nghiên cứu (n = 512)

Biến số		**Rất không hài lòng**		**Không hài lòng**		**Chấp nhận được**		**Hài lòng**		**Rất hài lòng**	
		SL	**%**	**SL**	**%**	**SL**	**%**	**SL**	**%**	**SL**	**%**
Nhóm tuổi	Từ 18 đến ≤ 30 tuổi	0	0,0	2	1,4	31	22,1	73	52,1	34	24,3
	Từ 31 đến 59 tuổi	0	0,0	1	0,4	64	23,6	134	49,4	72	26,6
	≥ 60 tuổi	0	0,0	0	0,0	18	17,8	53	52,5	30	29,7
Giới tính	Nam	0	0,0	2	0,9	73	33,3	84	38,4	60	27,4
	Nữ	0	0,0	1	0,3	40	13,7	176	60,1	76	25,9
Trình độ học vấn	> THPT	0	0,0	2	0,8	96	36,5	104	38,5	61	23,2
	≤ THPT	0	0,0	1	0,4	17	6,9	156	62,9	74	29,8
Tình trạng hôn nhân	Góa/ ly dị/ ly thân/chưa kết hôn	0	0,0	2	1,8	29	26,6	42	38,5	36	33,0
	Đã kết hôn	0	0,0	1	0,2	84	20,8	218	54,1	100	24,8
Nơi sống	Thành thị	0	0,0	2	0,6	82	23,8	171	49,6	90	26,1
	Nông thôn	0	0,0	1	0,6	31	18,6	89	53,3	46	27,5
Nghề nghiệp	Thất nghiệp/ già	0	0,0	1	0,7	24	16,1	76	51,0	48	32,2
	Có việc làm	0	0,0	2	0,6	89	24,5	184	50,7	88	24,2

Kết quả bảng 3.3 cho thấy: Nhìn chung bệnh nhân có tỷ lệ hài lòng khá cao về tư vấn sức khỏe của chatGPT. Trong đó, tỷ lệ hài lòng cao nhất ở bệnh nhân từ 18 đến ≤ 30 tuổi là 52,1%; rất hài lòng là 24,3% và chỉ có 1,4% không hài lòng. Chỉ có 0,9% bệnh nhân nam không hài lòng; 38,4% hài lòng và 27,4% rất hài lòng. Theo trình độ học vấn, tỷ lệ hài lòng của bệnh nhân >THPT là 38,5% và rất hài lòng (23,2%), nhóm trình độ ≤ THPT có tỷ lệ lần lượt là 62,9% hài lòng và 29,8% rất hài lòng. Bệnh nhân đã kết hôn có tỷ lệ hài lòng và rất hài lòng cao lần lượt là 54,1% và 24,8%, tỷ lệ không hài lòng rất thấp (0,2%). Về nghề nghiệp, tỷ lệ hài lòng ở nhóm có việc làm là 50,7% và rất hài lòng là 24,2%, chỉ có 0,6% không hài lòng.

Bảng 3.4. Phân tích hồi quy logistic một số yếu tố liên quan đến mức độ hài lòng của bệnh nhân với tư vấn sức khỏe của Bác sĩ lâm sàng (n=512)

Biến số		**Mô hình 1 (hồi quy đơn biến)**			**Mô hình 2 (hồi quy đa biến)**		
		OR	**95% CI**	**p**	**OR hiệu chỉnh**	**95% CI**	**p**
Nhóm tuổi	Từ 18 đến ≤ 30 tuổi	1			1		
	Từ 31 đến 59 tuổi	1,19	0,78 -1,82	0,43	1,20	0,72 – 2,01	0,49
	≥ 60 tuổi	1,40	0,81 – 2,42	0,23	0,61	0,27 – 1,35	0,22
Giới tính	Nam	1			1		
	Nữ	2,68	1,84 – 3,90	**<0,001**	2,58	1,73 – 3,86	**<0,001**
Trình độ học vấn	> THPT	1			1		
	≤ THPT	1,81	1,24-2,62	**0,002**	1,77	1,19 – 2,63	**0,005**
Tình trạng hôn nhân	Góa/ ly dị/ ly thân/chưa kết hôn	1			1		
	Đã kết hôn	1,32	0,85-2,04	0,22	0,98	0,58 – 1,68	0,95
Nơi sống	Thành thị	1			——	——	———
	Nông thôn	1,13	0,76 – 1,68	0,54			
Nghề nghiệp	Thất nghiệp/ già	1			1		
	Có việc làm	0,44	0,28-0,68	**<0,001**	0,28	0,14 – 0,54	**<0,001**

Kết quả bảng 3.2 cho thấy: Giới tính, trình độ học vấn, tình trạng hôn nhân, nghề nghiệp có liên quan đến mức độ hài lòng tư vấn sức khỏe của bệnh nhân với bác sĩ lâm sàng (p<0,05). Trong mô hình hồi quy logistic đa biến, bệnh nhân nữ có mức độ hài lòng với cách tư vấn sức khỏe của bác sĩ lâm sàng nhiều gấp 2,58 lần so với bệnh nhân nam (OR= 2,58; 95%CI: 1,73 – 3,86). Bệnh nhân có trình độ học vấn ≤ THPT có mức độ hài lòng nhiều gấp 1,77 lần người có trình độ > THPT (OR= 1,77; 95%CI: 1,19 – 2,63). Bệnh nhân có việc làm có mức độ hài lòng thấp hơn 0,28 lần so với người thất nghiệp/già (OR= 0,28; 95%CI: 0,14 – 0,54).

Bảng 3.5. Phân tích hồi quy logistic một số yếu tố liên quan đến mức độ hài lòng của bệnh nhân với tư vấn sức khỏe của chatGPT 4.0 (n=512)

Biến số		**Mô hình 1 (hồi quy đơn biến)**			**Mô hình 2 (hồi quy đa biến)**		
		OR	**95% CI**	**p**	**OR hiệu chỉnh**	**95% CI**	**p**
Nhóm tuổi	Từ 18 đến ≤ 30 tuổi	1			_____	_____	_____
	Từ 31 đến 59 tuổi	0,98	0,61 – 1,58	0,93			
	≥ 60 tuổi	1,42	0,75 – 2,70	0,28			
Giới tính	Nam	1			1		
	Nữ	3,20	2,08 – 4,93	**<0,001**	2,78	1,73 - 4,46	**<0,001**
Trình độ học vấn	> THPT	1			1		
	≤ THPT	7,62	4,44-13,09	**<0,001**	7,64	4,38-13,32	**<0,001**
Tình trạng hôn nhân	Góa/ ly dị/ ly thân/ chưa kết hôn	1			1		
	Đã kết hôn	1,49	0,92-2,40	0,11	1,17	0,68-2,01	0,57
Nơi sống	Thành thị	1			1		
	Nông thôn	1,36	0,86 – 2,15	0,19	1,15	0,69 – 1,91	0,59
Nghề nghiệp	Thất nghiệp/ già	1			1		
	Có việc làm	0,60	0,38-1,00	**0,043**	0,48	0,28 – 0,82	**0,007**

Kết quả bảng 3.3 cho thấy: Giới tính, trình độ học vấn, tình trạng hôn nhân, nghề nghiệp có liên quan đến mức độ hài lòng tư vấn sức khỏe của bệnh nhân với chatGPT (p<0,05). Trong mô hình hồi quy logistic đa biến, bệnh nhân nữ có mức độ hài lòng với cách tư vấn sức khỏe của chatGPT nhiều gấp 2,78 lần so với bệnh nhân nam (OR= 2,78; 95%CI: 1,73 -4,46). Bệnh nhân có trình độ học vấn ≤ THPT có mức độ hài lòng nhiều gấp 7,64 lần so với người có trình độ > THPT (OR= 7,64; 95%CI: 24,38 – 13,32). Bệnh nhân có việc làm có mức độ hài lòng thấp hơn 0,48 lần so với người thất nghiệp/già (OR= 0,48; 95%CI: 0,28-0,82).

Chương 4
BÀN LUẬN

ChatGPT sử dụng công nghệ xử lý ngôn ngữ tự nhiên tập trung vào sự tương tác giữa máy tính và ngôn ngữ tự nhiên của con người để cung cấp thông tin y tế trực tuyến, có khả năng cải thiện kết quả điều trị và sự hài lòng của bệnh nhân. Tuy nhiên, độ chính xác và chất lượng của thông tin cần được đánh giá cẩn thận. Trong nghiên cứu này, chúng tôi đã phân tích phản hồi Bác sĩ lâm sàng và ChatGPT 4.0 với 512 câu hỏi của 512 bệnh nhân điều trị nội trú tại Bệnh viện Đa khoa Tâm Trí Đà Nẵng và Tâm Trí Quảng Nam, từ đó đánh giá mức độ hài lòng của bệnh nhân về phản hồi của bác sĩ lâm sàng so với chatbot. Trong nghiên cứu của chúng tôi có 219/512 nam (chiếm 42,8%) và nữ có 293/512 (chiếm 57,3%). Tỷ lệ nam/nữ là 219/293 (1/1,3). Nhóm 31-59 tuổi chiếm đa số (52,9%), theo sau là 18-30 tuổi (27,3%) và 60 tuổi trở lên (19,7%). Hơn 1/2 bệnh nhân có trình độ >THPT (51,4%) và 70,9% bệnh nhân đang có việc làm.

Kết quả nghiên cứu cho thấy bệnh nhân có mức độ hài lòng với ChatGPT cao hơn so với bác sĩ lâm sàng, cụ thể: tỷ lệ không hài lòng thấp hơn (0,6% của chatGPT so với 1,8% của bác sĩ) và tỷ lệ rất hài lòng cao hơn (26,6% của chatGPT so với 18,0% của bác sĩ). Mức độ hài lòng người bệnh chủ yếu tập trung ở mức 4 (hài lòng) với chatGPT là 50,8% cao hơn so với Bác sĩ là 48,4%. Một số nghiên cứu chỉ ra rằng, Chatbot hoạt động gần ngưỡng vượt qua Kỳ thi cấp giấy phép y tế của Hoa Kỳ mà không cần bất kỳ khóa đào tạo chuyên môn nào và chúng cung cấp các câu trả lời mang tính đàm thoại, dễ hiểu và phù hợp với người dùng [3], [18]. Điều này có thể giải thích lý do ChatGPT 4.0 có mức độ hài lòng cao hơn so với bác sĩ lâm sàng trong nghiên cứu này. Ngoài ra, mức độ không hài lòng và chấp nhận được của bệnh nhân với cách tư vấn sức khỏe của bác sĩ lâm sàng chiếm tỷ lệ khá cao (cộng tổng 2 mức độ này là 33,6%), điều này có thể được giải thích là vì các bác sĩ bị áp lực với số lượng bệnh nhân lớn, không dành được nhiều thời gian để trả lời câu hỏi của bệnh nhân. Do đó những lời tư vấn sức khỏe của bác sĩ thường sẽ ngắn gọn, đi vào trọng tâm, không diễn giải chi tiết như chatGPT. Mức độ hài lòng và rất hài lòng của bệnh nhân khi được tiếp

nhận câu trả lời của chatGPT cao hơn đáng kể khi so với câu trả lời của bác sĩ lâm sàng (77,4% so với 66,4%), kết quả này của chúng tôi cũng khá tương đồng với nghiên cứu của John W. Ayers được thực hiện vào năm 2022, khi mà tỷ lệ phản hồi có mức độ hài lòng và rất hài lòng của đối tượng nghiên cứu đối với câu trả lời của chatGPT cao hơn so với câu trả lời của bác sĩ lâm sàng [10].

Nghiên cứu của chúng tôi phát hiện ra rằng các yếu tố bao gồm giới tính, trình độ học vấn, nghề nghiệp có liên quan đến mức độ hài lòng của bệnh nhân về sự tư vấn sức khỏe của bác sĩ lâm sàng và chatGPT 4.0 ($p<0,05$). Trong đó, bệnh nhân có trình độ học vấn thấp hơn có mức độ hài lòng về sự tư vấn của bác sĩ và chatGPT nhiều hơn so với những người có trình độ học vấn trên THPT. Khi phân tích mối tương quan giữa trình độ học vấn và tỷ lệ hài lòng người bệnh, nhận thấy trình độ học vấn tỷ lệ nghịch với sự hài lòng. Theo phân tích của chúng tôi, những người có trình độ học vấn cao thường có địa vị xã hội và điều kiện kinh tế tốt hơn, dẫn đến kỳ vọng cao hơn về chất lượng dịch vụ và sự tư vấn từ các bác sĩ. Vì vậy, khi những kỳ vọng này không được đáp ứng hoàn toàn, họ có xu hướng cảm thấy không hài lòng. Ngược lại, những người bệnh có trình độ học vấn thấp hơn thường có ít yêu cầu hơn, dẫn đến mức độ hài lòng cao hơn về sự tư vấn sức khỏe của bác sĩ.

Nghiên cứu này cho thấy ChatGPT 4.0 có tiềm năng trở thành một công cụ hữu ích cho việc tư vấn sức khỏe vì chúng có nguồn tài nguyên có giá trị để truy cập nhanh chóng thông tin y tế với phản hồi nhanh giúp nâng cao hiệu quả chăm sóc sức khỏe và ra quyết định phức tạp. Chatbot có thể hữu ích trong việc tạo ra thông tin giáo dục bệnh nhân được cá nhân hóa phù hợp với ngôn ngữ và mức độ hiểu biết về sức khỏe nhất định [19].

Ngoài ra khi có sự hỗ trợ bởi trợ lý AI có thể giải phóng năng suất làm việc cho nhân viên y tế, giúp họ sử dụng thời gian hiệu quả cho các nhiệm vụ phức tạp hơn. Hơn nữa, sử dụng AI giúp họ đưa ra phản hồi nhất quán và cải thiện kỹ năng giao tiếp hơn khi được tham khảo và chỉnh sửa từ bản nháp do AI viết. Câu trả lời nhanh chóng, đồng cảm và chất lượng cao của Chatbot đối với các câu hỏi của bệnh nhân có thể giảm số lần khám lâm sàng không cần thiết, giải phóng nguồn lực nhân viên y tế cho những bệnh nhân khác thực sự cần [10]. Đối với những người bệnh có hạn chế về khả năng

di chuyển hoặc lo ngại về chi phí y tế có thể dễ dàng tìm câu trả lời về vấn đề sức khỏe của họ thông qua Chatbot.

Tuy nhiên, không giống như nhiều lĩnh vực khác, thực hành y tế không thể dựa vào một công cụ và công cụ này thỉnh thoảng cung cấp câu trả lời không chính xác như chatGPT [20]. Ngoài ra còn có những lo ngại về tính minh bạch vì chatbot cung cấp các tài liệu tham khảo không đúng khi người dùng yêu cầu cung cấp nguồn. Vì vậy, không nên chỉ dựa vào các phiên bản hiện có của công cụ AI dành cho cộng đồng là nguồn thông tin y tế duy nhất mà cần có sự quản lý chặt chẽ bởi các chuyên gia và các thông tin y tế được kiểm duyệt (ví dụ: tài liệu y tế, cơ sở dữ liệu dược lý và hồ sơ y tế điện tử…).

KẾT LUẬN

Điểm trung bình chung về mức độ hài lòng của bệnh nhân khi được bác sĩ lâm sàng tư vấn sức khỏe là 3,83 ± 0,73 điểm, đối với chatGPT 4.0 là 4,03 ± 0,71 điểm. Trong đó, tỷ lệ bệnh nhân rất hài lòng với chatGPT và với bác sĩ lâm sàng lần lượt là 26,6% và 18,0%. Tỷ lệ bệnh nhân hài lòng với chatGPT là 50,8%, khá tương đương với tỷ lệ 48,4% của bác sĩ lâm sàng. Có 1,8% bệnh nhân không hài lòng với cách tư vấn sức khỏe của bác sĩ lâm sàng, tỷ lệ này ở chatGPT thấp hơn, chỉ có 0,6%.

Trong các yếu tố được khảo sát, giới tính, trình độ học vấn, tình trạng hôn nhân và nghề nghiệp có mối liên quan có ý nghĩa thống kê đến mức độ hài lòng của bệnh nhân khi nhận được những lời tư vấn sức khỏe của bác sĩ lâm sàng và của chatGPT ($p < 0,05$). Cụ thể, bệnh nhân nữ thường hài lòng hơn đáng kể so với bệnh nhân nam khi được nghe bác sĩ lâm sàng và chatGPT tư vấn sức khỏe, với tỷ lệ hài lòng cao hơn gần 3 lần. Bên cạnh đó, những người có trình độ học vấn từ THPT trở xuống cũng đạt được mức độ hài lòng cao hơn so với những người có trình độ học vấn trên THPT, với tỷ lệ hài lòng cao hơn gần gấp 2 lần đối với bác sĩ lâm sàng và gấp hơn 7 lần đối với chatGPT. Ngược lại, bệnh nhân có việc làm thường ít hài lòng với những câu trả lời của bác sĩ lâm sàng và chatGPT hơn so với những người không có việc làm hoặc đã nghỉ hưu.

KHUYẾN NGHỊ

Cần tiến hành thêm các nghiên cứu để xác định mức độ tin cậy của thông tin y tế do chatGPT cung cấp, bao gồm thử nghiệm với quần thể lớn các nhân viên y tế hoặc bệnh nhân và với nhiều loại câu hỏi khác nhau. Điều này sẽ giúp đánh giá sự phát triển của chatGPT theo thời gian và giải quyết các thách thức liên quan đến đạo đức, tính minh bạch, mức độ bảo mật và quyền riêng tư dữ liệu, cũng như các vấn đề pháp lý trong y tế.

Để đảm bảo tính hiệu quả và an toàn khi triển khai AI trong chăm sóc sức khỏe, việc sử dụng các nguồn thông tin y tế uy tín và những khuyến cáo được cập nhật, hình ảnh, và thông tin từ các bảng biểu là thiết yếu. Ngoài ra, cần phải thiết lập các tiêu chuẩn và quy định vừa linh hoạt vừa chặt chẽ, để quản lý việc áp dụng công nghệ AI nói chung và chatGPT nói riêng, nhằm đảm bảo rằng ứng dụng này mang lại lợi ích tối đa và giảm thiểu rủi ro trong ngành y tế.

TÀI LIỆU THAM KHẢO

1. Hu Krystal J (2023), "ChatGPT sets record for fastest-growing user base-analyst note", *Reuters*. 12, p. 2023.
2. Weise, et al (2023), Inside the A.I. Arms Race That Changed Silicon Valley Forever, *The New York Times*.
3. Kung T. H., Cheatham M., Medenilla A., et al. (2023), "Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models", *PLOS Digit Health*. 2(2), p. e0000198
4. Gilson, Aidan, et al (2023), "How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment", *JMIR medical education*. 9. e45312. 10.2196/45312.
5. Resnick M. J. (2019), "Re: Physicians' Well-Being Linked to In-Basket Messages Generated by Algorithms in Electronic Health Records", *J. Urol*. 202, p. 859.
6. Lock and Samantha (2022), *What is AI chatbot phenomenon ChatGPT and could it replace humans?*, accessed
7. WHO (1948), "Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference".
8. Đàm Khai Hoàn (2007), *Giáo dục và nâng cao sức khỏe*, Nhà xuất bản Y học.
9. Oliver R L. and W.O. Bearden (1985), "Disconfirmation processes and consumer evaluations in product usage", *Journal of business Research*. vol. 13, 235-246.
10. Ayers J. W., Poliak A., et al. (2023), "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum", *JAMA Intern Med*. 183(6), pp. 589-596.
11. Bashshur R. L., Howell J. D., Krupinski E. A., et al. (2016), "The Empirical Foundations of Telemedicine Interventions in Primary Care", *Telemed J E Health*. 22(5), pp. 342-75.
12. Hernandez C. A., Vazquez Gonzalez A. E., Polianovskaia A., et al. (2023), "The Future of Patient Education: AI-Driven Guide for Típe 2 Diabetes", *Cureus*. 15(11), p. e48919.
13. Haver H. L., Ambinder E., et al. (2023), "Appropriateness of Breast Cancer Prevention and Screening Recommendations Provided by ChatGPT", *Radiology*. 307(4), p. e230424.

14. Haver HL, Lin CT, et al (2023), "Evaluating ChatGPT's Accuracy in Lung Cancer Prevention and Screening Recommendations", *Radiology: Cardiothoracic Imaging*. 5(4).
15. H. L. Walker, et al (2023), "Reliability of Medical Information Provided by ChatGPT: Assessment Against Clinical Guidelines and Patient Information Quality Instrument", *J. Med. Internet Res*. Vol. 25
16. Maheshwari G. (2023), "Factors influencing students' intention to adopt and use ChatGPT in higher education: A study in the Vietnamese context", *Education and Information Technologies*. 10.1007/s10639-023-12333-z.
17. Bursac, ZoranGauss, C. HeathWilliams, et al. (2008), "Purposeful selection of variables in logistic regression", *Source Code for Biology and Medicine*. 3(1), p. 17.
18. Gilson, Aidan Safranek, ConradHuang, et al. (2022), "How Does ChatGPT Perform on the Medical Licensing Exams? The Implications of Large Language Models for Medical Education and Knowledge Assessment", *medRxiv*, p. 2022.12.23.22283901.
19. E. Idowu, K. Potter (2024), "AI-Powered Chatbots for Patient Education and Engagement", *Healthcare*.
20. Goodman, S.Patrinely, J. R.Stone, et al. (2023), "Accuracy and Reliability of Chatbot Responses to Physician Questions", *JAMA Netw Open*. 6(10), p. e2336483.

PHỤ LỤC

PHIẾU THU THẬP SỐ LIỆU

Mã số phiếu:.....

1. Họ và tên bệnh nhân:.....

2. Ngày, tháng, năm sinh:.....Tuổi:…….

3. Giới: ☐ Nam ☐ Nữ

4. Nơi sống: ☐ Thành thị ☐ Nông thôn

5. Nghề nghiệp: ☐ Thất nghiệp/già ☐ Có việc làm

6. Trình độ học vấn: ☐ Trung học phổ phông và thấp hơn ☐ Trên trung học phổ thông

7. Tình trạng hôn nhân: ☐ Góa/ly dị/ly thân/chưa kết hôn. ☐ Đã kết hôn

8. Câu hỏi của bệnh nhân:.....

8.1 Câu trả lời của Bác sĩ A (Bác sĩ lâm sàng):.....

*Mức độ hài lòng đối với câu trả lời của Bác sĩ A (Bác sĩ lâm sàng):

☐ Rất không hài lòng ☐ Không hài lòng ☐ Chấp nhận được ☐ Hài lòng ☐ Rất hài lòng

8.2. Câu trả lời của Bác sĩ B (ChatGPT-4):.....

*Mức độ hài lòng đối với câu trả lời của Bác sĩ B (ChatGPT-4):

☐ Rất không hài lòng ☐ Không hài lòng ☐ Chấp nhận được ☐ Hài lòng ☐ Rất hài lòng