

## Số lần xét nghiệm cần thiết virus SARS-COV-2 để khẳng định một âm tính thật

A necessary number of SARS-COV-2 virus testing for confirming a true negative

Võ Nhân Văn<sup>a,b\*</sup>, Nguyễn Phúc Minh Tú<sup>a,b</sup>, Trần Bàn Thạch<sup>a,b</sup>, Lưu Văn Hiền<sup>a,b</sup>,  
Nguyễn Quang Vinh<sup>c</sup>, Nguyễn Tấn Trung<sup>d</sup>  
Vo Van Nhan<sup>a,b\*</sup>, Nguyen Phuc Minh Tu<sup>a,b</sup>, Tran Thach Ban<sup>a,b</sup>, Luu Van Hien<sup>a,b</sup>,  
Nguyen Quang Vinh<sup>c</sup>, Nguyen Tan Trung<sup>d</sup>

<sup>a</sup>Khoa Công nghệ thông tin, Đại học Duy Tân, Đà Nẵng, Việt Nam

<sup>a</sup>Faculty of Information Technology, Duy Tan University, Da Nang, 550000, Vietnam

<sup>b</sup>Viện Nghiên cứu và Phát triển Công nghệ Cao, Đại học Duy Tân, Đà Nẵng, Việt Nam

<sup>b</sup>Institute of Research and Development, Duy Tan University, Da Nang, 550000, Vietnam

<sup>c</sup>Tổng công ty Điện lực Tp.HCM, Hồ Chí Minh, Việt Nam

<sup>c</sup>HoChiMinh city Power Corporation, Ho Chi Minh, Vietnam

<sup>d</sup>Trường Cao đẳng Việt Nam- Hàn Quốc, Bình Dương, Việt Nam

<sup>d</sup>Vietnam - Korea college of Binhduong province, Binh Duong, Vietnam

(Ngày nhận bài: 18/5/2021, ngày phản biện xong: 26/6/2021, ngày chấp nhận đăng: 19/9/2021)

### Tóm tắt

Trong bài báo này, các dự đoán về kết quả xét nghiệm Virus SARS-COV-2 sử dụng xác suất có điều kiện theo định lý Bayes sẽ được phân tích. Đầu tiên, bài báo sẽ định nghĩa một số thuật ngữ được dùng trong dự đoán xác suất bị lây nhiễm Virus. Sau đó, bài báo khảo sát một số ví dụ để phân tích việc dự đoán khả năng âm tính giả và dương tính giả từ các kết quả xét nghiệm bằng cách áp dụng định lý Bayes. Cuối cùng, các khuyến cáo về số lần xét nghiệm để đạt được xác suất âm tính thật là chính xác.

*Từ khóa:* xác suất có điều kiện; định lý Bayes; âm tính giả; dương tính giả.

### Abstract

In this paper, the predictions about SARS-COV-2 virus test results are analyzed by using a conditional probability. First, the paper defines some terms to estimate the probability that a person is infected with the virus. Then, the paper considers some examples to analyze the predictions of false negative and false positive from the test results by applying Bayes' theorem. Finally, the number of negative test results in a row is recommended to guarantee the exact true negative cases.

*Keywords:* Conditional Probability; Bayes' theorem; False negative; False positive.

### 1. Đặt vấn đề

Đến giữa năm 2021, Việt Nam đã trải qua bốn đợt cao điểm lây truyền trong cộng đồng

(tháng 1, tháng 7 năm 2020, tháng 1, và tháng 4 năm 2021) và báo cáo tổng cộng 21.943 trường hợp nhiễm SARS-CoV-2 được PCR xác nhận

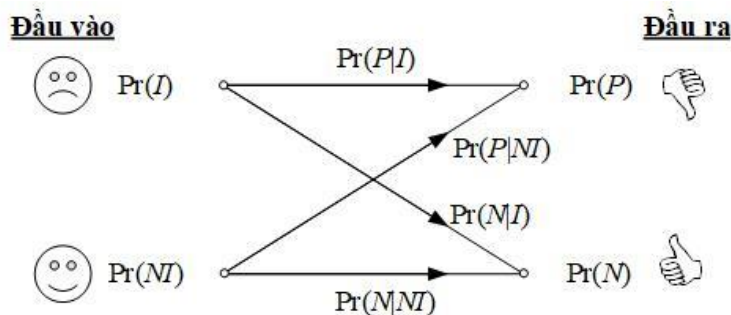
\*Corresponding Author: Van Nhan Vo, Faculty of Information Technology, Duy Tan University, 55000, Danang, Vietnam; Institute of Research and Development, Duy Tan University, 550000, Danang, Vietnam  
Email: vonhanvan@dtu.edu.vn

với 70 trường hợp tử vong vào ngày 09/7/2021 [1]. Trong đó, vào ngày 28/4 xuất hiện một ca nhiễm tại Hà Nam sau 3 lần kiểm tra âm tính và bắt đầu cho đợt lây nhiễm trong cộng đồng cao điểm thứ tư tại Việt Nam [2]. Vậy câu hỏi đặt ra là liệu rằng một trường hợp cách ly sau 3 lần kiểm tra âm tính thì có phải người đó không nhiễm Virus?

Để trả lời câu hỏi này, dựa theo hướng dẫn của một chuyên gia kỹ thuật Giáo Sư DesMcLernon (Đại học Queen, Belfast, Anh), trước tiên chúng ta hãy phân biệt hai khái niệm là trạng thái nhiễm Virus và kết quả xét nghiệm. Trạng thái nhiễm Virus là một cá nhân hoặc là bị nhiễm, hoặc là không bị nhiễm. Tuy đơn giản như vậy, nhưng làm thế nào để xác nhận được là trường hợp nào bị nhiễm và trường hợp nào không bị nhiễm thì không phải là một điều đơn giản. Nguyên nhân là trạng thái

nhiễm và không nhiễm là một hiện tượng chúng ta khó (hoặc không) quan sát được. Do đó, chúng ta phải dùng các phương pháp gián tiếp là xét nghiệm bằng các phương pháp sinh hóa, ở đây sử dụng phương pháp PRC định tính. Kết quả của xét nghiệm cho chúng ta hai giá trị: dương tính (Positive) hay âm tính (Negative) [3].

Tuy nhiên, không có một phương pháp xét nghiệm gián tiếp nào, dù kỹ thuật hiện đại đến đâu đi nữa, là hoàn hảo và chính xác tuyệt đối do có sự nhiễm bẩn và/hoặc lỗi của con người trong quá trình xử lý mẫu hoặc dữ liệu (rất khó để loại bỏ hoàn toàn). Do đó, có thể một số người có kết quả dương tính, nhưng thực sự không bị nhiễm Virus và có thể một số người có kết quả âm tính, nhưng trong thực tế lại nhiễm Virus. Đến đây, một mô hình tính xác suất hậu nghiệm được trình bày như Hình 1, trong đó:



Hình 1. Mô hình tính xác suất hậu nghiệm

- Xác suất tiên nghiệm (Prior Probability) là khả năng dự đoán một biến cố sẽ xuất hiện theo một cách nhất định dựa trên tỉ lệ phần trăm. Ở đây, xác suất mà một người nào đó được chọn ngẫu nhiên trong quần thể bị nhiễm Virus  $Pr(I)$  hoặc không bị nhiễm Virus  $Pr(NI)$ , trong đó  $Pr(I) + Pr(NI) = 1$ .

- Xác suất dương tính thật (True Positive hay Sensitivity)  $Pr(P|I)$ : Xác suất xét nghiệm dương tính khi mà bệnh nhân đã bị nhiễm Virus. Lý tưởng nhất là gần bằng 1.

- Xác suất dương tính giả (False Positive)  $Pr(P|NI)$ : Xác suất xét nghiệm dương tính

khi mà bệnh nhân chưa bị nhiễm Virus. Hoặc xác suất dương tính giả mà bệnh nhân đã không bị nhiễm Virus. Lý tưởng nhất là gần bằng 0.

- Xác suất âm tính thật (True Negative hay Specificity)  $Pr(N|NI)$ : Xác suất xét nghiệm âm tính khi mà bệnh nhân đã bị nhiễm Virus. Lý tưởng nhất là gần 1.

- Xác suất âm tính giả (False Negative)  $Pr(N|I)$ : Xác suất xét nghiệm âm tính khi mà bệnh nhân bị nhiễm Virus. Lý tưởng nhất là gần bằng 0.

## 2. Một số ví dụ về việc phân tích kết quả xét nghiệm

Trong phần này, các ví dụ sẽ được phân tích để kiểm tra tính đúng đắn của kết quả xét nghiệm. Đầu tiên, để hiểu rõ về xác suất hậu nghiệm trong quá trình xét nghiệm, một ví dụ được nêu ra như sau: Một người tới trung tâm để xét nghiệm Covid-19, bác sĩ nói: “Nếu bạn bị nhiễm Virus thì kết quả xét nghiệm dương tính là 98%”. Sau khi xét nghiệm, người được xét nghiệm đó có kết quả dương tính và bác sĩ nói: “Xin chia buồn, nhưng tôi chắc chắn là bạn đã nhiễm Virus vì xác suất là 98%”. Điều này có đúng không? Để trả lời câu hỏi này, chúng ta phải hiểu rằng 98% là xác suất bệnh nhân sẽ có kết quả dương tính khi người đó đã nhiễm

Virus, tức là  $\Pr(P|I)$ . Giả sử như vào ngày 07 tháng 09 năm 2021, có 21.943 người bị nhiễm Virus tại Việt Nam trên tổng dân số là 98.058.068 [5]. Điều này có nghĩa là:

$$\Pr(I) = \frac{21943}{98058068} = 0.000223776 \text{ và}$$

$\Pr(NI) = 0.999776224$ . Ngoài ra, xác suất bệnh nhân bị nhiễm Virus cho ra kết quả 98% tức là  $\Pr(P|I) = \Pr(N|NI) = 0.98$ , nên  $\Pr(P|NI) = 1 - \Pr(N|NI) = 0.02$ . Do đó, dựa vào định lý Bayes, xác suất hậu nghiệm (Posterior Probability) mà bệnh nhân bị nhiễm bệnh khi kết quả xét nghiệm dương tính được định nghĩa như sau [4]:

$$\Pr(I|P) = \frac{\Pr(P|I) \times \Pr(I)}{\Pr(I) \times \Pr(P|I) + \Pr(NI) \times \Pr(P|NI)} = 0.0108 \quad (1)$$

Như vậy, người bác sĩ ở trên đã nói chưa chính xác, khi người bệnh nhân nhận kết quả dương tính thì ông nên nói là: “Trước khi xét nghiệm thì bạn có 0.02238% khả năng bị nhiễm Virus và sau khi nhận kết quả xét nghiệm dương tính thì khả năng bị nhiễm Virus của bạn chỉ tăng lên 1.08% tại Việt Nam”.

Trong một trường hợp khác, giả sử xác suất dương tính thật là 0.7 và âm tính thật là 0.95, nếu có 2 bệnh nhân A (không có triệu chứng)

và B (có triệu chứng ho và mất vị giác) đều xét nghiệm âm tính. Vậy xác suất âm tính giả là bao nhiêu? Dựa trên các thông tin này, ta có  $\Pr(I) = 0.000223776$ ,  $\Pr(NI) = 0.999776224$ ,  $\Pr(P|I) = 0.7$ ,  $\Pr(N|NI) = 0.95$ ,  $\Pr(P|NI) = 0.05$ ,  $\Pr(N|I) = 0.3$ . Do đó, xác suất mà bệnh nhân A thực sự bị nhiễm khi kết quả xét nghiệm âm tính là [6]:

$$\Pr(I|N) = \frac{\Pr(N|I) \times \Pr(I)}{\Pr(I) \times \Pr(N|I) + \Pr(NI) \times \Pr(N|NI)} = 0.00007 \quad (2)$$

Như vậy, nếu ngưỡng không nhiễm Coronavirus là 1/10000 thì có thể khẳng định rằng bệnh nhân A không nhiễm Coronavirus. Đối với bệnh nhân B với các triệu chứng nhiễm Virus và theo khảo sát xác suất tiên nghiệm tại Việt Nam là  $\Pr(I) = 0.8$  thì xác suất mà bệnh nhân B thực sự bị nhiễm khi kết quả xét nghiệm âm tính là  $\Pr(I|N) \approx 0.56$ , nghĩa là khả năng bệnh nhân B có kết quả xét nghiệm âm tính giả lên đến 56%.

Dựa vào 2 ví dụ trên, ta thấy rằng để có được sự đánh giá chính xác bệnh nhân có thực sự bị nhiễm Virus hay không thì chúng ta cần cải thiện các kết quả xét nghiệm. Có rất nhiều cách để cải thiện như cải thiện quá trình PCR, liên kết các dữ liệu từ nhiều nguồn, ... Trong đó, việc tăng số lần xét nghiệm là cách dễ dàng để có được sự đánh giá chính xác khả năng bệnh nhân có thật sự bị nhiễm Virus hay không. Và câu hỏi đặt ra là bao nhiêu lần xét nghiệm

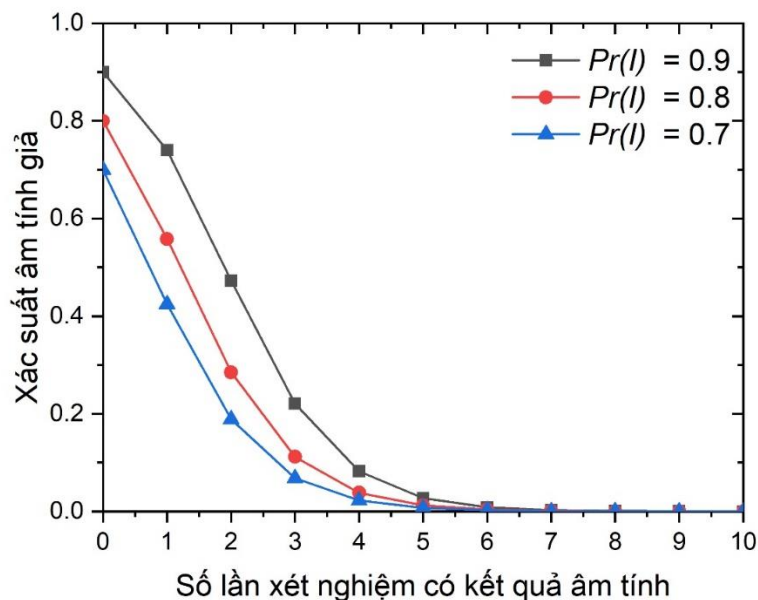
liên tiếp sẽ cho ra kết quả chính xác. Ở đây, xác suất bệnh nhân thật sự bị nhiễm Virus sau khi đã thực hiện  $n$  xét nghiệm và có kết quả  $n$  lần

liên tiếp là âm tính, tức là  $\Pr[I|N(n)]$ , được tính toán như sau [6]:

$$\Pr[I|N(n)] = \frac{\Pr(I) \times [\Pr(N|I)]^n}{\Pr(I) \times [\Pr(N|I)]^n + \Pr(NI) \times [\Pr(N|NI)]^n} \quad (3)$$

Từ công thức (3), mối tương quan của số lần xét nghiệm có kết quả âm tính và xác suất âm tính giả sau khi xét nghiệm âm tính  $n$  lần liên tiếp được minh họa ở Hình 2. Dễ dàng nhận thấy rằng, với xác suất tiên nghiệm là 0.9, 0.8, và 0.7, nếu một bệnh nhân có kết quả xét nghiệm âm tính 3 lần liên tiếp thì khả năng người đó có thể bị nhiễm bệnh lần lượt là 22.083%, 11.187%, và 6.845%. Điều đó có thể giải thích lý do vì sao bệnh nhân ở Hà Nam sau

khi có kết quả 3 lần liên tiếp âm tính thì vẫn bị nhiễm Virus. Cũng ở trong hình này, chúng ta cũng có thể nhìn thấy rằng sau 5 lần xét nghiệm âm tính liên tiếp thì khả năng người được xét nghiệm có khả năng âm tính giả theo xác suất tiên nghiệm 0.9, 0.8, và 0.7 lần lượt là 2.749%, 1.241%, và 0.727%. Tức là để dự đoán được xác suất bệnh nhân nhiễm bệnh dưới ngưỡng 3% thì phải có kết quả âm tính 5 lần liên tiếp.



Hình 2. Xác suất âm tính giả sau khi xét nghiệm âm tính  $n$  lần liên tiếp

### 3. Kết luận

Trong bài báo này, một phương pháp tiếp cận dựa trên kỹ thuật tính xác suất trong dự đoán các ca nhiễm Covid Virus được trình bày. Cụ thể, một số khái niệm được định nghĩa để áp dụng định lý Bayes nhằm tính ra xác suất các ca âm tính giả. Ngoài ra, dựa trên phương pháp tính xác suất, số lần xét nghiệm được khuyến cáo sử dụng để đạt được xác suất âm tính giả

dưới ngưỡng cho phép nhằm hạn chế việc bệnh nhân tiềm ẩn bị nhiễm Virus SARS-COV-2 đi ra xã hội khi mà kết quả xét nghiệm là âm tính. Tuy nhiên, trên đây chỉ là một góc nhìn kỹ thuật dựa trên xác suất, để có được kết quả chính xác thì vẫn còn phụ thuộc vào nhiều yếu tố khác như thời điểm thu mẫu hay khoảng cách giữa các lần thu mẫu [7].

## Tài liệu tham khảo

- [1] <https://www.24h.com.vn/tin-tuc-trong-ngay/trua-9-7-them-603-ca-mac-covid-19-trong-nuoc-c46a1268986.html>
- [2] [https://danang.gov.vn/chi-tiet?id=43380&\\_c=100000150,3,9](https://danang.gov.vn/chi-tiet?id=43380&_c=100000150,3,9)
- [3] Mai HK, Trieu NB, Long TH. (2021). Long-Term Humoral Immune Response in Persons with Asymptomatic or Mild SARS-CoV-2 Infection, Vietnam. *Emerg Infect Dis.*;27(2):663-666. doi:10.3201/eid2702.204226.
- [4] Z. Liu, Z. Zhang, S. Liu, Y. Ke and J. Chen (2011). A Trust Model Based on Bayes Theorem in WSNs. 7th International Conference on Wireless Communications, Networking and Mobile Computing, 2011, pp. 1-4, doi: 10.1109/wicom.2011.6040294.
- [5] <https://danso.org/viet-nam/>
- [6] Des McLernon. (2021). An introductory tutorial on testing for sars-cov-2\* virus from an engineering perspective. <https://www.youtube.com/watch?v=yxT1ciCnHM>
- [7] Sethuraman N, Jeremiah SS, Ryo A. (2020) Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA.* 323(22):2249–2251. doi:10.1001/jama.2020.8259