Original Article

JHSMR Journal of Health Science and Medical Research

Clinical Prediction Model of Long COVID During the Delta and the Omicron Variant Dominant Waves in Thailand

Chonlawat Chaichan, M.Sc.[1,2], Sirinda Sritipsukho[3], Sasinuch Rutjanawech, M.D.[4], Paskorn Sritipsukho, M.D., M.Sc.[5,6], Thammanard Charernboon, M.D., Ph.D.[1,5,7]

[1]Department of Clinical Epidemiology, Faculty of Medicine, Thammasat University, Khlong Luang, Pathumthani 12120, Thailand.
[2]Department of Research and Medical Innovation, Faculty of Medicine, Vajira Hospital, Navamindradhiraj University, Dusit, Bangkok 10300, Thailand.
[3]Medical Student Program, Faculty of Medicine, Chulalongkorn University, Khlong Luang, Bangkok 12120, Thailand.
[4]Department of Internal Medicine, Faculty of Medicine, Thammasat University, Khlong Luang, Pathumthani 12120, Thailand.
[5]Center of Excellence in Applied Epidemiology, Thammasat University, Khlong Luang, Pathumthani 12120, Thailand.
[6]Research Group in Clinical Epidemiology, Faculty of Medicine, Thammasat University, Khlong Luang, Pathumthani 12120, Thailand.
[7]Department of Clinical Epidemiology, Faculty of Medicine, Thammasat University, Khlong Luang, Pathumthani 12120, Thailand.

Received 10 April 2025 ● Revised 27 May 2025 ● Accepted 11 June 2025 ● Published online 17 October 2025

Abstract:

Objective: Long coronavirus disease (long COVID) represents a significant burden on healthcare systems and requires enhanced management strategies. There is a critical need for more comprehensive care and targeted healthcare services for affected populations. This study aimed to develop a clinical prediction scoring system for long COVID in patients recovering from COVID-19.

Material and Methods: This prospective cohort study collected data at Thammasat University Hospital and the Thammasat Field Hospital during the Delta- and Omicron-variant-dominant epidemics. Phone interviews regarding long COVID symptoms were conducted with 2516 patients at 3 months post-infection. A stepwise logistic regression model was employed to develop the final predictive model for long COVID.

Results: In total, 40.46% of patients exhibited long COVID symptoms 3 months after infection. Our model comprised 5 predictors: dyspnea, healthcare worker status, female gender, severity of acute illness, and variant dominant wave. With a sensitivity of 57.1% and a specificity of 67.3% at 3 months, the risk score exhibited an area under the receiver operating characteristic curve of 0.62 for long COVID prediction. The probability of long COVID for each risk score point

Contact: Thammanard Charernboon, M.D., Ph.D.
Faculty of Medicine, Thammasat University, Pathumthani 12120, Thailand.
E-mail: dr.thammanard@gmail.com

J Health Sci Med Res 2026;44(4):e20251262
doi: 10.31584/jhsmr.20251262
www.jhsmr.org

was also reported. The Hosmer–Lemeshow test (p-value=0.49) indicated good model calibration, with closely aligned observed and expected frequencies.

Conclusion: The predictive risk score demonstrated satisfactory accuracy in identifying COVID-19 patients at high risk of developing long COVID 3 months post-infection.

Keywords: clinical prediction model, Delta, long COVID, Omicron, Thailand

Introduction

The prevalence of long COVID at 3 months based on a pooled estimation was approximately 42%[1,2], and it continues to burden patients during their recovery, with severity and duration varying by individual factors. Certain patients with long-lasting COVID-19 exhibited only moderate symptoms or were asymptomatic throughout the acute phase of the infection; this resulted in the delayed onset of some symptoms by several weeks up to 2 years[3,4]. Despite the declining emergency status of COVID-19, physicians and clinical researchers are actively updating their knowledge and practices regarding long COVID management. Fatigue, cognitive dysfunction, dyspnea, and sleep disturbances are among the most frequently reported symptoms of long COVID, with fatigue and neurocognitive impairment often persisting for over 6 months in a substantial proportion of patients[5-7]. Prolonged symptom duration can significantly impair physical functioning, occupational performance, and psychosocial well-being, particularly in working-age individuals. These long-term effects pose a considerable burden on healthcare systems and societal productivity[8,9]. At present, no disease-specific treatment has been established, and current management strategies remain largely supportive[10]. Considering these challenges, preventive measures including vaccination, early identification of high-risk individuals, and mitigation of reinfection are essential to reduce the incidence and duration of long COVID[11].

Previous studies, systematic reviews, and meta-analyses had a follow-up at approximately 4 months, collecting 49 symptoms that met the inclusion criteria for long COVID. This study reported common systemic symptoms of long COVID, including fatigue and weakness, pain, and neurological, gastrointestinal, and cardiopulmonary symptoms, along with mental health issues. Long COVID follow-up times varied widely depending on individual studies[2].

Being female was the highest impact predictor for long COVID. A meta-analysis revealed that women experienced cell cycle process changes, immune dysregulation, and histone modifications[12]. Additional variables, including comorbidities and older age, remain controversial. Nevertheless, ethnicity served as a predictor in numerous studies on long COVID. Studies from different countries included a range of district predictors that varied across studies, which could be attributed to differences in each country's population, which inevitably had an impact on the prediction model[13].

Patients after survival to hospital discharge or recovery from COVID-19 have been frequently found to have long COVID, especially in cases of moderate-to-severe acute illness. Most such patients were reported to have long COVID[14]. In contrast, individuals with asymptomatic or mild acute infection have not demonstrated a consistent association between initial illness severity and the duration of long COVID[15]. Recovery in these cases may be either brief or prolonged, and the severity of the acute phase alone does not reliably predict the course or duration of long COVID. Some studies have described this presentation as delayed symptom onset[16].

Prediction models for long COVID are currently limited, with most studies originating from Western countries and lacking standardized case definitions and outcome measures[17,18]. Differences in symptom presentation, prevalence, geography, and culture across populations highlight the need for localized prediction models—particularly in underrepresented settings such as Thailand, where research on long COVID remains scarce[19,20]. In response to this gap, our study aimed to develop a simple and practical clinical prediction model to identify patients at risk of long COVID 3 months after recovery. This model may support more effective screening and clinical decision-making, especially in primary care settings or resource-limited environments.

Material and Methods

Setting

A prospective cohort study was conducted at Thammasat University Hospital (TUH) and its affiliate, the Thammasat Field Hospital (TFH), to examine the frequency and features of long COVID in patients in Thailand from May 2021 to June 2022, when the Delta and Omicron variants were most prevalent. These were crucial periods in the local spread of the disease. TUH, a tertiary-care academic hospital with a capacity of 700 beds, caters to a wide population in northern Bangkok and central Thailand. TFH, with 490 beds, was established to provide medical services to patients with asymptomatic or mild COVID-19. This ensured that patients at all levels of severity received thorough and complete care. All subjects gave informed consent for participation on their first visit.

Participants

The study population included adults aged ≥18 years who visited TUH or TFH and tested positive for SARS-CoV-2 via nasopharyngeal RT–PCR. Individuals with positive RT–PCR results were approached for consent to participate in the study, resulting in a total of 2,516 participants.

They were interviewed by telephone 3 months after their COVID-19 diagnosis. The participants included 1,018 individuals diagnosed during the Delta-dominant wave and 1,498 individuals from the Omicron-dominant wave. Participants were COVID-19 patients aged ≥18 years diagnosed with COVID-19 via a positive nasopharyngeal Real-time Polymerase Chain Reaction (RT-PCR) test for SARS-CoV-2 at TUH and TFH. Regarding the representativeness of the sampling, we approached all patients who tested positive for SARS-CoV-2 via RT-PCR during the study period and invited them to participate based on informed consent. However, we did not have data on the specific COVID-19 variant or sub-lineage for each case. Based on data from the Thai Ministry of Public Health, the Delta variant predominated from July to December 2021, while Omicron became dominant from January 2022 onward. National surveillance reported Delta in 88%–100% of cases through mid-December 2021, with Omicron rising to 94%–100% by mid-January 2022. This trend aligns with our findings, where Omicron was identified in 97.8% of 363 adult cases (BA.1: 64.8%, BA.2: 35.2%)[21]. Exclusion criteria included death or loss to follow-up within 3 months, inability to communicate in Thai, refusal to provide phone consent, or being unreachable by phone.

Patients with long COVID and data collection

Our study utilized 2 data sources. First, demographic parameters, comorbidities, history of COVID-19 immunization, and pertinent clinical data, including gender, age, smoking history, highest level of education, vaccination history (including number of doses), and comorbidities, such as cardiovascular diseases (CVD), chronic lung diseases, diabetes mellitus, chronic kidney disease (CKD),

stroke, and cancer, were obtained for each participant from the TUH database. The classification of disease severity followed the guidelines set by the National Institutes of Health (NIH), categorizing cases as asymptomatic, mild, moderate, severe, or critical. Treatment for all patients adhered to standard protocols recommended by the Department of Medical Services, Ministry of Public Health (MOPH). We confirmed each patient's vaccination status before COVID-19 infection using the MOPH immunization database. Full vaccination was defined as receiving at least 2 doses, following Thai governmental guidelines.

Data on acute COVID-19 severity and symptoms were recorded. Symptoms were categorized as follows: respiratory (e.g., cough, sore throat, rhinorrhea, sputum production, dyspnea), neurological (headache), musculoskeletal (myalgia), gastrointestinal (diarrhea), ear, nose, and throat (loss of smell, loss of taste), and dermatological (rash, red eye).

Second, due to the delay in the release of the ICD-10 with the code U09.9 for long COVID for standardized use, a symptom-based structured questionnaire created by the Thai Ministry of Public Health was used to assess long COVID through telephone interviews. The question asks whether the symptoms that the respondents observed in themselves were ongoing or newly occurring in order to exclude symptoms that might not be related to long COVID. Research assistants were trained in patient interviewing to ensure consistent interviews. The timing of the long COVID interviews was the same for both periods—at 3 months post-infection—during an ongoing emergency situation. The questionnaire covered symptoms: dyspnea, cough, chest tightness, palpitation, headache, attention deficit, memory loss, insomnia, diarrhea, myalgia, arthralgia, dizziness, lack of appetite, loss of smell, loss of taste, rash, alopecia, depression, stress, exhaustion, and weakness.

Statistical analysis

Common long COVID symptoms were selectively classified per systematic reviews and meta-analyses. The 13 symptoms compiled included fatigue and weakness; neurological symptoms included headache, attention deficit, memory loss, and insomnia; cardiopulmonary symptoms included dyspnea, cough, chest constriction, and palpitations; gastrointestinal symptoms included diarrhea; and mental-health symptoms included depression and stress.

Patient characteristics were described using frequencies and percentages, and those with and without common long COVID were compared using the chi-square test. All potential predictors for the prediction score were identified by comparing patients with and without long COVID at 3 months. All predictors were included in stepwise backward logistic regression models using a p-value≤0.1 as the cutoff. In multivariable analysis, all variables were classified into binaries. A simple point system was developed for the score based on the coefficients from the final logistic model. The lowest beta coefficient was used to standardize and scale the other coefficients, assigning point values proportionally.

After assigning scores, a cut-off point for differentiating the risk of long COVID at 3 months was determined using the Liu method. This statistical approach identifies the optimal threshold by maximizing the product of sensitivity and specificity, thereby balancing true positive and true negative rates[22]. The prediction accuracy of the model was evaluated using the area under the receiver operating characteristic curve (AuROC), sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratio for a positive test. The calibration was assessed by calibration plots, and goodness-of-fit tests were performed to compare observed and predicted probabilities. The analyses were conducted using STATA version 14.0 (StataCorp, College Station, TX, USA) [StataCorp, 2015].

Ethical considerations

This study was approved by the Ethics Committee of Thammasat University (Medicine, MTU-EC-PE-1-332/64).

Results

Patient characteristics

A total of 2,516 patients were followed up 3 months after contracting COVID-19 during the Delta- and Omicron-dominant waves. Of these, 40.46% (1018 patients) had long COVID as defined by the NICE guidelines, and 70.33% of these long COVID patients were female. As seen in Table 1, significant characteristics included female gender, being a healthcare worker, severe and critical acute illness, less than 2 doses of the vaccine, and symptoms of acute illness, including cough, sore throat, and dyspnea.

Model development

In Table 2, we present the multivariable logistic regression model of all the predictors. The final backward logistic regression model is presented in Table 3. The final model includes 10 predictors: female sex, obesity with body mass index (BMI) ≥25 kg/m^2, occupation in the healthcare sector, severe and critical symptoms during acute illness, the Omicron-dominant infection wave, cough, myalgia, dyspnea, and loss of smell during acute illness. Of these characteristics, female gender, the Omicron-dominant wave, and occupation in the healthcare sector appeared to have the highest predictive power, with odds ratios of 2.29, 2.05, and 1.64, respectively.

The total score for the predictive model was 10.5 points from 10 predictors. We calculated the summary risk score by summing the scores of all 5 items (Table 3). The 10.5-point total score includes 2.4 points for being female, 1.0 point for being a healthcare worker, 3.4 points for experiencing severe or critical illness during the acute phase, 2.3 points for being infected during the Omicron-dominant wave, and 1.4 points for having dyspnea during acute illness. For example, if a female patient was infected with COVID-19 during the Omicron-dominant wave and experienced severe illness during the acute phase, the cumulative risk score for long COVID would be 2.4 (female) + 2.3 (Omicron) + 3.4 (severe/critical illness), resulting in a total of 8.1 points.

Model validation & calibration

The total score of 10.5 points was divided by a 2.85-point threshold. Individuals with a score greater than or equal to 2.85 were categorized as having long COVID. The AuROC for the final regression model was 0.62. The model's sensitivity and specificity were 57.1% and 67.3%, respectively. Additionally, the positive predictive value was 54.2%, and the negative predictive value was 69.7% (Table 4).

A calibration plot was graphed between the predicted probability and observed probability. The plot revealed a chi-square correlation coefficient of 4.40. The item score for each variable was determined by dividing its regression coefficients by the lowest coefficient value (0.33) of the model and rounding the result to one decimal place (Figure 1).

The risk scores were plotted against the actual proportions of long COVID at each individual risk score point (Figure 1). The calculated likelihood of the risk score closely matched the observed patterns in the prevalence of long COVID, indicating that the risk score was well calibrated. Moreover, the goodness of fit between observed and predicted probabilities was acceptably correlated with p-value 0.4936.

Discussion

Our study developed a clinical prediction scoring system for Long COVID in patients recovering from COVID-19. The strengths of our research include the

Table 1 Baseline characteristics of monitoring common long COVID at 3 months during the Delta and the Omicron dominant variant waves

	Total (N=2,516)	Long COVID (N=1,018)	No long COVID (N=1,498)	p-value[a]
Variant of concern				<0.001*
Delta	1128 (44.83)	345 (30.59)	783 (69.41)	
Omicron	1388 (55.17)	673 (48.49)	715 (51.51)	
Sex – no (%)	(100%)			<0.001*
Female	1892 (41.02)	716 (70.33)	768 (51.27)	
Male	2681 (58.98)	302 (29.67)	730 (48.73)	
Age – no (%)				0.984
Older 60 years	386 (15.34)	156 (15.32)	230 (15.35)	
Under 60	2130 (84.66)	862 (84.68)	1268 (84.65)	
Occupation – no (%)				
Healthcare	268 (10.65)	144 (14.15)	124 (8.28)	<0.001*
Obesity (BMI ≥25 kg/m^2)	975 (38.86)	403 (39.70)	572 (38.29)	0.475
Education level – no (%)				0.882
Primary & secondary	1370 (54.78)	552 (54.60)	818 (54.90)	
Bachelor& postgraduate	1131 (45.22)	459 (45.40)	672 (45.10)	
Comorbidities – no (%)				
Cardiovascular disease	410 (16.30)	169 (16.60)	241 (16.09)	0.732
Chronic lung disease	96 (3.82)	41 (4.03)	55 (3.67)	0.647
Diabetes mellitus	237 (9.42)	142 (9.48)	95 (9.33)	0.901
Chonic kidney disease	51 (2.03)	23 (2.26)	28 (1.87)	0.565
Stroke	40 (1.59)	21 (2.06)	19 (1.27)	0.143
Cancer	42 (1.67)	19 (1.87)	23 (1.54)	0.530
Severity at acute illness				<0.001*
Asymptomatic & mild	2068 (82.23)	813 (79.86)	1255 (83.83)	
Moderate	350 (13.92)	145 (14.24)	205 (13.69)	
Severe & critical	97 (3.86)	60 (5.89)	37 (2.47)	
Vaccine history				<0.001*
2 doses	883 (35.10)	297 (29.17)	586 (39.12)	
More than 2 doses	1633 (64.90)	721 (70.83)	912 (60.88)	
Common symptoms at acute illness				
Cough	1329 (52.82)	567 (55.70)	762 (50.87)	0.017*
Sore throat	1121 (44.55)	490 (48.13)	631 (42.12)	0.003*
Myalgia	214 (8.51)	98 (9.63)	116 (7.74)	0.097
Rhinorrhea	519 (20.63)	219 (21.51)	300 (20.03)	0.366
Sputum production	253 (10.06)	104 (10.22)	149 (9.95)	0.825
Dyspnea	183 (7.27)	183 (9.14)	90 (6.01)	0.003*
Headache	455 (19.08)	196 (19.25)	259 (17.29)	0.209
Diarrhea	55 (2.19)	25 (2.46)	30 (2.00)	0.446
Loss of smell	168 (6.68)	62 (6.09)	106 (7.08)	0.331
Loss of taste	70 (2.78)	27 (2.65)	43 (2.87)	0.744

*Statistical significance, Data are numbers (%), BMI=body mass index, COVID=coronavirus disease
[a]Comparison between long COVID and no long COVID

Table 2 Full model of long COVID by multivariable logistic regression model

	AOR [95% CI]	p-value[a]
Female sex	2.25 (1.88–2.68)	<0.001*
Age ≥60 years	0.85 (0.64–1.13)	0.260
Healthcare worker	1.26 (0.96–1.72)	0.090
Obesity (BMI ≥25 kg/m^2)	1.16 (0.97–1.39)	0.095
Education level		
Primary & secondary	1.20 (1.00–1.44)	0.047*
Bachelor& postgraduate	Ref	
Vaccine more than 2 doses	1.25 (0.99–1.57)	0.066
Omicron variant dominant wave	2.11 (1.69–2.64)	<0.001*
Cardiovascular diseases (CVD)	0.97 (0.74–1.29)	0.852
Chronic lung diseases	0.93 (0.59–1.46)	0.741
Diabetes mellitus	0.86 (0.62–1.19)	0.352
Chronic kidney diseases (CKD)	0.94 (0.49–1.79)	0.848
Stroke	1.34 (0.66–2.74)	0.417
Cancer	1.09 (0.54–2.20)	0.800
Severity at acute illness		
Asymptomatic & mild	Ref	
Moderate	1.29 (1.00–1.67)	0.050*
Severe & critical	3.71 (2.27–6.08)	<0.001*
Symptoms during acute illness		
Cough	1.09 (0.91–1.30)	0.350
Sore throat	1.05 (0.88–1.26)	0.577
Myalgia	1.32 (0.98–1.78)	0.072
Rhinorrhea	1.05 (0.85–1.30)	0.626
Sputum production	0.82 (0.61–1.08)	0.161
Dyspnea	1.49 (1.07–2.08)	0.019*
Headache	1.06 (0.85–1.32)	0.627
Diarrhea	1.13 (0.64–2.01)	0.669
Loss of smell	1.12 (0.75–1.66)	0.582
Loss of taste	1.28 (0.71–2.30)	0.407

Adjusted odds ratio at 3 months by female sex, age, healthcare worker, obesity, education level, vaccine doses, dominant variant wave, CVD, chronic lung diseases, diabetes mellitus, chronic kidney diseases, stroke, cancer, severity and symptoms during acute illness
AOR=Adjusted odds ratio by multivariable logistic regression
[a]Comparison between common long COVID and no common long COVID at 3 months

utilization of RT-PCR to diagnose all patients, resulting in a large sample size for predicting long COVID at 3 months, covering both the Delta and Omicron variant waves in Thailand. The predictive model demonstrated that our 10.5-point risk score was both moderately sensitive and valid in predicting the risk of developing common long COVID 3 months after infection.

The predictors included in our risk score are largely consistent with those identified in previous studies as significant risk factors for long COVID. In the medical literature, female gender, severe and critical acute illness, dyspnea, and severe acute illness are frequently cited as risk factors[23-25]. Severe and critical acute illness was the most significant risk factor in our risk score, receiving the highest possible point value of 3.4. Consequently, the interpretation is more straightforward for COVID-19 patients who experience severe or critical acute illnesses. Individuals with additional risk factors are more susceptible to developing long COVID. However, our study found that the Omicron variant was associated with a higher risk of long COVID compared to the Delta variant, whereas the majority of studies report the opposite[26]. The hypothesis regarding the prevalence of long COVID and its underlying biological mechanisms remains unclear. Interestingly, our findings are consistent with those from Maharat Nakhon Ratchasima Hospital in Thailand, which also reported a higher prevalence of long COVID during the Omicron wave compared to the Delta wave. They attributed this to greater genetic variation in the Omicron spike protein, particularly within the receptor-binding domain and receptor-binding motif, which may enhance transmissibility and immune evasion. Furthermore, differences in the types and doses of vaccines used in Thailand may complicate direct comparisons with international data[9].

Our risk score could be beneficial for screening COVID-19 patients at high risk for common long COVID at 3 months. At-risk patients can then be prioritized for close follow-up, and innovative interventions can then be developed for them.

Previous studies of prediction models for long COVID using statistical and machine learning techniques have

Table 3 Final multivariable logistic regression model with long COVID as the outcome variable

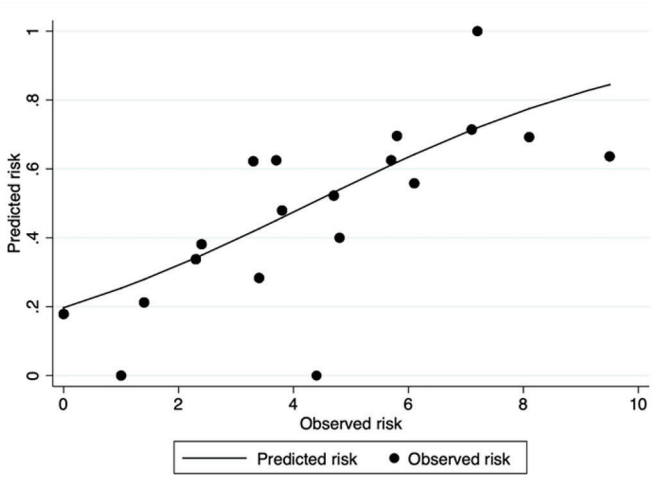
	Regression Coefficient	Odds ratio	95% CI of odds ratio	p-value	Assigned score
Sex					
Male	Ref				+0
Female	0.80	2.23	1.87–2.65	<0.001	+2.4
Occupation					
Non-healthcare	Ref				+0
Healthcare	0.33	1.39	1.06–1.81	0.017	+1.0
Severity at acute illness					
Asymptomatic & mild	Ref				+0
Severe & critical	1.12	3.07	1.94–4.87	<0.001	+3.4
Variant of concern					
Delta	Ref				+0
Omicron	0.76	2.14	1.80–2.55	<0.001	+2.3
Dyspnea symptoms					
No	Ref				+0
Yes	0.45	1.56	1.13–2.16	0.007	+1.4

Pseudo R^2=0.06, AuROC=0.62

Table 4 Accuracy of the risk score

Cutoff point	AuROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)
≥2.85	0.62 (0.60–0.64)	57.1 (54.0–60.1)	67.3 (64.8–69.6)	54.2 (51.2–57.3)	69.7 (67.3–72.1)	1.74 (1.59–1.91)

AuROC=area under receiver operating characteristic curve, PPV=positive predictive value, NPV=negative predictive value, LR+=likelihood ratio if test positive



Calibration plot of observed vs predicted risk for developing common long COVID; Hosmer–Lemeshow chi-square test on modeling showed the goodness of fit for logistic regression analysis (X^2=4.40, p-value=0.494).

Figure 1 Calibration plot (observed probability versus predicted probability) for long COVID

identified different predictors[27-31]. Antony et al. developed a predictive model from electronic health records to predict long COVID by machine learning. Their model included 8 predictors: age, female gender, cough, fatigue, albuterol, obesity, diabetes, and chronic lung diseases. The AuROC was 0.76 and 0.75 as per a logistic regression and random forest model, respectively. Moreover, Kessler et al. developed a model using a gradient-boosting classifier from machine learning. It had the highest recall score at 72%, with a specificity of 80% in a test data set separated from the total data set (20%). The model collected 14 predictors including dominant variant waves, physician practice, age, diagnostic and treatment management, length of stay, sex, vaccine history, somatoform disorders, migraine, back pain, asthma, malaise, fatigue, and cough[28]. In contrast, Honchar et al. found that worse physical function during acute illness was not associated with long COVID. They developed and reported a model with 7 predictors comprising age, sex, CRP levels in-hospital, eGFR, need for oxygen supplementation, symptoms after discharge with assessment using the 6-Minute Walk Test, and Medical Research Council dyspnea score[32]. Although the AUROC of 0.62 indicates limited discriminatory performance, and the sensitivity (57.1%) and specificity (67.3%) reflect modest accuracy, these values are not uncommon in early prediction models for complex, heterogeneous conditions like long COVID. It is important to note that this model was developed using real-world data during an emergency setting, where limitations in data quality and availability were necessary. Nonetheless, the model provides a foundation for identifying potential risk factors, including novel predictors such as healthcare-worker status. With further refinement, such as incorporating additional validated variables, using larger or more diverse datasets, and applying more advanced modeling techniques, its predictive accuracy may be improved for practical clinical use.

These models were similar to our model in terms of sex, severity, and dyspnea during acute illness as predictors. Being female was a particularly strong predictor, which our model rated 1.4 out of a total score of 10.5. Our model also identified being a healthcare worker as an important predictor in the final model, with 1.0 point of the total 10.5-point score. A systematic review discusses vulnerable workers in a healthcare setting[15]. Myalgia during acute illness was another predictor of long COVID. Kessler et al. reported that the Omicron variant predicted long COVID. Although age in our study was not a predictor in the model, it might be variably categorized such that older adults are not a discrete variable. Myalgia during acute illness was another predictor of long COVID. Kessler et al. also reported that the Omicron variant predicted long COVID.

Our study has several limitations. First, the predictive model might have been improved if additional data had been collected, such as other comorbidities (e.g., asthma), laboratory results, type of vaccination, access to care, and treatments received during the acute phase of illness. Second, due to the emergency epidemic situation, our physicians were unable to conduct face-to-face follow-up visits to clinically confirm the presence of long COVID symptoms or to obtain laboratory results. This limitation may have introduced potential inaccuracies or inconsistencies in identifying and classifying long COVID cases. To minimize data collection errors, however, we provided thorough training to research assistants involved in administering the questionnaires and managing follow-up communications. Additionally, the questionnaire used to assess long COVID was not a formally validated diagnostic instrument and relied entirely on patients' self-reported information without clinical corroboration. This approach may have led to the inclusion of vague or unrelated symptoms not directly associated with long COVID. Third, the study did not include biomarker measurements to support the prediction model—such as

C-reactive protein and serum cytokines (e.g., IFN, IL-6, IL-10, IL-1, and TNF)—despite many studies having attempted to use these biomarkers to predict long COVID. Lastly, we did not have data on the sub-lineages of the SARS-CoV-2 variants.

Conclusion

The predictive risk score exhibited acceptable accuracy in identifying long COVID and effectively identifying individuals at a high risk of developing long COVID 3 months after infection. These results may help inform patient management by predicting long COVID.

Acknowledgement

This study received support from the Health Systems Research Institute (Grant No. 65-050) and the Faculty of Medicine, Thammasat University. We used ChatGPT for language refinement. The authors are solely responsible for the content and interpretations presented in this paper.

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Woodrow M, Carey C, Ziauddeen N, Thomas R, Akrami A, Lutje V, et al. Systematic review of the prevalence of long COVID. Open Forum Infect Dis 2023;10:ofad233.
2. Natarajan A, Shetty A, Delanerolle G, Zeng Y, Zhang Y, Raymont V, et al. A systematic review and meta-analysis of long COVID symptoms. Syst Rev 2023;12:88.
3. Demko ZO, Yu T, Mullapudi SK, Varela Heslin MG, Dorsey CA, Payton CB, et al. Two-year longitudinal study reveals that long COVID symptoms peak and quality of life nadirs at 6–12 months postinfection. Open Forum Infect Dis 2024;11:ofae027.
4. Kim Y, Bae S, Chang HH, Kim SW. Long COVID prevalence and impact on quality of life 2 years after acute COVID-19. Sci Rep 2023;13:11207.
5. Lippi G, Sanchis-Gomar F, Henry BM. COVID-19 and its long-term sequelae: what do we know in 2023? Pol Arch Intern Med 2023;133.
6. Almeria M, Cejudo JC, Deus J, Krupinski J. Long neurocognitive and neuropsychiatric sequelae in participants with post-COVID-19 Infection: a longitudinal study. Neurol Int 2024;16:853-68.
7. Schild AK, Scharfenberg D, Regorius A, Klein K, Kirchner L, Yasemin G, et al. Six-month follow-up of multidomain cognitive impairment in non-hospitalized individuals with post-COVID-19 syndrome. Eur Arch Psychiatry Clin Neurosci 2024;274:1945-57.
8. Müller K, Poppele I, Ottiger M, Wastlhuber A, Weber RC, Stegbauer M, et al. Long-term course and factors influencing work ability and return to work in post-COVID patients 12 months after inpatient rehabilitation. J Occup Med Toxicol 2024;19:43.
9. Xiang J, Zheng H, Cai Y, Chen S, Wang Y, Chen R. Cumulative social disadvantage and its impact on long COVID: insights from a U.S. national survey. BMC Med 2025;23:207.
10. Cheng AL, Herman E, Abramoff B, Anderson JR, Azola A, Baratta JM, et al. Multidisciplinary collaborative guidance on the assessment and treatment of patients with Long COVID: A compendium statement. PM R 2025;17:684-708.
11. Deforth M, Gebhard CE, Bengs S, Buehler PK, Schuepbach RA, Zinkernagel AS, et al. Development and validation of a prognostic model for the early identification of COVID-19 patients at risk of developing common long COVID symptoms. Diagn Progn Res 2022;6:22.
12. Rusu EC, Monfort-Lanzas P, Bertran L, Barrientos-Riosalido A, Solé E, Mahmoudian R, et al. Towards understanding post-COVID-19 condition: a systematic meta-analysis of transcriptomic alterations with sex-specific insights. Comput Biol Med 2024;175:108507.
13. Mkoma GF, Agyemang C, Benfield T, Rostila M, Cederström A, Petersen JH, et al. Risk of long COVID and associated symptoms after acute SARS-COV-2 infection in ethnic minorities: A nationwide register-linked cohort study in Denmark. PLoS Med 2024;21:e1004280.

14. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis. EClinicalMedicine 2023;55:101762.
15. Goel N, Goyal N, Spalgais S, Mrigpuri P, Varma-Basil M, Khanna M, et al. Initial COVID-19 severity and long-COVID manifestations: an observational analysis. Thorac Res Pract 2023;24:22-8.
16. van Kessel SAM, Olde Hartman TC, Lucassen P, van Jaarsveld CHM. Post-acute and long-COVID-19 symptoms in patients with mild diseases: a systematic review. Fam Pract 2022;39:159-67.
17. Bai J, Wang J. Modeling long COVID dynamics: impact of underlying health conditions. J Theor Biol 2024;576:111669.
18. Chou R, Herman E, Ahmed A, Anderson J, Selph S, Dana T, et al. Long COVID definitions and models of care: a scoping review. Ann Intern Med 2024;177:929-40.
19. Krongthaeo S, Partiprajak S, Piaseu N, Ckumdee S, Taaon C, Kongsuwan A. Prevalence and exploratory factor analysis of long COVID-19 symptoms among experienced infected population in Bangkok, Thailand. BMC Public Health 2024;24:2863.
20. Rutjanawech S, Sritipsukho P, Sritipsukho S. Long COVID at 3 and 6 months after covid-19 infection in Thailand. Sci Rep 2024;14:23441.
21. Petcharat S, Supataragul A, Hirunpatrawong P, Torvorapanit P, Klungthong C, Chinnawirotpisan P, et al. High transmission rates of early omicron subvariant BA.2 in Bangkok, Thailand. Adv Virol 2023;2023:4940767.
22. Hoque MA, Kibria BMG. Performance of some estimators for the multicollinear logistic regression model: theory, simulation, and applications. Res Stat 2024;2:2364747.
23. Bai F, Tomasoni D, Falcinella C, Barbanotti D, Castoldi R, Mulè G, et al. Female gender is associated with long COVID syndrome: a prospective cohort study. Clin Microbiol Infect 2022;28:611.e9-16.
24. Vimercati L, De Maria L, Quarato M, Caputi A, Gesualdo L, Migliore G, et al. Association between long COVID and overweight/obesity. J Clin Med 2021;10.
25. Jassat W, Mudara C, Vika C, Welch R, Arendse T, Dryden M, et al. A cohort study of post COVID-19 condition across the Beta, Delta and Omicron waves in South Africa: 6-month follow up of hospitalised and non-hospitalised participants. Int J Infect Dis 2022.
26. Du M, Ma Y, Deng J, Liu M, Liu J. Comparison of long COVID-19 caused by different SARS-CoV-2 strains: a systematic review and meta-analysis. Int J Environ Res Public Health 2022;19.
27. Jin W, Hao W, Shi X, Fritsche LG, Salvatore M, Admon AJ, et al. Using multi-modal electronic health record data for the development and validation of risk prediction models for long COVID using the super learner algorithm. J Clin Med 2023;12.
28. Kessler R, Philipp J, Wilfer J, Kostev K. Predictive attributes for developing long COVID-a study using machine learning and real-world data from primary care physicians in Germany. J Clin Med 2023;12.
29. Nair P, Nair CV, Kulirankal KG, Corley EM, Edathadathil F, Gutjahr G, et al. Characterization and predictive risk scoring of long COVID in a south indian cohort after breakthrough COVID infection; a prospective single centre study. BMC Infect Dis 2023;23:670.
30. Socia D, Larie D, Feuerwerker S, An G, Cockrell C. Prediction of long COVID based on severity of initial COVID-19 infection: differences in predictive feature sets between hospitalized versus non-hospitalized index infections. medRxiv 2023.
31. Myska V, Genzor S, Mezina A, Burget R, Mizera J, Stybnar M, et al. Artificial-intelligence-driven algorithms for predicting response to corticosteroid treatment in patients with post-acute COVID-19. Diagnostics (Basel) 2023;13.
32. Honchar O, Ashcheulova T, Chumachenko T, Chumachenko D, Bobeiko A, Blazhko V, et al. A prognostic model and pre-discharge predictors of post-COVID-19 syndrome after hospitalization for SARS-CoV-2 infection. Front Public Health 2023;11:1276211.