A Practical Clinical Score Predicting Respiratory Failure in COVID-19 Patients

Moshe Ashkenazi MD MBA[1,6]*, Eyal Zimlichman MD[2,6]*, Noa Zamstein PhD[5], Galia Rahav MD[3,6], Reut Kassif Lerner MD[1,6], Yael Haviv MD[4,6], and Itai M. Pessach MD PhD MPH[1,6]

[1]Safra Children's Hospital, Sheba Medical Center, [2]Central Management, [3]Department of Infectious Disease, and [4]Intensive Care Unit, Sheba Medical Center, Tel Hashomer, Israel
[5]MDClone LTD, Beer Sheva, Israel
[6]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

ABSTRACT **Background:** The coronavirus disease 2019 (COVID-19) pandemic resulted in repeated surges of patients, sometimes challenging triage protocols and appropriate control of patient flow. Available models, such as the National Early Warning Score (NEWS), have shown significant limitations. Still, they are used by some centers to triage COVID-19 patients due to the lack of better tools.

Objectives: To establish a practical and automated triage tool based on readily available clinical data to rapidly determine a distinction between patients who are prone to respiratory failure.

Methods: The electronic medical records of COVID-19 patients admitted to the Sheba Medical Center March–April 2020 were analyzed. Population data extraction and exploration were conducted using a MDClone (Israel) big data platform. Patients were divided into three groups: non-intubated, intubated within 24 hours, and intubated after 24 hours. The NEWS and our model where applied to all three groups and a best fit prediction model for the prediction of respiratory failure was established.

Results: The cohort included 385 patients, 42 of whom were eventually intubated, 15 within 24 hours or less. The NEWS score was significantly lower for the non-intubated patients compared to the two other groups. Our improved model, which included NEWS elements combined with other clinical data elements, showed significantly better performance. The model's receiver operating characteristic curve had area under curve (AUC) of 0.92 with of sensitivity 0.81, specificity 0.89, and negative predictive value (NPV) 98.4% compared to AUC of 0.63 with NEWS. As patients deteriorate and require further support with supplemental O2, the need for re-triage emerges. Our model was able to identify those patients on supplementary O2 prone to respiratory failure with an AUC of 0.86 sensitivity 0.95, and specificity 0.7 NPV 98.9%, whereas NEWS had an AUC of 0.76. For both groups positive predictive value was approximately 35%.

Conclusions: Our model, based on readily available and simple clinical parameters, showed an excellent ability to predict negative outcome among patients with COVID-19 and therefore might be used as an initial screening tool for patient triage in emergency departments and other COVID-19 specific areas of the hospital.

IMAJ 2022; 24: 327–331

KEY WORDS: coronavirus disease 2019 (COVID-19), mechanical ventilation, National Early Warning Score (NEWS), respiratory failure

*These authors contributed equally to this study

The coronavirus disease 2019 (COVID-19) outbreak was characterized as a public health emergency of international concern by the World Health Organization on 30 January 2020, and on 11 March 2020 as a pandemic [1]. The disease spread from China to the rest of the world infecting around 11.5 million and killing more than 530,000 individuals as of 5 July 2020. The novel coronavirus was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the coronavirus study group [2]. The presenting clinical symptoms were mostly fever, dry cough, and fatigue [3], whereas the COVID-19 patients had severe clinical course such as greater severity of pneumonia and markedly higher rate of acute respiratory distress syndrome (ARDS) compared to seasonal flu [4,5] and mortality rate of 11–15% among hospitalized patients [6,7]. However, the main challenge is the high patient density. As a result, early identification and management of severe patients became a critical need of the healthcare system.

The pandemic caused some countries to experience scenarios of an ongoing mass casualty incident (MCI) with limited medical resources. In Italy, up to 60–70% of the patients presenting to the emergency department (ED) were admitted [8]. Due to the severe clinical course of COVID-19 patients, there was an unprecedented load on the healthcare system. In any MCI, ethical decisions have to be considered to allocate the scarce resources to the patients who will profit the most. Better understanding of the disease and the ability to set prognostic factors are crucial in this scenario [9,10].

Clinical scoring is a well-known method that tries to quantify patient's status and to categorize it into risk groups. Some of the most established scores are those who created to identify early sepsis, e.g the quick Sepsis-related Organ Failure Assessment (qSOFA), Systemic Inflammatory Response Syndrome (SIRS), Early Warning Score (EWS) and its derivatives, National EWS (NEWS), and Modified EWS (MEWS). In a study published in 2018, a comparison between qSOFA, SIRS, and NEWS revealed that NEWS was the most accurate scoring system for the detection of sepsis among adults [11]. NEWS was also shown to be a good predictor to detect early clinical deterioration 24 hours after discharge from the intensive care unit (ICU) [12]. An additional study found NEWS superior to qSOFA in pre-

dicting early, unplanned escalation of care for ED patients [13]. However, as long-term mortality predictors, NEWS and MEWS had poor predictive powers among visitors in ED [14]. The literature is scarce with studies implementing those clinical scores in COVID-19 patients. There was a report from Sichuan, China, about using EWS in those patients [15]. The NEWS score was introduced by the Royal College of Physicians of London and is based on respiratory rate, oxygen saturation, supplemental oxygen, temperature, systolic blood pressure, and heart rate. Unlike scores of ICU patients, the NEWS does not include laboratory values and is based only on clinical parameters. We tried to strengthen the NEWS with additional laboratory values that were described as distinctive in COVID-19 patients [2] as lactate dehydrogenase (LDH), D-Dimer, C-reactive protein (CRP), total protein, lymphocyte count, and background diseases such as diabetes and hypertension.

Few studies have been conducted to analyze models that predict respiratory failure among COVID-19 patients. One of them [16] used machine learning to predict respiratory decompensation among 197 COVID-19 patients. The results were promising, showing higher sensitivity and specificity in comparison with MEWS. Liang et al. [17] developed a clinical risk score that achieved promising result of area under curve (AUC) 0.88 and which included independent predictive factors as hemoptysis, history of cancer, age, dyspnea, and more. The information was collected from data of 1590 COVID-19 patients. Outcome was death, mechanical ventilation, or admission to the ICU. Furthermore, an additional study [18] analyzing data from 208 patients described a novel scoring model named as CALL (co-morbidity, age, lymphocytes, and LDH), which is composed of co-morbidities, age, lymphocyte count, and LDH, showed AUC of 0.91.

The aim of our study was to explore the strength of our new score in predicting deterioration to mechanical ventilation among COVID-19 patients.

PATIENTS AND METHODS

A retrospective case study was performed, collecting data from electronic medical records from both an acute care academic tertiary medical center (Sheba Medical Center, Israel) and from a national health fund (Maccabi Healthcare Services). The study was approved by the local ethics committee (approval SMC-20-7112). No informed consent was needed. Inclusion criteria were patients who were hospitalized and were positive to SARS-CoV-2 in all different hospitalization wards at Sheba Medical Center for the time period from 19 February until 11 May 2020. The research data were obtained using MDClone's big data platform, which allows fast and flexible extraction of desired cohorts and their characteristics (MDClone LTD, Beer Sheva, Israel). The parameters extracted for the study were age, gender, co-morbid conditions, clinical and laboratory parameters, and clinical course. The patients were divided into three different groups: those who arrived intubated or were intubated in the first 24 hours, those who were intubated after the first 24 hours, and those who were not intubated at all. For additional analysis we created a new subgroup; those who needed supplementary oxygen during the hospitalization. Outcome parameter was intubation and mechanical ventilation.

The NEWS was calculated for all patients (respiratory rate, oxygen saturation, supplemental oxygen, temperature, systolic blood pressure, and heart rate). To strengthen our model, additional variables were collected including LDH, CRP, total protein, calcium, and co-morbid conditions such as diabetes and hypertension. The additional parameters were not integrated in the NEWS but were added to the model and calculated as categorical parameter. Points were given when exceeding the cut-off value. In addition, for the building of the model, the NEWS score was taken as a categorical variable, with 10 as the cut-off between normal and abnormal. For the patients who were intubated, the worst score from the 12–48 hours preceding the intubation was recorded. For those who were not intubated, the worst calculated score was taken from the all hospitalization. For predicting ventilation after respiratory support, the data was collected up to 24 hours after the onset of respiratory support.

Candidate predictors based on univariable analysis were incorporated into a multivariable logistic regression using stepwise forward and backward selection [Table 2, Table 3]. Then, multivariate logistic regression models were performed in order to predict mechanical ventilation out of the entire cohort, and the fraction of the cohort who had prior respiratory support, respectively. Due to the relatively small size of the cohorts, we avoided dividing them into train and test sets. Last, regarding the NEWS score, to compare the three groups we constructed ANOVA-Welch test with a post-hoc Games-Howell test to

Table 1: Patient data by status of ventilation during hospitalization

	Ventilated > 24 hours, n=27	Non-ventilated, n=343	*P* value
Average age in years	63.2 ± 14.3	54.4 ± 20.18	0.027
Males (%)	21 (77.8)	193 (56.3)	0.008
National Early Warning Score	9.81 ± 3.3	7.5 ± 4.45	0.008
Co-morbidities			
Arterial hypertension (%)	15 (55)	114 (33)	0.019
Diabetes mellitus (%)	9 (33.3)	74 (21.6)	0.158
Laboratory results			
Protein (g/dl)	6.23 ± 0.54	7.13 ± 0.58	< 0.001
Lactate dehydrogenase (IU/l)	610 ± 170	383 ± 182	< 0.001
C-reactive protein (g/l)	164 ± 91	93 ± 89	< 0.001
Calcium (mg/dl)	8.3 ± 0.37	9.1 ± 0.6	< 0.001

Table 2: Predictors of ventilation from multivariable logistic regression Analysis for the whole patients' population

Variable	Odds ratio (95% confidence interval)	*P* value
NEWS (abnormal vs. normal)	42 (7.4–237)	0.00002
Arterial hypertension (yes vs. no)	9.8 (1.55–61.8)	0.015
Diabetes mellitus (yes vs. no)	0.037 (0.0007–1.87)	0.1
Protein (maximal value)	0.94 (0.19–4.6)	0.93
Calcium (maximal value)	2.6 (0.37–18.2)	0.33
LDH (maximal value)	1.87 (0.17–20)	0.61
NEWS × arterial hypertension	0.01 (0.001–0.11)	0.00018
Maximal protein × maximal LDH	128 (6.3- 2586)	0.0015
Maximal NEWS × maximal D-dimer	0.056 (0.007–0.47)	0.008
Arterial hypertension × diabetes mellitus	4.2 (0.2–87)	0.35
Arterial hypertension × maximal calcium	2.2 (0.17–28)	0.54
Arterial hypertension × maximal LDH	0.06 (0.003–1.17)	0.063
diabetes mellitus × maximal CRP	5.7 (0.12–286)	0.38
Arterial hypertension × maximal CRP	2.04 (0.08–50.8)	0.66
Intercept	0.009	

CRP = C-reactive protein, LDH = lactate dehydrogenase, NEWS = National Early Warning Score

Table 3: Predictors of ventilation from multivariable logistic regression analysis for patients who needed supplementary oxygen

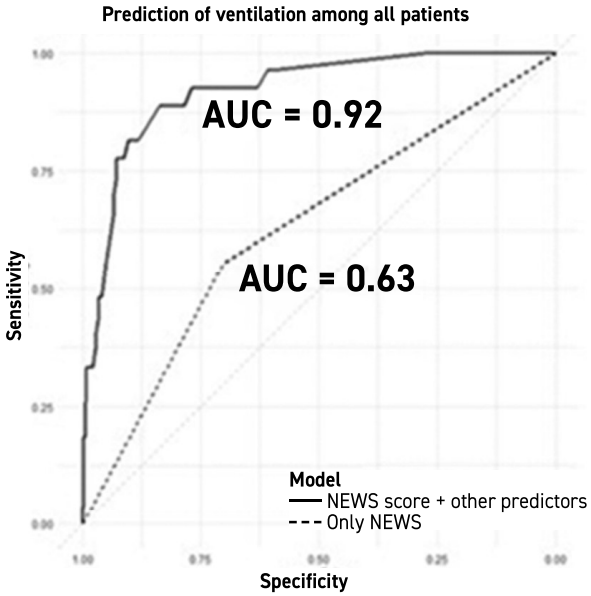
Variable	Odds ratio (95% confidence interval)	*P* value
NEWS (abnormal vs. normal)	147 (7–3032)	0.0012
LDH (maximal value)	57.6 (4.5–740)	0.0018
Arterial hypertension (yes vs. no)	4.98 (0.35–71)	0.24
Diabetes mellitus (yes vs. no)	0.98 (0.32–2.95)	0.97
Maximal NEWS × maximal LDH	0.04 (0.002–0.74)	0.03
Maximal NEWS × arterial hypertension	0.088 (0.005–1.6)	0.1
Intercept	0.0004	

LDH = lactate dehydrogenase, NEWS = National Early Warning Score

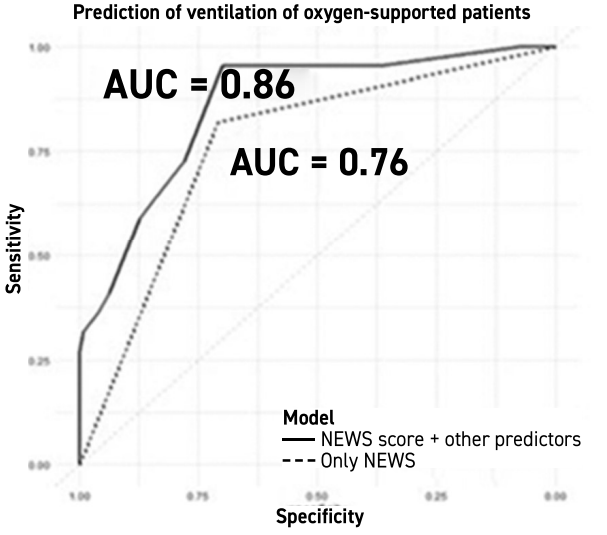
say which pairs of groups are different. Sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) were added to our score for two groups: total patients and those who needed supplementary oxygen.

Figure 1: Receiver operating characteristic curve for patients

[A] All patients



[B] For patients needed supplementary oxygen



RESULTS

We collected data from 385 patients (227 males), with an average age of 55.2 ± 19.8 years [Table 1]. Out of the total group, 42 (10.9%) were eventually intubated, 15 (35.7%) in the first 24 hours and 27 (64.3%) during the rest of the hospitalization; 127 patients (33%) needed supplementary oxygen and were not ventilated during the hospitalization. The calculated odds ratio of men for ventilation was 3.23 (95% confidence interval [95%CI] 1.455–7.194, *P* value = 0.004).

The predictors of ventilation from multivariable logistic regression analysis are show in Table 2(for all study's population) and Table 3 (for patients needed supplementary oxygen).

The NEWS for the three different groups was the highest for those who were ventilated in the first 24 hours (10.9 ± 2.53) and the lowest for those who were not ventilated [Table 1]. The NEWS for both groups of ventilated patients (< 24 hours from admission, > 24 hours) was significantly higher than the non-ventilated group.

Our initial goal was to predict the chances of patient deterioration due to respiratory failure. Confusion matrix for all patients reveals 306 patients who were true negative for our model, 22 who were true positive, and number of patients with false prediction was 37 and 5 (positive and negative, respectively). The calculated sensitivity was 0.81, specificity 0.89, NPV 98.4%, and PPV 37.3 %.

As the clinical course of COVID-19 associated illness is dynamic and sometimes progressive, the continuous monitoring of the patients and their appropriate allocation to different levels of care is crucial. After the patient has been admitted, the next significant triage point in the clinical course of COVID-19 patents is once supplemental oxygen therapy is initiated. Hence, we implemented our model also on patients needing supplementary oxygen. In this group of 149 patients, the calculated sensitivity for the prediction of respiratory failure was 0.95 with the specificity of 0.7. The NPV and PPV were 98.9% and 35.6%, respectively.

In practice, among those 149 patients only 1 patient (1.1%) who developed respiratory failure was missed by the model. However, 38 of the 59 patients (64%) predicted to develop respiratory failure actually needing intubation.

The receiver operating characteristic (ROC) curves of our model, compared to the NEWS score, are shown in Figure 1A. The results showed an AUC of 0.92 for our model and 0.63 for the NEWS score for the whole population. Similarly, for the patient on supplemental oxygen [Figure 1B] an AUC of 0.86 Vs 0.76 were observed for our model in comparison to the MEWS score in predicting respiratory failure.

DISCUSSION

We have developed a new model for the prediction of respiratory failure among COVID-19 patients. In light of its performance, the model might be used as a risk stratification and triage tool for the ED and clinical service bedside provider: low score ensures in high probability that the patient is not going to require mechanical ventilation at least in the next 12 hours. In a situation of limited medical resources, a predictive tool like this, with very high NPV, might help to allocate resources wisely. Furthermore, compared to other newly developed models [17,18] our model shows higher AUC. The comparison between our model to the CALL [18] is interesting; CALL includes age, co-morbidities (extended than ours) and LDH, whereas ours includes bigger variety of laboratory values than the CALL. The constant parameters (age, co-morbidities) constitute more than 50% of the points in CALL; therefore, the risk group has bigger weight in CALL score.

Five patients were categorized as false negatives. A careful analysis of these five patients revealed that LDH levels were not measured in these cases, significantly limiting the model's performance.

In surge capacity conditions the need for on-going triage and proper allocation of patients to the appropriate level of care is key. Hence, once a patient starts to deteriorate and needs increased support (i.e., supplemental oxygen) the bedside provider has to quickly decide the appropriate placement. Our model's sensitivity was also high, and higher than NEWS in this group of patients, while maintaining a similar NPV.

The odds ratio for men to be ventilated was high and statistically significant; however, after performing logistic regression, the male sex was not determined to be a risk factor. We attribute the finding to the higher rate of co-morbidity (arterial hypertension and diabetes mellitus) among males.

The low PPV is due to a significant false positive fraction; however, in a screening tool aimed at predicting deterioration it is reasonable to allow some false alerts (meaning patients will be monitored further) as long as the false negative fraction is minimal (so possibly deteriorating patients are not missed).

The NEWS score itself had good predictive ability [Figure 1, Figure 2, Table 1]. It is based on clinical parameters only (no laboratory data are needed) and hence can be immediately calculated at the bedside. However, our model requires only minimal laboratory data that is taken from the basic panels that are almost universally acquired for every hospitalized COVID-19 patient, allowing for a significantly more robust performance.

There are a number of limitations to our study that deserve attention. The primary one being that it is a retrospective, single center study on a relatively small group of patients and thus, the study is exposed to possible confounders and cannot be generalized to other hospitals.

CONCLUSIONS

We developed a model that is based on basic clinical and laboratory parameters that may allow for ongoing automatic triage of patients during pandemic surge conditions. The use of such tools may play an important role in patient flow management under these conditions. Further validation and studies on a wider population are suggested.

Correspondence

Dr. M. Ashkenazi
Safra Children's Hospital, Sheba Medical Center, Tel Hashomer 5262000, Israel
Phone: (972-3) 530-2884
Fax: (972-3) 530-9334
email: moshe.ashkenazi@sheba.health.gov.il

References

1. Siordia JA Jr. Epidemiology and clinical features of COVID-19: a review of current literature [published online ahead of print, 2020 Apr 10]. *J Clin Virol* 2020; 127: 104357.
2. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* 2020; 91: 264-6.
3. Wang Y, Lu X, Li Y, et al. Clinical course and outcomes of 344 intensive care patients with COVID-19. *Am J Respir Crit Care Med* 2020; 201 (11): 1430-4.
4. Cevik M, Bamford CGG, Ho A. COVID-19 pandemic-a focused review for clinicians. *Clin Microbiol Infect* 2020; 26 (7): 842-7.
5. Tolksdorf K, Buda S, Schuler E, Wieler LH, Haas W. Influenza-associated pneumonia as reference to assess seriousness of coronavirus disease (COVID-19). *Euro Surveill* 2020; 25 (11): 2000258.
6. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395 (10223): 497-506.
7. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020; 395 (10223): 507-13.
8. Giwa AL, Desai A, Duca A. Novel 2019 coronavirus SARS-CoV-2 (COVID-19): an overview for emergency clinicians. *PediatrEmerg Med Pract* 2020; 17 (5): 1-24.
9. Kirkpatrick JN, Hull SC, Fedson S, Mullen B, Goodlin SJ. Scarce-resource allocation and patient triage during the COVID-19 pandemic: JACC review topic of the week. *J Am Coll Cardiol* 2020; 76 (1): 85-92.
10. Rello J, Storti E, Belliato M, Serrano R. Clinical phenotypes of SARS-CoV-2: implications for clinicians and researchers. *Eur Respir J* 2020; 55 (5): 2001028.
11. Usman OA, Usman AA, Ward MA. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the Emergency Department. *Am J Emerg Med* 2019; 37 (8): 1490-7.
12. Uppanisakorn S, Bhurayanontachai R, Boonyarat J, Kaewpradit J. National Early Warning Score (NEWS) at ICU discharge can predict early clinical deterioration after ICU transfer. *J Crit Care* 2018; 43: 225-9.
13. Sutherland ME, Yarmis SJ, Lemkin DL, Winters ME, Dezman ZDW. National Early Warning Score is modestly predictive of care escalation after emergency department-to-floor admission. *J Emerg Med* 2020; 58 (6): 882-91.
14. Ahn JH, Jung YK, Lee JR, et al. Predictive powers of the Modified Early Warning Score and the National Early Warning Score in general ward patients who activated the medical emergency team. *PLoS One* 2020; 15 (5): e0233078.
15. Liao X, Wang B, Kang Y. Novel coronavirus infection during the 2019-2020 epidemic: preparing intensive care units-the experience in Sichuan Province, China. *Intensive Care Med* 2020; 46 (2): 357-60.
16. Burdick H, Lam C, Mataraso S, et al. Prediction of respiratory decompensation in COVID-19 patients using machine learning: The READY trial. *Comput Biol Med* 2020; 124: 103949.
17. Liang W, Liang H, Ou L, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients With COVID-19. *JAMA Intern Med* 2020; 180 (8): 1-9.
18. Ji D, Zhang D, Xu J, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL Score. *Clin Infect Dis* 2020; 71 (6): 1393-9.

Capsule

KIR-full goodbye to autoimmunity

Ly49+CD8+ T cells are a subset of CD8+ T cells that show immunoregulatory activity in mice. **Li** et al. report the existence of a similar CD8+ T cell subset in humans that expresses killer cell immunoglobulin-like receptors (KIRs), a functional parallel of the mouse Ly49 family. These cells, which can suppress self-reactive CD4+ T cells, were more abundant in patients with autoimmune conditions such as celiac disease, multiple sclerosis, and lupus, as well as in patients infected with influenza virus or severe acute respiratory syndrome coronavirus 2. When mice selectively deficient in Ly49+CD8+ T cells were infected with viruses, they showed normal antiviral immune responses but eventually developed symptoms of autoimmune disease. KIR+CD8+ T cells may therefore be an important therapeutic target for the control of autoimmune diseases such as long COVID that emerge after viral infections.

Science 2022; 376: 265
Eitan Israeli

Capsule

Bat coronaviruses related to SARS-CoV-2 and infectious for human cells

SARS-CoV-2 progenitor bat viruses genetically close to SARS-CoV-2 and able to enter human cells through a human ACE2 (hACE2) pathway have not yet been identified, although they would be key in understanding the origin of the epidemic. **Temman** and colleagues showed that such viruses circulate in cave bats living in the limestone karstic terrain in northern Laos, in the Indochinese peninsula. The authors found that the receptor-binding domains of these viruses differ from that of SARS-CoV-2 by only one or two residues at the interface with ACE2, bind more efficiently to the hACE2 protein than that of the SARS-CoV-2 strain isolated in Wuhan from early human cases, and mediate hACE2-dependent entry and replication in human cells, which is inhibited by antibodies that neutralize SARS-CoV-2. None of these bat viruses contains a furin cleavage site in the spike protein. These findings therefore indicate that bat-borne SARS-CoV-2-like viruses that are potentially infectious for humans circulate in *Rhinolophus* spp. in the Indochinese peninsula.

Nature 2022; 604: 330
Eitan Israeli