

ONLINE FIRST PUBLICATION

Online first papers have undergone full scientific review and copyediting, but have not been typeset or proofread. To cite this article, use the DOIs number provided. Mandatory typesetting and proofreading will commence with regular print and online publication of the online first papers of the *SMJ*.

A risk prediction score to identify patients at low risk for COVID-19 infection

Wui Mei Chew¹, MB BCh BAO, MRCP, Chee Hong Loh¹, MB BCh BAO, Aditi Jalali², MBBS, Grace Shi En Fong², MBBS, Loshini Senthil Kumar², MBBS, Rachel Hui Zhen Sim², MBBS, Russell Pinxue Tan², MBBS, Sunil Ravinder Gill², MBBS, Trilene Ruiting Liang², MBBS, Jansen Meng Kwang Koh¹, MBBS, FRCPE, Tunn Ren Tay¹, MBBS, MRCP

¹Department of Respiratory and Critical Care Medicine, ²Department of General Medicine, Changi General Hospital, Singapore

Correspondence: Dr Wui Mei Chew, Associate Consultant, Level 6 Medical Centre, Department of Respiratory and Critical Care Medicine, Changi General Hospital, 2 Simei Street 3, Singapore 529889. chew.wui.mei@gmail.com

Singapore Med J 2021, 1–18

<https://doi.org/10.11622/smedj.2021019>

Published ahead of print: 12 March 2021

Online version can be found at
<http://www.smj.org.sg/online-first>

ABSTRACT

Introduction: Singapore's enhanced surveillance programme for COVID-19 identifies and isolates hospitalised patients with acute respiratory symptoms to prevent nosocomial spread. We developed risk prediction models to identify patients with low risk for COVID-19 from this cohort of hospitalised patients with acute respiratory symptoms.

Methods: This was a single-centre retrospective observational study. Patients admitted to our institution's respiratory surveillance wards from 10 February to 30 April 2020 contributed data for analysis. Prediction models for COVID-19 were derived from a training cohort using variables based on demographics, clinical symptoms, exposure risks and blood investigations fitted into logistic regression models. The derived prediction models were subsequently validated on a test cohort.

Results: Of the 1,228 patients analysed, 52 (4.2%) were diagnosed with COVID-19. Two prediction models were derived, the first based on age, presence of sore throat, dormitory residence, blood haemoglobin level (Hb), and total white blood cell counts (TW), and the second based on presence of headache, contact with infective patients, Hb and TW. Both models had good diagnostic performance with areas under the receiver operating characteristic curve of 0.934 and 0.866, respectively. Risk score cut-offs of 0.6 for Model 1 and 0.2 for Model 2 had 100% sensitivity, allowing identification of patients with low risk for COVID-19. Limiting COVID-19 screening to only elevated-risk patients reduced the number of isolation days for surveillance patients by up to 41.7% and COVID-19 swab testing by up to 41.0%.

Conclusion: Prediction models derived from our study were able to identify patients at low risk for COVID-19 and rationalise resource utilisation.

Keywords: COVID-19 infection, infection control, respiratory infections

INTRODUCTION

Since its emergence in December 2019, the spread of COVID-19 has been rapid, with global numbers surpassing 100 million cases as of 9 February 2021.⁽¹⁾ Given its high infectivity, healthcare systems worldwide are under immense pressure to expediently identify and isolate COVID-19 cases to curb its spread. At the start of the COVID-19 outbreak when cases originated from a limited number of affected regions, history of travel to affected regions or close contact with a COVID-19 case proved useful for case detection.⁽²⁾ However, as local transmissions intensified, previously important epidemiological risk factors became increasingly inadequate.⁽³⁾ Compounding the problem, the clinical spectrum of COVID-19 resembles many common respiratory illnesses, making accurate identification of cases even more challenging. Some healthcare systems, such as Singapore's, have turned to widespread screening of patients with acute respiratory symptoms as a means to contain the virus.

Since early February 2020, Singapore's Ministry of Health required all public hospitals in Singapore to set up surveillance wards in order to cohort and screen all patients who present with pneumonia and acute respiratory symptoms for COVID-19.⁽⁴⁾ Although this 'leave no stone unturned' approach aids in containment, it is highly resource intensive and may not be sustainable in the long run.⁽⁵⁾ In addition to requiring large quantities of tests for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), dedicated hospital wards and even isolation facilities are occupied by such screening exercises.⁽⁵⁾

Prediction tools to stratify patients according to their risk of having COVID-19 would enable healthcare systems to target screening efforts at high-risk patients only, dispensing with testing for low-risk patients. This balances the need for early disease containment with judicious resource utilisation. However, few prediction tools are currently available.⁽⁶⁾ This study aims to develop a prediction model to identify patients who are at low risk of having COVID-19 to better guide resource allocation.

METHODS

This was a retrospective observational study conducted at Changi General Hospital, a university-affiliated 1,000-bed hospital located in the eastern region of Singapore.

In response to the government's call for enhanced surveillance, all patients with pneumonia and symptoms of acute respiratory infection (ARI) who did not fulfil the Ministry of Health's official suspect case criteria (based largely on travel history to high-risk regions and COVID-19 contact; Appendix) were admitted to our hospital's Pneumonia and Acute Respiratory Infection (PARI) wards. The admission criteria were patients with pneumonia (with or without ARI symptoms) and patients with only ARI symptoms. These wards consisted of single-occupancy rooms with en-suite bathrooms. Patients fulfilling the official suspect case criteria were admitted to a separate isolation ward.

All patients admitted to our PARI wards from 10 February 2020 to 30 April 2020 were included in the study. For patients with multiple admissions during the study period, analysis was limited to the first PARI admission. Patients admitted to the PARI wards underwent two SARS-CoV-2 polymerase chain reaction (PCR) tests with combined nasal and oropharyngeal swabs on two consecutive days. Positive patients were transferred out to separate wards designated for COVID-19 patients after confirmation of positive PCR testing. Patients that tested negative were transferred to non-isolation wards or discharged.

Data on demographics, clinical symptoms, exposure risks, vital parameters, blood investigations and radiological investigations was collected as part of our hospital's audit process. A positive contact history referred to self-reported contact with someone who had symptoms of ARI but no direct COVID-19 contact, as the latter group would be admitted to separate isolation wards. We included residence in a migrant workers' dormitory (i.e. dormitory residence) as an exposure risk factor in our analysis, because COVID-19 infections in Singapore's migrant worker community rose significantly during the study period. The spike

in cases was largely attributed to living in close quarters in these dormitories. Study analyses were performed on de-identified datasets.

Positive clinical symptoms and exposure risks indicated in the medical records were recorded as 'present' in our dataset, while negative symptoms and risk factors or those that were not mentioned in the medical records were recorded as 'absent'. Clinical variables were selected from features of respiratory tract infection. The first set of vital parameter readings available from the emergency department or ward was used for analysis. Blood and radiological investigations were based on the first available results within 24 hours of hospital admission.

Only patients with available SARS-CoV-2 PCR test results were included for analysis. Missing values were encountered only for blood investigations, and these values were imputed using multiple imputation techniques. Continuous variables were compared using *t*-test or Mann-Whitney *U* test as appropriate, while categorical variables were compared using chi-square tests. A *p*-value < 0.05 was taken to be statistically significant. All statistical analyses were performed using IBM SPSS Statistics version 24.0 (IBM Corp, Armonk, NY, USA).

A subset of patients (60%) were randomly selected from the study population to form the training set for model development, while the remainder (40%) formed the testing set for model validation. Candidate predictors were selected after considering (a) univariate analyses between COVID-19 positive and negative cases; (b) predictors identified in prevailing literature and data available in our study population;⁽⁶⁾ and (c) clinical relevance. Variables that showed multicollinearity were excluded. To achieve a more parsimonious model, candidate predictors were entered into a logistic regression model, and only variables with a *p*-value < 0.1 were retained in the final model. The predicted risk score of being COVID-19 positive was $100 * 1/(1 + e^{-z})$, where *z* was the derived value from the prediction model equation.

The performance of the model was assessed by measuring the area under the receiver operating characteristic (ROC) curve. An area under the curve (AUC) of 0.7–0.8 was considered acceptable, 0.8–0.9 considered excellent, and > 0.9 considered outstanding. Cut-off values for predicting COVID-19 status were determined based on Youden's index. ROC statistics were employed to establish the performance of the model for the test set. In addition, cut-off values derived from the training set were used to determine the predictive accuracy of the model for the validation cohort.

As this was a retrospective study based on de-identified patient data collected for a hospital audit, an ethics review was not required. This study complies with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis reporting guideline.⁽⁷⁾

RESULTS

A total of 1,232 unique patients were admitted to PARI wards over the study period. As SARS-CoV-2 testing was not performed in four patients, 1,228 patients contributed data for the analyses (Fig. 1). COVID-19 was diagnosed in 52 (4.2%) patients.

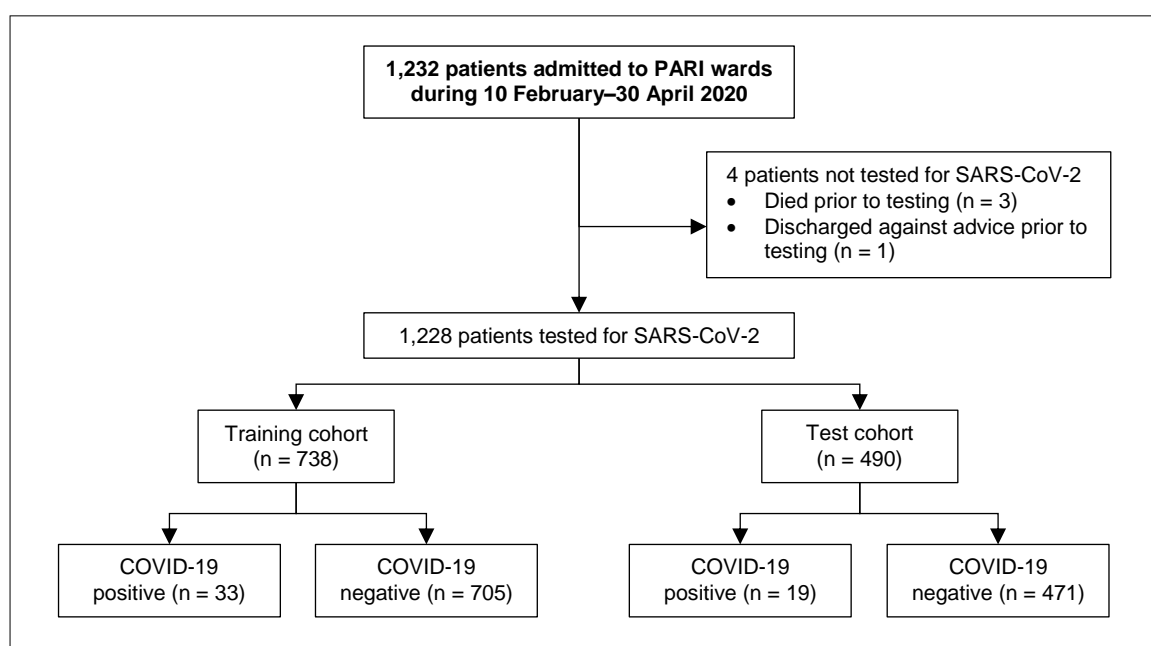


Fig. 1 Flowchart shows patient inclusion in the study. PARI: Pneumonia and Acute Respiratory Infection; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

The training set comprised 738 (60.1%) patients and the test set comprised 490 patients (39.9%). Their baseline characteristics are shown in Table I. 33 (4.5%) patients in the training set were diagnosed with COVID-19. They were more likely to be male, non-Chinese and younger than patients without COVID-19. Comparisons of the clinical characteristics of patients with and without COVID-19 are shown in Table II.

Table I. Baseline characteristics of the study cohort.

Characteristic	No. (%)		
	Total (n = 1,228)	Derivation cohort (n = 738)	Validation cohort (n = 490)
Age (yr)	63 ± 20	63 ± 20	62 ± 20
Male gender	735 (59.9)	444 (60.2)	291 (59.4)
Ethnicity			
Chinese	712 (58.0)	423 (57.3)	289 (59.0)
Malay	226 (18.4)	134 (18.2)	92 (18.8)
Indian	118 (9.6)	72 (9.8)	46 (9.4)
Others	172 (14.0)	109 (14.8)	63 (12.9)
Smoking status			
Current	122 (9.9)	75 (10.2)	47 (9.6)
Never	314 (25.6)	179 (24.3)	135 (27.6)
Ex-smoker	76 (6.2)	39 (5.3)	37 (7.6)
No data	716 (58.3)	445 (60.3)	271 (55.3)
Comorbidities			
Diabetes mellitus	412 (33.6)	241 (32.7)	171 (34.9)
Hypertension	644 (52.4)	383 (51.9)	261 (53.3)
Ischaemic heart disease	266 (21.7)	155 (21.0)	111 (22.7)
Stroke	162 (13.2)	93 (12.6)	69 (14.1)
End-stage renal failure	56 (4.6)	35 (4.7)	21 (4.3)
Active cancer	66 (5.4)	39 (5.3)	27 (5.5)
Asthma	122 (9.9)	62 (8.4)	60 (12.2)
COPD	81 (6.6)	39 (5.3)	42 (8.6)
Other chronic lung disease*	79 (6.4)	44 (6.0)	35 (7.1)

*Other chronic lung diseases included bronchiectasis and diffuse parenchymal lung diseases. COPD: chronic obstructive pulmonary disease

Table II. Univariate analysis of the clinical features, exposure history, blood investigations and radiology of the training cohort.

Parameter	No. (%)			p-value
	Total (n = 738)	COVID-19 positive (n = 33)	COVID-19 negative (n = 705)	
Age* (yr)	63 ± 20	48 ± 15	64 ± 20	< 0.001
Clinical feature				
Fever	368 (49.9)	21 (63.6)	347 (49.2)	0.105
Cough	449 (60.8)	22 (66.7)	427 (60.6)	0.483
Sore throat	118 (16.0)	15 (45.5)	103 (14.6)	< 0.001
Rhinorrhoea	112 (15.2)	6 (18.2)	106 (15.0)	0.622
Anosmia	1 (0.1)	0 (0)	1 (0.1)	1
Breathlessness	211 (28.6)	4 (12.1)	207 (29.4)	0.032
Headache	32 (4.3)	7 (21.2)	25 (3.5)	< 0.001
Chest discomfort	87 (11.8)	6 (18.2)	81 (11.5)	0.264
Duration of symptom[†] (day)	3 (1–7)	3 (2–7)	3 (1–7)	0.908
Exposure history				
Contact with persons having ARI [‡]	43 (5.8)	12 (36.4)	31 (4.4)	< 0.001
Dormitory residence	42 (5.7)	19 (57.6)	23 (3.3)	< 0.001
Nursing home residence	46 (6.2)	0 (0)	46 (6.5)	0.257
Travel history in past 1 mth	23 (3.1)	1 (3.0)	22 (3.1)	1
Blood results*				
Urea (mmol/L)	7.5 ± 6.2	4.4 ± 2.2	7.6 ± 6.3	0.003
Sodium (mmol/L)	137 ± 6	137 ± 4	137 ± 6	0.808
Bicarbonate (mmol/L)	22 ± 4	23 ± 3	22 ± 4	0.051
Creatinine (µmol/L)	120 ± 132	78 ± 29	122 ± 135	0.058
ALT (U/L)	51 ± 78	59 ± 52	50 ± 79	0.548
AST (U/L)	52 ± 69	50 ± 39	52 ± 70	0.885
C-reactive protein (mg/L)	51.2 ± 67	47.3 ± 70.6	51.5 ± 66.6	0.728
Procalcitonin (µg/L)	2.54 ± 8.28	1.75 ± 3.31	2.58 ± 8.31	0.726
Haemoglobin (g/dL)	12.8 ± 2.3	14.5 ± 1.8	12.7 ± 2.3	< 0.001
White blood cell (10 ³ /µL)	10.7 ± 4.7	7.1 ± 3.3	10.9 ± 4.7	< 0.001
Platelets (10 ³ /µL)	254 ± 104	207 ± 84	256 ± 104	< 0.001
Neutrophils (10 ³ /µL)	8.22 ± 5.55	4.65 ± 2.96	8.39 ± 5.59	< 0.001
Lymphocytes (10 ³ /µL)	1.63 ± 1.30	1.46 ± 0.76	1.64 ± 1.32	0.442
Monocytes (10 ³ /µL)	0.81 ± 0.59	0.73 ± 0.36	0.82 ± 0.59	0.39
CXR suggesting infection	374 (50.7)	13 (39.4)	361 (51.2)	0.185
Pattern of CXR changes[§]				0.127
Unilateral changes	254 (67.9)	6 (46.2)	248 (68.7)	
Bilateral changes	120 (32.1)	7 (53.8)	113 (31.3)	

Data presented as *mean ± standard deviation, †median (interquartile range). ‡Contact with persons with ARI symptoms excluded those with direct COVID-19 contact. §Determination of pattern was based on the chest radiography report. ALT: alanine transaminase; ARI: acute respiratory infection; AST: aspartate transaminase; CXR: chest radiograph

Gender, ethnicity, age, sore throat, dyspnoea, headache, an ARI contact history, dormitory residence, urea level, blood haemoglobin level (Hb), total white blood cell count (TW) and platelet count were identified as candidate predictors based on univariate results. In addition, we included the presence of fever and absolute lymphocyte count as well as serum procalcitonin as candidate predictors in our model. The final predictors that were retained in our first model (Model 1) were age, presence of sore throat, dormitory residence, Hb and TW (Table III). The prediction equation (z_1) was $-5.134 + (0.016 * \text{age}) + (1.257 * \text{sore throat}) + (3.887 * \text{dorm}) + (0.216 * \text{Hb}) - (0.36 * \text{TW})$.

As dormitory residence was regarded a specific risk factor for COVID-19 in Singapore, a repeat regression modelling without dormitory residence, gender and ethnicity as candidate predictors was performed to obtain a more generalisable model. The latter two variables were excluded because they were highly associated with dormitory residence – 100% and 90% of patients from dormitory residence were males and non-Chinese, respectively. The final predictors that were retained in our second model (Model 2) were presence of headache, contact with ARI patients, Hb and TW (Table III). The prediction equation (z_2) was $-5.951 + (1.573 * \text{headache}) + (1.793 * \text{ARI contact}) + (0.385 * \text{Hb}) - (0.331 * \text{TW})$.

Table III. Predictors of COVID-19 retained in final models.

Parameter	B	OR	95% CI
Model 1			
Age	0.160	1.02	0.99–1.05
Sore throat	1.257	3.52	1.35–9.17
Dormitory residence	3.887	48.76	14.06–169.06
Hb	0.216	1.24	0.98–1.57
TW	-0.360	0.006	0–0.37
Model 2			
Headache	1.573	4.82	1.62–14.35
ARI contact	1.793	6.01	2.42–14.94
Hb	0.385	1.47	1.19–1.81
TW	-0.331	0.72	0.62–0.84

ARI: acute respiratory infection; CI: confidence interval; Hb: haemoglobin; OR: odds ratio; TW: total white blood cell count

The AUC for Model 1 was 0.934 (95% confidence interval [CI] 0.891–0.978). A risk score cut-off value of 4 gave the best diagnostic accuracy with sensitivity and specificity of 90.9% (95% CI 75.7%–98.1%) and 84.5% (95% CI 81.7%–87.1%), respectively. As the intended utility of the risk score was to identify patients with a low risk of COVID-19, a cut-off value of 0.6 provided a sensitivity of 100.0% (95% CI 89.4%–100.0%) and corresponding specificity of 43.0% (95% CI 39.3%–46.7%).

The AUC for Model 2 was 0.866 (95% CI 0.797–0.936). The best diagnostic accuracy was with a risk score cut-off value of 4, giving a sensitivity and specificity of 84.9% (95% CI 68.1–94.9%) and 74.3% (95% CI 70.9%–77.5%), respectively. A cut-off value of 0.2 had a sensitivity of 100.0% (95% CI 89.4%–100%) and specificity of 17.7% (95% CI 15.0%–20.8%).

As such, we propose a cut-off value of 0.6 for Model 1 and 0.2 for Model 2 to identify patients with low risk of COVID-19. Risk calculators for both models can be found at https://docs.google.com/spreadsheets/d/1whvKbsjfA1wkHPTk2qpzgnm70x-1_SYmGpGaFX8JZss.

Both models performed well in the test cohort. The AUC for Model 1 and Model 2 were 0.906 (95% CI 0.830–0.982) and 0.925 (95% CI 0.863–0.986), respectively. The diagnostic accuracies of these cut-off values to identify patients at low risk of COVID-19 are shown in Table IV.

Table IV. Diagnostic accuracies using cut-off values of 0.6 for Model 1 and 0.2 for Model 2 on training and test cohorts.

Diagnostic accuracy	Training cohort	Test cohort
Model 1		
Sensitivity	100.0%	94.7%
Specificity	43.0%	43.5%
Model 2		
Sensitivity	100.0%	100.0%
Specificity	17.7%	20.6%

Our study population of 1,228 patients had a median length of stay of 1.5 (IQR 1.0–1.8) days and a cumulative length of stay of 1,765 days in isolation facilities. An average of two SARS-CoV-2 PCR tests were performed per patient. If Model 1 was employed, a risk score of 0.6 would identify 509 (41.4%) patients as being at low risk for COVID-19. 736 isolation days and 1,017 PCR tests would have been avoided if low-risk patients were not subjected to PARI wards admission, although one case of COVID-19 would be missed (Fig. 2). If Model 2 was employed using a risk score of 0.2, 222 (18.1%) patients would be identified as being at low risk for COVID-19, avoiding 321 isolation days, 445 PCR tests, without missing any COVID-19 cases (Fig. 2).

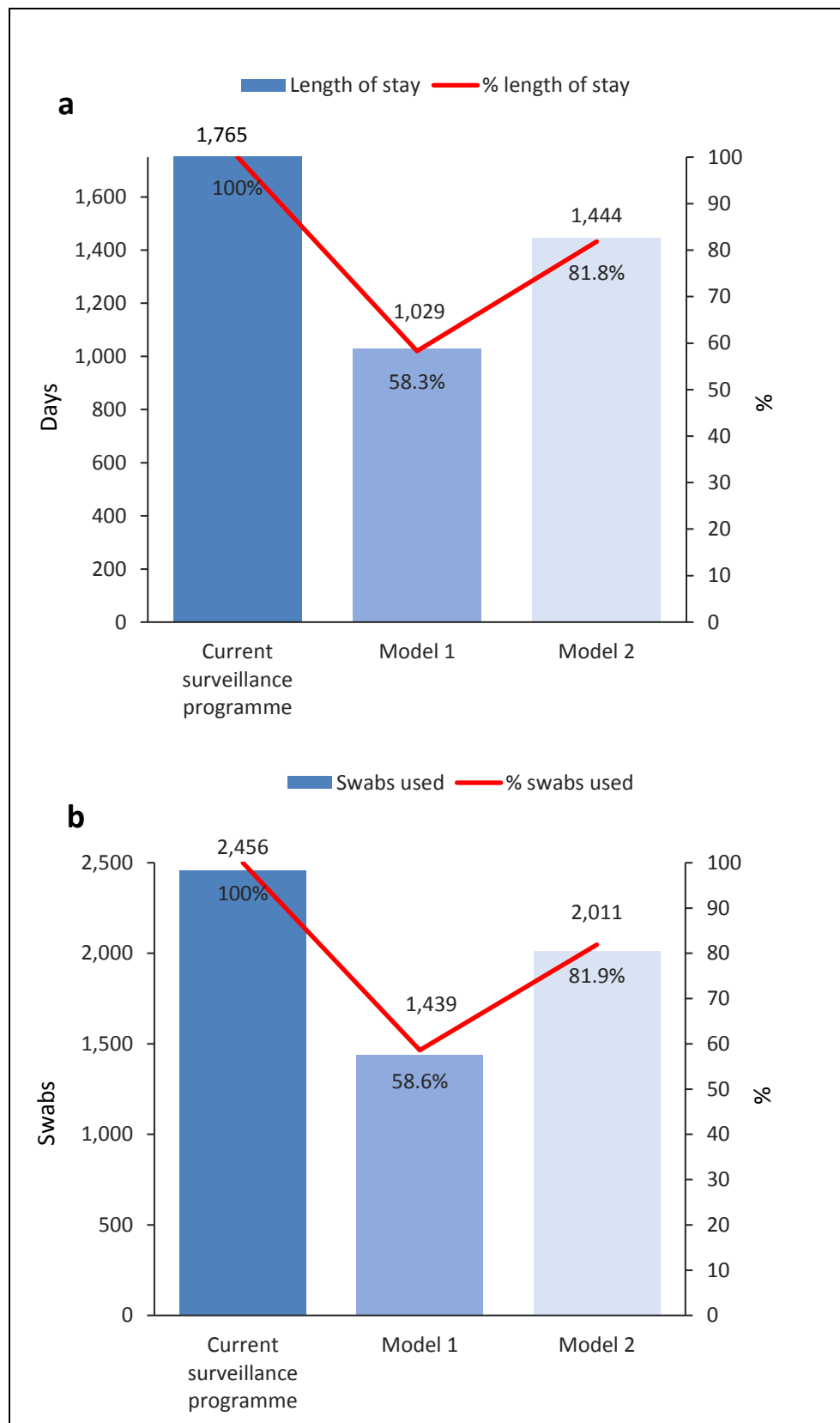


Fig. 2 Bar charts show a comparison of (a) the cumulative duration of stay in the PARI wards in days, (b) and cumulative number of swabs used based on the application of the current surveillance model and two clinical prediction models. A reduction in the number of days of isolation and swabs used were seen with the use of both models.

DISCUSSION

Over a period of 12 weeks, systematic screening for SARS-CoV-2 in all patients presenting with pneumonia and symptoms of ARI had a diagnostic yield of 4.2%. We developed two models (Model 1 and Model 2) based on data readily available from medical history and basic blood investigations to predict the risk of COVID-19 in this patient cohort; both showed good diagnostic performance with an AUC of 0.934 and 0.866, respectively. The country-specific risk factor of dormitory residence was excluded during the development of Model 2 with the intention of making this model more generalisable to other healthcare settings. Cut-off values of 0.6 for Model 1 and 0.2 for Model 2 provided diagnostic sensitivities of 100%, allowing identification of patients at low risk of disease who may not require COVID-19 testing.

Enhanced surveillance through systematic screening of patients with ARI is an effective but highly resource-intensive exercise. Up to 10% of acute hospital beds, including isolation rooms, may be used for such surveillance programmes alone.⁽⁵⁾ Our risk prediction score would have obviated the need for isolation and testing in up to 41.4% of our cohort of patients with pneumonia and ARI. Missed cases of COVID-19 are expected trade-offs of this risk stratification strategy, but there have been reassuring reports that basic infection control measures encompassing surgical masks and standard hand hygiene practices are effective in minimising the risks of nosocomial spread to healthcare workers and patients with inadvertent exposure to COVID-19 cases.^(8,9)

Dormitory residence was the single largest predictor of COVID-19 in our study and a reflection of an epidemiological phenomenon that is unique to Singapore. While the first wave of COVID-19 infections in Singapore were mostly imported cases, the huge surge in cases observed from April 2020 was due to outbreaks in migrant worker dormitories. At the time of writing, more than 90% of COVID-19 cases in Singapore were from migrant workers' dormitories.⁽¹⁰⁾ Widespread transmission of COVID-19 has also been reported in other

congregated settings such as skilled nursing facilities,⁽¹¹⁾ correctional facilities⁽¹²⁾ and cruise ships.⁽¹³⁾ The exact risk posed by residence in a congregated setting is, however, highly dependent on the local epidemiology and public health response. Although COVID-19 cases were also reported in a few nursing home residences in Singapore, early preventive strategies likely mitigated spread within this community.⁽¹⁴⁾ As such, a nursing home residence, unlike a dormitory residence, was not a risk predictor in our study.

Exposure history is another important risk predictor. In a cohort of at-risk patients referred to Singapore's National Centre for Infectious Diseases for SARS-CoV-2 testing, travel to Wuhan, China, the epicentre of COVID-19 outbreak during the study period, was the greatest risk factor for COVID-19.⁽¹⁵⁾ As our study included only patients who did not fulfil the official Ministry of Health's criteria for COVID-19 suspect cases, which was largely based on travel history to high-risk regions (Appendix), having an overseas travel history did not predict risk of COVID-19. However, we did observe that positive contact with someone having an ARI, regardless of COVID-19 status, increased the risk of COVID-19 and was a major predictor in Model 2.

Interestingly, patients who were COVID-19 positive were more likely to have an accompanying headache at presentation in our study. Based on a meta-analysis, headache was reported in 12% (95% CI 4%–23%) of patients with COVID-19,⁽¹⁶⁾ but most studies did not compare the prevalence of headache between patients with and without COVID-19. In a survey of healthcare workers with mild respiratory symptoms who were tested for SARS-CoV-2 in the Netherlands, headache was reported in 72% of participants with COVID-19 and 41.5% of those without.⁽¹⁷⁾ It is hypothesised that headache in COVID-19 may be due to involvement of trigeminal nerve endings and vascular endothelial cells by the SARS-CoV-2 virus via angiotensin-converting enzyme 2 (ACE-2) receptors.⁽¹⁸⁾ As respiratory symptoms in COVID-

19 are generally indistinguishable from other respiratory infections, non-respiratory manifestations (mediated by ACE-2 receptors) may have more diagnostic utility.

An earlier systematic review of prediction models for diagnosis of COVID-19 concluded that existing, non-peer-reviewed studies had substantial selection bias.⁽⁶⁾ In addition, several of the prediction models included variables (e.g. interleukin-6 and computed tomography of the chest) that are not routinely performed in clinical practice.^(19,20) Since publication of the systematic review, several peer-reviewed studies on risk predictors for COVID-19 became available. One was a case-control study conducted in Singapore; the full model consisting of 16 variables had excellent performance with an AUC of 0.91.⁽¹⁵⁾ However, a degree of selection bias remains, as the study cohort consisted of patients referred specifically for SARS-CoV-2 testing, including cases identified through national contact tracing. The reported diagnostic yield of 6.9% is supportive of a higher risk profile compared to our study. Another study performed in the United States also derived a prediction model based on high-risk patients who underwent SARS-CoV-2 testing, with an overall diagnostic yield of 8.6%.⁽²¹⁾

Our prediction models were based on an unbiased cohort, as all patients who were admitted for pneumonia and symptoms of ARI underwent SARS-CoV-2 testing. Consequently, our results are likely to be more useful for clinical triaging, as the main challenge lies in risk-stratifying patients with no travel history to high-risk regions or contact history with COVID-19 cases. In addition, both models utilise variables that are readily available from routine clinical history and blood investigations, facilitating practical implementation.

We acknowledge several limitations of this study. Firstly, there are limitations inherent in a retrospective study. Secondly, the prediction models may not be generalisable to patients with pneumonia and ARI in the outpatient setting, or to those who present either much earlier or much later than our study cohort, as some of the blood indices included in our models may fluctuate during the course of illness.⁽²²⁾ Thirdly, our models need to be validated in other

cohorts to confirm their diagnostic performance. We anticipate that Model 1 is likely to perform more poorly in non-local settings because dormitory residence as a risk factor is specific to Singapore. However, Model 2 utilises non-country-specific risk factors and is likely to be more generalisable. We wish to stress that prediction models are meant to complement the decision-making process, and individual clinicians should exercise discernment by considering prevailing as well as emerging epidemiological risk factors.

We anticipate that the prediction models derived in this study would be most useful in identifying patients at low risk for COVID-19 infection, which would better rationalise the use of healthcare resources. Validation of the models in external cohorts is recommended before adopting their use in clinical practice.

REFERENCES

1. World Health Organization. Coronavirus disease (COVID-19) weekly report. Available at: <https://www.who.int/publications/m/item/weekly-epidemiological-update---9-february-2021>. Accessed February 10, 2021.
2. Wee LE, Fua TP, Chua YY, et al. Containing COVID-19 in the emergency department: the role of improved case detection and segregation of suspect cases. *Acad Emerg Med* 2020; 27:379-87.
3. Ng Y, Li Z, Chua YX, et al. Evaluation of the effectiveness of surveillance and containment measures for the first 100 patients with COVID-19 in Singapore - January 2-February 29, 2020. *MMWR Morb Mortal Wkly Rep* 2020; 69:307-11.
4. Lee VJ, Chiew CJ, Khong WX. Interrupting transmission of COVID-19: lessons from containment efforts in Singapore. *J Travel Med* 2020; 27:taaa039.
5. Wee LE, Hsieh JYC, Phua GC, et al. Respiratory surveillance wards as a strategy to reduce nosocomial transmission of COVID-19 through early detection: the experience

- of a tertiary-care hospital in Singapore. *Infect Control Hosp Epidemiol* 2020; 41:820-5.
6. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of COVID-19 infection: systematic review and critical appraisal. *BMJ* 2020; 369:m1328.
 7. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Cancer* 2015; 112:251-9.
 8. Ng K, Poon BH, Puar THK, et al. COVID-19 and the risk to health care workers: a case report. *Ann Intern Med* 2020; 172:766-7.
 9. Wong SCY, Kwong RTS, Wu TC, et al. Risk of nosocomial transmission of coronavirus disease 2019: an experience in a general ward setting in Hong Kong. *J Hosp Infect* 2020; 105:119-27.
 10. Ministry of Health, Singapore. 20 June 2020 daily report on COVID-19. Available at: <https://www.moh.gov.sg/docs/librariesprovider5/2019-ncov/situation-report--20-jun-2020.pdf>. Accessed June 20, 2020.
 11. Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N Engl J Med* 2020; 382:2081-90.
 12. Rubin R. The challenge of preventing COVID-19 spread in correctional facilities. *JAMA* 2020; 323:1760-1.
 13. Russell TW, Hellewell J, Jarvis CI, et al. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. *Euro Surveill* 2020; 25:2000256.
 14. Tan LF, Seetharaman SK. COVID-19 outbreak in nursing homes in Singapore. *J Microbiol Immunol Infect* 2020 May 13. <https://doi.org/10.1016/j.jmii.2020.04.018>. [Epub ahead of print]

15. Sun Y, Koh V, Marimuthu K, et al. Epidemiological and clinical predictors of COVID-19. *Clin Infect Dis* 2020; 71:786-92.
16. Borges do Nascimento IJ, Cacic N, Abdulazeem HM, et al. Novel coronavirus infection (COVID-19) in humans: a scoping review and meta-analysis. *J Clin Med* 2020; 9:941.
17. Tostmann A, Bradley J, Bousema T, et al. Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020. *Euro Surveill* 2020; 25:2000508.
18. Bolay H, Gül A, Baykan B. COVID-19 is a real headache! *Headache* 2020 May 15. <https://doi.org/10.1111/head.13856>. [Epub ahead of print]
19. Feng C, Huang Z, Wang L, et al. A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. *medRxiv* 2020 Mar 20. Available at: <https://www.medrxiv.org/content/10.1101/2020.03.19.20039099v1>. Preprint.
20. Song CY, Xu J, He JQ, Lu YQ. COVID-19 early warning score: a multi-parameter screening tool to identify highly suspected patients. *medRxiv* 2020 Mar 8. Available at: <https://doi.org/10.1101/2020.03.05.20031906>. Preprint.
21. Jehi L, Ji X, Milinovich A, et al. Individualizing risk prediction for positive coronavirus disease 2019 testing: results from 11,672 patients. *Chest* 2020; 158:1364-75.
22. Mei Y, Weinberg SE, Zhao L, et al. Risk stratification of hospitalized COVID-19 patients through comparative studies of laboratory results with influenza. *EClinicalMedicine* 2020; 26:100475.

APPENDIX

Ministry of Health Singapore's COVID-19 suspect case definition

4 Feb 2020	23 Feb 2020	9 March 2020
<p>(a) A person with clinical signs and symptoms suggestive of pneumonia or severe respiratory infection with breathlessness AND travel to mainland China within 14 days before onset of illness;</p> <p>OR</p> <p>(b) A person with an acute respiratory illness of any degree of severity who, within 14 days before onset of illness had:</p> <ul style="list-style-type: none"> • Been to Hubei province (including Wuhan city) or Zhejiang province (including Hangzhou city), China; OR • Been to a hospital in mainland China; OR • Had close contact with a case of 2019 novel coronavirus infection; OR • Had frequent or close contact during work with recent travellers from mainland China (travel history in the last 14 days). 	<p>(a) A person with clinical signs and symptoms suggestive of pneumonia or severe respiratory infection with breathlessness AND within 14 days before onset of illness had:</p> <ul style="list-style-type: none"> • Been to mainland China; OR • Been to Daegu city or Cheongdo county, South Korea. <p>(b) A person with an acute respiratory illness of any degree of severity who, within 14 days before onset of illness had:</p> <ul style="list-style-type: none"> • Been to Hubei province (including Wuhan city) or Zhejiang province (including Hangzhou city), China; OR • Been to a hospital in mainland China; OR • Had close contact with a case of COVID-19 infection. 	<p>(a) A person with clinical signs and symptoms suggestive of pneumonia or severe respiratory infection with breathlessness AND who within 14 days before onset of illness had travelled abroad (i.e. to any country outside of Singapore).</p> <p>(b) A person with an acute respiratory illness of any degree of severity who, within 14 days before onset of illness had:</p> <ul style="list-style-type: none"> • Been to any of the areas requiring heightened vigilance; OR • Been to any hospital abroad; OR • Close contact with a case of COVID-19 infection.