

Integrated Model for COVID-19 Diagnosis Based on Computed Tomography AI, and Clinical Features: A Multicenter Cohort Study

Yuki Kataoka (✉ youkiti@gmail.com)

1. Department of Internal Medicine, Kyoto Min-Iren Asukai Hospital <https://orcid.org/0000-0001-7982-5213>

Yuya Kimura

Clinical Research Center, Department of Respiratory Medicine, National Hospital Organization Tokyo National Hospital

Tatsuyoshi Ikenoue

Human Health Sciences, Kyoto University Graduate School of Medicine

Yoshinori Matsuoka

Department of Emergency Medicine, Kobe City Medical Center General Hospital

Junji Kumasawa

Department of Critical Care Medicine, Sakai City Medical Center

Kentaro Tochatani

Department of Infectious Diseases, Kyoto City Hospital

Hiraku Funakoshi

Department of Emergency and Critical Care Medicine, Tokyobay Urayasu Ichikawa Medical Center

Tomohiro Hosoda

Department of Infectious Disease, Kawasaki Municipal Kawasaki Hospital

Aiko Kugimiya

Department of Emergency and Critical Care Medicine, Yamanashi Prefectural Central Hospital

Michinori Shirano

Department of Infectious Diseases, Osaka City General Hospital

Fumiko Hamabe

Department of Radiology, National Defense Medical College Hospital

Sachiyo Iwata

Division of Cardiovascular Medicine, Hyogo Prefectural Kakogawa Medical Center

Shingo Fukuma

Division of Cardiovascular Medicine, Hyogo Prefectural Kakogawa Medical Center

Japan COVID-19 AI team

Keywords: COVID-19, diagnosis, Decision support tool, machine learning, LightGBM, test accuracy

Posted Date: October 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-979599/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

We developed and validated a machine learning diagnostic model for novel coronavirus (COVID-19) disease, integrating artificial-intelligence-based computed tomography (CT) imaging and clinical features.

Methods

We conducted a retrospective cohort study in 11 Japanese tertiary care facilities that treated COVID-19 patients. Participants were tested using both real-time reverse transcription polymerase chain reaction (RT-PCR) and chest CT between January 1 and May 30, 2020. We chronologically split the dataset in each hospital into training and test sets, containing patients in a 7:3 ratio. Light Gradient Boosting Machine model was used for analysis.

Results

A total of 703 patients were included with two models – the full model and the A-blood model – developed for their diagnosis. The A-blood model included eight variables (the Ali-M3 confidence, along with seven clinical features of blood counts and biochemistry markers). The areas under the receiver-operator curve of both models (0.91, 95% confidence interval (CI), 0.86 to 0.95 for the full model and 0.90, 95% CI, 0.86 to 0.94 for the A-blood model) were better than that of the Ali-M3 confidence (0.78, 95% CI, 0.71 to 0.83) in the test set.

Conclusions

The A-blood model, a COVID-19 diagnostic model developed in this study, combines machine-learning and CT evaluation with blood test data and is better than the Ali-M3 framework existing for this purpose. This would significantly aid physicians in making a quicker diagnosis of COVID-19.

1. Introduction

Since the discovery of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) species – and the resulting novel coronavirus disease (COVID-19) – in December 2019, humanity has been plunged into a global pandemic (1). Although the devastating effects of the virus have been mitigated by vaccination (2,3), breakthrough infections caused by new variants of the variants prevent the pandemic from coming to an end (4).

The gold standard for COVID-19 diagnosis is the real-time reverse transcription polymerase chain reaction (RT-PCR) test. However, RT-PCR has several drawbacks: in several cases, it is known to be insufficiently

sensitive to the virus even in symptomatic patients, leading to false negatives (5,6). In addition, its diagnosis in facilities that require specimen transport takes a long turnaround time (7). These aspects reveal the need for a more accurate and timely diagnosis; for this, several diagnostic models have been developed using clinical characteristics, laboratory data, and radiographic images. However, most such models have not been validated with datasets external to the development phase (8). Moreover, methodological flaws and/or underlying biases, making it difficult to determine the model validity. In consequence, there are no diagnostic models using chest computed tomography (CT) with potential clinical use (9).

There is no diagnostic system that automatically interprets the CT and clinical features. In addition, to overcome these limitations in diagnostic models using chest CT, we have externally validated a deep-learning-based, CT diagnostic system for COVID-19 (Ali-M3) (10). To further improve its accuracy, it is important to properly diagnose COVID-19 patients without pneumonia detectable by CT. For this purpose, we integrated the Ali-M3 model with the clinical characteristics of patients suspected of having COVID-19 using machine learning, and validated this new system.

2. Materials And Methods

We used datasets for the external validation of the Ali-M3. The details of the datasets were published elsewhere (10). We used the guidelines of the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement to report this study (Supplementary Table 1) (11). The institutional review board of each facility approved of our study and waived the need to obtain written informed consent.

2.1 Study Design

This was a retrospective cohort study conducted in 11 Japanese tertiary care facilities that provided treatment for patients with COVID-19.

2.2 Participants

We included patients who tested both RT-PCR and chest CT for the diagnosis of COVID-19. Potentially eligible participants were identified as those who had, on the advice of physicians, taken both the RT-PCR and chest CT tests when they presented with symptoms or were suspected of having COVID-19. RT-PCR results were extracted from the patients' medical records at each facility. We selected patients by using consecutive sampling methods between January 1 and May 30, 2020. We excluded patients when the time-interval between chest CT and the first RT-PCR assay exceeded 7 days.

2.3 Chest CT and Artificial Intelligence

We considered, for each patient, the CT image that was taken closest to the onset of symptoms. All these images featured the patient in a supine position.

Ali-M3 is a three-dimensional, deep-learning framework for the detection of COVID-19 infections, developed from 7,000 chest CT scans (12). It predicts COVID-19 infections with confidence levels in the range of 0–1. The learning of Ali-M3 was halted before our evaluation (10) and the investigators who entered data from the CT images into Ali-M3 were blinded to the corresponding RT-PCR results. The area under the curve (AUC) of Ali-M3 for predicting a COVID-19 diagnosis was 0.797 (95% confidence intervals [CI]: 0.762-0.833) (10).

2.4 Clinical Characteristics

We extracted, from electronic medical records, those clinical characteristics that were recorded at a time closest to the date of the chest CT scans. At the time of data acquisition, the turnaround time of RT-PCR was a few days; therefore, all predictive variables were recorded without the RT-PCR results.

2.5 Reference Standard

COVID-19 was diagnosed by the RT-PCR test, which detected the presence of the nucleic acid of SARS-CoV-2 in the sputum, throat swabs, and secretions of the lower respiratory tract (13). This test was established as the primary reference standard. Although the findings of chest CT, interpreted by radiologists, were included as a reference standard in the AI development phase of this framework, we did not include it as the reference standard in this study.

2.6 Model Development

We used the machine learning model, Light Gradient Boosting Machine (LightGBM), which is also a highly effective gradient-boosting decision tree algorithm (13). In the boosting algorithm, a weak classifier (decision tree) is sequentially created to minimize the prediction errors made by the previous classifier (14). The result is a powerful ensemble classifier with superior predictive performance. To avoid overfitting, parameters specific to the algorithm (known as *hyperparameters*) must be well tuned before fitting them to the final model, which also needs to treat missing data as such.

The creation of prediction models consists of three steps. First, the data set in each hospital was chronologically split into a training set (with 70% of the patients) and a test set (with the remaining 30% of subjects). Second, hyperparameters were tuned to maximize the area under the receiver-operator curve (AUROC) by performing a five-fold cross-validation on the training set using stratified splitting in equally sized groups. A Bayesian optimization algorithm was used for tuning, with the search parameters and spaces given as follows: “num_leaves” (maximum number of leaves in one tree) at 10–150, “max_depth” (maximum tree depth) at 10–150, “learning_rate” (learning rate) at 0.005–0.5, “subsample_for_bin” (number of data sampled to construct feature-discrete bins) at 20000–300000, “min_child_samples” (minimal number of data in one leaf) at 10–100, “reg_alpha” (L1 regularization) at 0.0–1.0, “reg_lambda” (L2 regularization) at 0.0–1.0, “colsample_bytree” (the rate of features selected in training each tree) at 0.5–1.0, “subsample” (the rate of data selected in training each tree) at 0.5–1.0, and “is_unbalanced”, True or False. Optimal “n_estimator” (number of trees) was automatically determined by employing early stopping (“early_stopping_rounds” = 100). Third, we used the entire training set to fit two final models,

whose hyperparameters were also tuned. One model – the full model – included all the above-mentioned variables, while the other model – the A-blood model – included only eight limited variables (the Ali-M3 confidence variable, in addition to seven variables pertaining to blood test results: white blood cell, hemoglobin, platelet, aspartate aminotransferase, alanine aminotransferase, lactate dehydrogenase, and C-reactive protein). These blood test variables were selected owing to the ease of availability of their data and due to their relative importance in the full model, which was computed as Shapley Additive exPlanations (SHAP) values (13). SHAP values quantify the association between each variable and the outcome of each patient.

We used Python version 3.7.11 (15) with lightgbm version 2.2.3 (16) and hyperopt version 0.1.2 (17).

2.7 Model Validation

We differentiated between the confidence of the machine-learning models and the Ali-M3 framework by using AUROC in the test set, with 95% confidence intervals calculated with bootstrapped resampling (1000 samples). AUROC is an effective measure of overall diagnostic accuracy, which is deemed to be “outstanding” if $AUROC \geq 0.9$, “excellent” if $0.8 < AUROC < 0.9$, and “acceptable” if $0.7 < AUROC < 0.8$ (18). Calibration was assessed using the Brier score (19) and a calibration plot. The formulation of the Brier score for a binary prediction is given by:

Brier Score = $\frac{1}{N} \sum_{i=1}^N (p_i - Y_i)^2$, where the score predicts the occurrence of the outcome, ranging from 1 for an outcome that definitely occurs and 0 for one that definitely does not occur, where smaller values indicate superior model performance. The AUROC values and Brier scores of the machine-learning models were compared with those of the Ali-M3 confidence using bootstrapped resampling (1000 samples). We calculated the SHAP values and presented them in the figures.

3. Results

3.1 Patient Characteristics

A total of 703 patients were included in the study, including 326 PCR-positive and 377 PCR-negative patients. In training set, we included 490 patients including 247 PCR-positive patients. In test set, we included 213 patients including 79 PCR-positive patients. Patient characteristics are shown in Table 1.

3.2 Model Performance

We developed two models – a full model and an A-blood model. Details on the A-blood model are accessible in web calculators online (20). We show the model discrimination and calibration of the test data in Table 2. The AUROC values of both the full model (0.91, 95% CI, 0.86 to 0.95) and the A-blood model (0.90, 95% CI, 0.86 to 0.94) were better than that of the Ali-M3 confidence (0.78, 95% CI, 0.71 to 0.83) in the test set. The calibration evaluated by the Brier scores of both the full model (0.10, 95% CI, 0.07 to 0.13) and the A-blood model (0.12, 95% CI, 0.08 to 0.16) were better than that of the Ali-M3

confidence (0.23, 95% CI, 0.19 to 0.27) in the test set. The ROCs of the test data are shown in Figure 1, with SHAP values shown in Figures 2 and 3. Figure 2 shows all of the predictive variables that we used. Figures 4–6 show the calibration plots for the Ali-M3 framework, the full model, and the A-blood model.

4. Discussion

We developed and validated two integrated diagnostic models of the Ali-M3 framework with the clinical characteristics of patients with suspected COVID-19. Based on the relative importance of each variable, we shrank the full model to a more compact A-blood model, whose parameters included the Ali-M3 confidence and eight routinely collected blood markers. This A-blood model showed better discrimination and calibration performance than the full model.

Our diagnostic model is the first to automatically interpret clinical data in conjunction with CT scans. Several problems faced by existing diagnostic models, such as separate collection of cases and controls, lack of external validation, and insufficient reporting (8,9), have been overcome in this study with rigorous methodology, with our model achieving good discrimination and calibration performance.

The A-blood model would allow for quicker diagnoses. Even if the RT-PCR test existed in the facility, the A-blood model would be a better option because of its lower turnaround time, which requires only general blood test and CT results. When the RT-PCR results are not known, the “A-blood” model could help physicians determine indications for timely treatment with antibody drugs (21). For patients for whom COVID-19 infection cannot be ruled out based on a single RT-PCR negative, physicians may be able to use the “A-blood” model to determine if a patient can be released from quarantine. These clinical implications need to be evaluated in further studies (22).

This study has several limitations. First, the dataset used in this study is from the first wave of infections in the spring of 2020, which does not include vaccinated patients or the latter variants of the SARS-CoV2 virus. Therefore, it is necessary to further expand on this external validation. A second limitation is the occurrence of false negatives, which includes patients falsely regarded as COVID-negative with a single negative PCR result. This misclassification may affect the accuracy, but the magnitude of this bias cannot be predicted. Further studies are required with datasets that also include sufficient follow-up.

In conclusion, we developed the A-blood model, which is a COVID-19 diagnostic tool that combines machine learning and CT evaluation with blood test data. Physicians would be able to use this model for the rapid diagnosis of COVID-19. Further validation studies, especially those including SARS-CoV-2 variants and subjects inoculated with different vaccines, are warranted.

Declarations

Reporting Checklist:

The authors have completed the TRIPOD reporting checklist.

Data Sharing Statement:

Data will be shared by the reasonable request to the corresponding author.

Conflicts Of Interest:

All authors have completed the ICMJE unified disclosure form. Junichi Matsumoto received lecture fee from M3 Inc. The other authors declare no conflicts of interest.

Funding information:

This study was partially supported by the Kyoto University managing fund for English editing. The article processing fee was supported by Scientific Research Works Peer Support Group (SRWS-PSG). Funders played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Author contributions:

(I) Conception and design: YKa and TI.

(II) Administrative support: TI

(III) Provision of study materials or patients: YKa, YM, JM, JK,KT, HF, TH, AK, MS, FH, SI, SF, SH, TH, SY, HS, KN, HN, KA, HS, HS, HN, TA, NY, RO, TK, YI, and YM

(IV) Collection and assembly of data: YKa and TI.

(V) Data analysis and interpretation: YKa, YKi, and TI

(VI) Manuscript writing: All authors

(VII) Final approval of the manuscript: All authors

Acknowledgements

We thank M3 Inc. and Clinical Porter for providing free Ali-M3 analysis and data storage, although they did not participate in the preparation of the protocol and the manuscript. To access Ali-M3, the reader can contact M3 (m3-ai-lab@m3.com). We thank Editage (www.editage.jp) for the English language review. We also thank Ms. Kyoko Wasai, who assisted in retrieving the data.

References

1. Carvalho T, Krammer F, Iwasaki A. The first 12 months of COVID-19: a timeline of immunological insights. *Nat Rev Immunol* [Internet]. 2021 Apr 15;21(4):245–56. Available from:

- <http://www.nature.com/articles/s41577-021-00522-1>
2. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N Engl J Med* [Internet]. 2021 Feb 4;384(5):403–16. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa2035389>
 3. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med* [Internet]. 2020 Dec 31;383(27):2603–15. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa2034577>
 4. Bergwerk M, Gonen T, Lustig Y, Amit S, Lipsitch M, Cohen C, et al. Covid-19 Breakthrough Infections in Vaccinated Health Care Workers. *N Engl J Med* [Internet]. 2021 Jul 28;NEJMoa2109072. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa2109072>
 5. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. *BMJ* [Internet]. 2020 May 12;m1808. Available from: <https://www.bmj.com/lookup/doi/10.1136/bmj.m1808>
 6. Kortela E, Kirjavainen V, Ahava MJ, Jokiranta ST, But A, Lindahl A, et al. Real-life clinical sensitivity of SARS-CoV-2 RT-PCR test in symptomatic patients. Ricci S, editor. *PLoS One* [Internet]. 2021 May 21;16(5):e0251661. Available from: <https://dx.plos.org/10.1371/journal.pone.0251661>
 7. Kohli A, Joshi A, Shah A, Jain RD, Gorlawar A, Dhapare A, et al. Does CT help in reducing RT-PCR false negative rate for COVID-19? *Indian J Radiol Imaging* [Internet]. 2021 Jan 13;31(S 01):S80–6. Available from: http://www.thieme-connect.de/DOI/DOI?10.4103/ijri.IJRI_739_20
 8. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* [Internet]. 2020 Apr 7;m1328. Available from: <https://www.bmj.com/lookup/doi/10.1136/bmj.m1328>
 9. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* [Internet]. 2021 Mar 15;3(3):199–217. Available from: <http://www.nature.com/articles/s42256-021-00307-0>
 10. Ikenoue T, KATAOKA Y, Matsuoka Y, Matsumoto J, Kumasawa J, Tochitani K, et al. Accuracy of deep learning based computed tomography diagnostic system of COVID-19: a consecutive sampling external validation cohort study. *medRxiv* [Internet]. 2020 Jan 1;2020.11.15.20231621. Available from: <http://medrxiv.org/content/early/2020/11/18/2020.11.15.20231621.abstract>
 11. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med* [Internet]. 2015 Jan 6;162(1):55. Available from: <http://annals.org/article.aspx?doi=10.7326/M14-0697>
 12. TAD A. Assisted Analysis Based On Chest CT Imaging. *Alibaba DAMO Acad*. 2020;2.
 13. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med* [Internet]. 2020 Aug 18;173(4):262–7. Available from: <https://www.acpjournals.org/doi/10.7326/M20-1495>

14. Friedman JH, Popescu BE. Importance Sampled Learning Ensembles. 2003;
15. Welcome to Python.org [Internet]. [cited 2021 Sep 10]. Available from: <https://www.python.org/>
16. GitHub - microsoft/LightGBM: A fast, distributed, high performance gradient boosting (GBT, GBDT, GBRT, GBM or MART) framework based on decision tree algorithms, used for ranking, classification and many other machine learning tasks. [Internet]. [cited 2021 Sep 10]. Available from: <https://github.com/Microsoft/LightGBM>
17. Hyperopt Documentation [Internet]. [cited 2021 Sep 10]. Available from: <http://hyperopt.github.io/hyperopt/>
18. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* [Internet]. 2010 Sep;5(9):1315–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20736804>
19. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* [Internet]. 2010 Jan;21(1):128–38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20010215>
20. A-blood model web calculator [Internet]. [cited 2021 Oct 14]. Available from: <https://youkiti.shinyapps.io/ablood-for-Dx-COVID-19/>
21. Siemieniuk RA, Bartoszko JJ, Díaz Martinez JP, Kum E, Qasim A, Zeraatkar D, et al. Antibody and cellular therapies for treatment of covid-19: a living systematic review and network meta-analysis. *BMJ* [Internet]. 2021 Sep 23;n2231. Available from: <https://www.bmj.com/lookup/doi/10.1136/bmj.n2231>
22. Lijmer JG, Leeflang M, Bossuyt PMM. Proposals for a Phased Evaluation of Medical Tests. *Med Decis Mak* [Internet]. 2009 Sep 15;29(5):E13–21. Available from: <http://journals.sagepub.com/doi/10.1177/0272989X09336144>

Tables

Table1. Patient Characteristics

Variables	COVID-19 PCR: positive (N=326)	COVID-19 PCR: negative (N=377)
Age, years	55 [43–68]	68 [45–79]
Sex, male, N (%)	197 (60.4)	220 (58.4)
Smoking, current/ex-smoker, N (%)	46 (14.1)	64 (17.0)
Contact history, N (%)		
With family patients	39 (12.0)	6 (1.6)
With non-family patients	78 (23.9)	34 (9.0)
None	209 (64.1)	337 (89.3)
Travel overseas, N (%)	39 (12.0)	14 (3.7)
Duration of symptom, days	6 [4–9]	4 [2–9]
Missing data, N (%)	15 (4.6)	15 (4.0)
Symptoms, N (%)		
Cough	123 (37.7)	109 (28.9)
Chill	48 (14.7)	43 (11.4)
Sore throat	81 (24.8)	78 (20.7)
Diarrhea	41 (12.6)	25 (9.3)
Muscle pain	29 (8.9)	17 (4.5)
Conjunctivitis	18 (5.5)	12 (3.2)
Taste disorder	33 (10.1)	22 (5.8)
Complications, N (%)		
Coronary arterial diseases	12 (3.7)	38 (10.1)
Cerebrovascular diseases	25 (7.7)	46 (12.2)
Chronic heart failures	21 (6.4)	59 (15.6)
Chronic kidney diseases	17 (5.2)	53 (14.1)
COPD	16 (4.9)	62 (16.4)
Malignancy	29 (8.9)	87 (23.1)
Immune disorders	5 (1.5)	28 (7.4)
Hypertension	52 (16.0)	85 (22.5)

Diabetes mellitus	56 (17.1)	84 (22.3)
Others	59 (18.1)	158 (41.9)
Vital signs		
Body temperature, degrees Celsius	37.2 [36.6–38.1]	37.2 [36.7–38.0]
Missing data, N (%)	14 (4.3)	25 (6.6)
Systolic blood pressure, mmHg	126 [113–138]	130 [114–148]
Missing data, N (%)	20 (6.1)	38 (10.1)
Diastolic blood pressure, mmHg	79 [70–89]	77 [67–87]
Missing data, N (%)	20 (6.1)	38 (10.1)
Heart rate, beats per minute	86 [78–98]	93 [80–108]
Missing data, N (%)	11 (3.4)	35 (9.3)
Respiratory rate, breaths per minute	18 [16–21]	20 [16–24]
Missing data, N (%)	65 (19.9)	138 (36.6)
Laboratory data		
White blood cell, counts $\times 10^3/\mu\text{L}$	4.1 [1.8–5.2]	9.2 [6.4–12.5]
Missing data, N (%)	14 (4.3)	44 (11.7)
Hemoglobin, g/dL	14.0 [12.9–15.2]	12.2 [10.3–13.5]
Missing data, N (%)	23 (7.1)	60 (15.9)
Platelet, counts $\times 10^4/\mu\text{L}$	18.9 [15.1–25.0]	23.5 [16.4–30.0]
Missing data, N (%)	16 (4.9)	45 (11.9)
Aspartate aminotransferase, U/L	32 [24–54]	27 [19–40]
Missing data, N (%)	13 (4.0)	43 (11.4)
Alanine aminotransferase, U/L	30 [17–46]	20 [13–34]
Missing data, N (%)	13 (4.0)	43 (11.4)
Lactate dehydrogenase, U/L	282 [216–403]	244 [186–324]
Missing data, N (%)	15 (4.6)	54 (14.3)
C-reactive protein, mg/dL	3.7 [0.5–9.5]	5.5 [1.4–11.9]
Missing data, N (%)	14 (4.3)	57 (15.1)

Computed tomography data

Ali-M3 confidence, % 0.93 [0.52–1.00] 0.25 [0.01–0.71]

All continuous variables are not normally distributed and are presented as median [interquartile range]; categorical variables are presented as N (%). COPD, chronic obstructive pulmonary disease.

Table2. Model discrimination and calibration in the test data

Model	Score		Difference between Ali-M3 confidence	
	AUROC (95% CI)	Brier score (95% CI)	AUROC (95% CI)	Brier score (95% CI)
Full model*	0.91 (0.86 to 0.95)	0.10 (0.07 to 0.13)	0.13 (0.07 to 0.19)	-0.13 (-0.17 to -0.09)
A-blood model†	0.90 (0.86 to 0.94)	0.12 (0.08 to 0.16)	0.12 (0.07 to 0.18)	-0.11 (-0.15 to -0.07)
Ali-M3 confidence	0.78 (0.71 to 0.83)	0.23 (0.19 to 0.27)		

AUROC, area under receiver operator curve; CI, confidence interval.

* Machine learning model using all variables

† Machine learning model using 8 variables including Ali-M3 Confidence, White blood cell, Hemoglobin, Platelet, Aspartate aminotransferase, Alanine aminotransferase, Lactate dehydrogenase, and C-reactive protein.

Figures

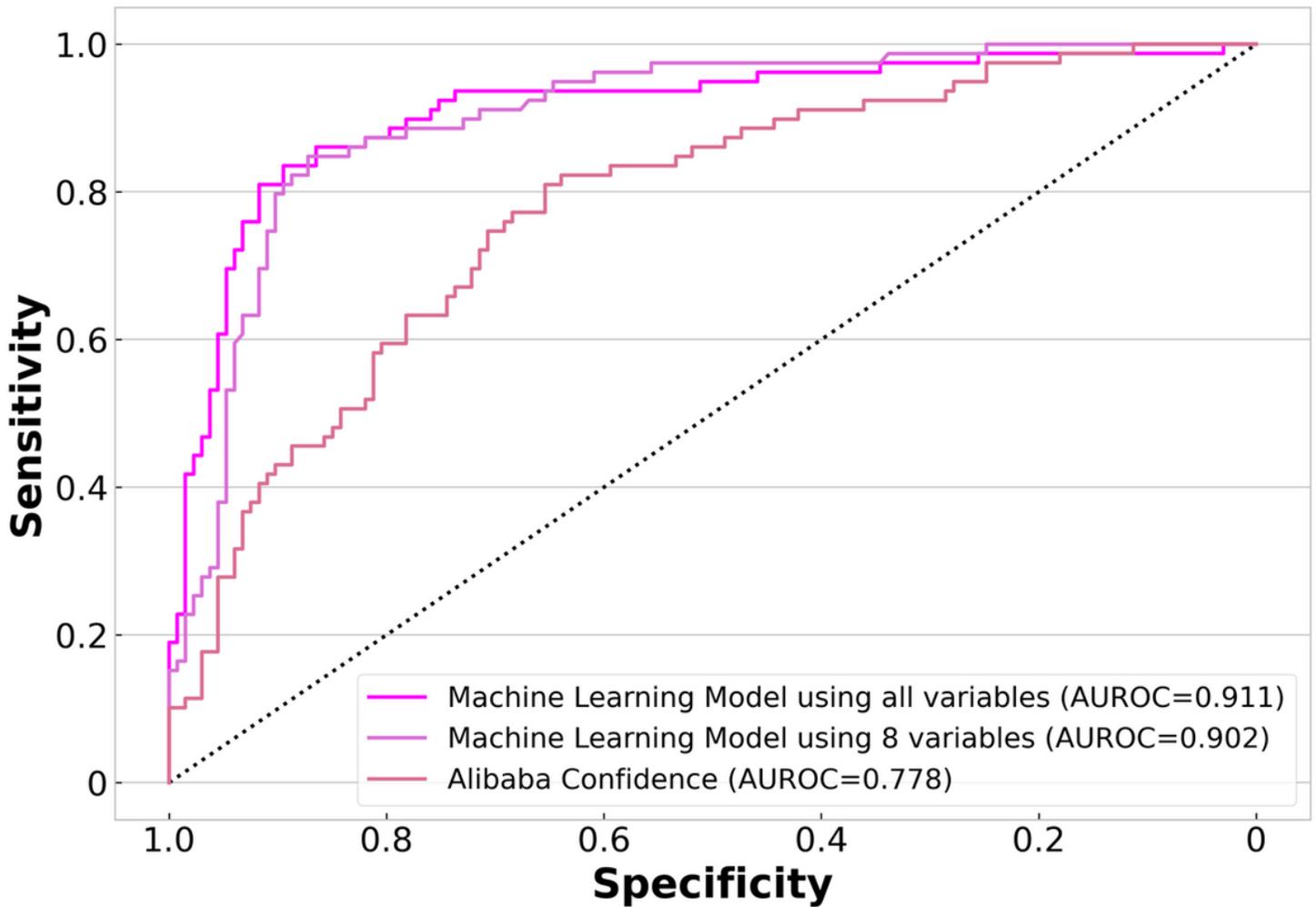


Figure 1

Receiver operator curves (ROCs) of COVID-19 PCR positive prediction models in the test data. Models include machine learning model using all variables, machine learning model using 8 variables (Ali-M3 confidence; White blood cell; Hemoglobin; Platelet; Aspartate aminotransferase; Alanine aminotransferase; Lactate dehydrogenase; C-reactive protein), and Ali-M3 confidence. AUROC: area under the receiver operator curves

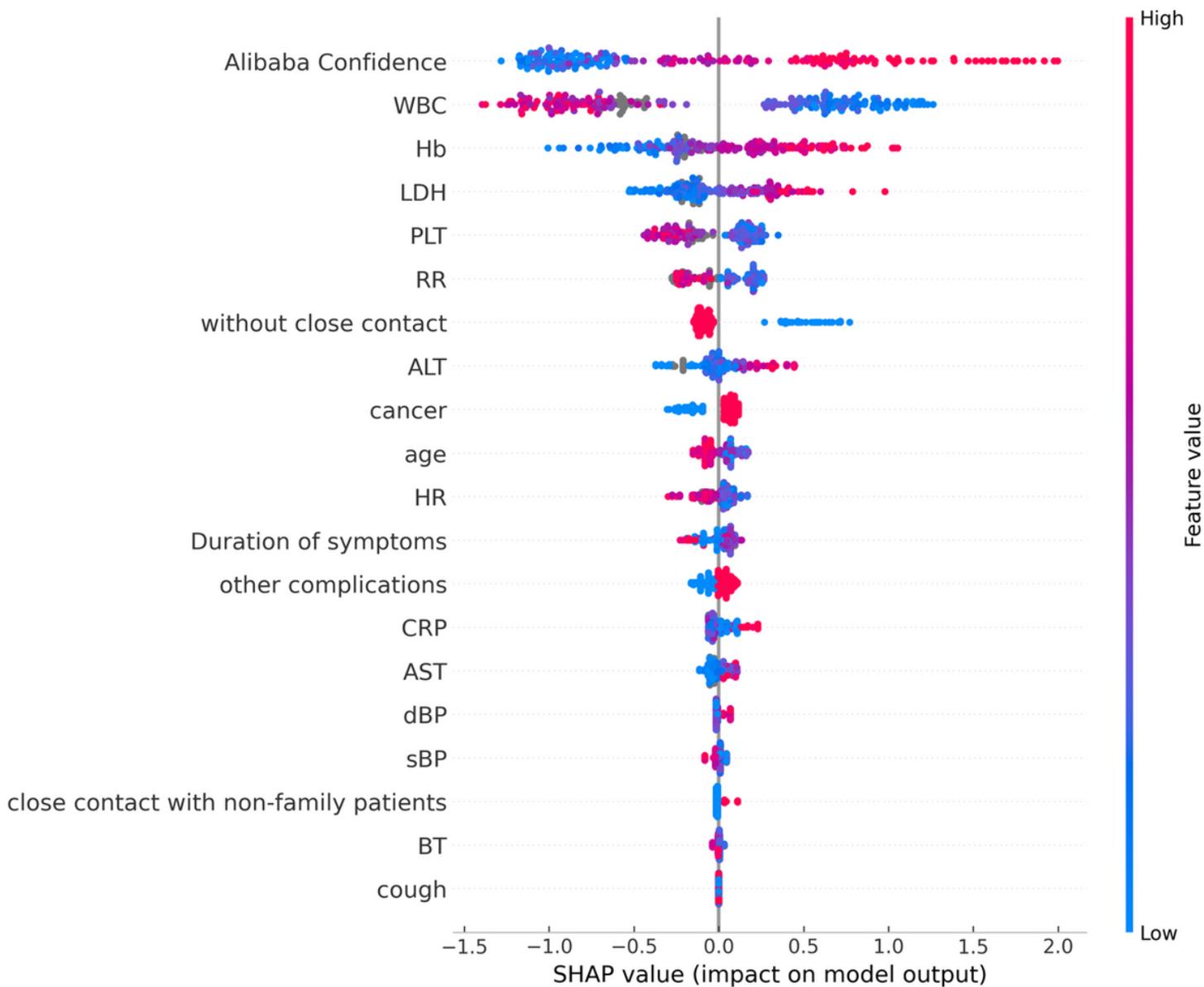


Figure 2

Variables that demonstrated the greatest association with COVID-19 real-time reverse transcription polymerase chain reaction (RT-PCR) positive in machine learning model using all variables in the test data. WBC: white blood cells, Hb: hemoglobin, LDH: lactate dehydrogenase, PLT: platelet, RR: respiratory rate, ALT: alanine aminotransferase, HR: heart rate, CRP: C-reactive protein, AST: aspartate aminotransferase, dBP: diastolic blood pressure, sBP: systolic blood pressure, BT: body temperature, SHAP: shapley additive explanations

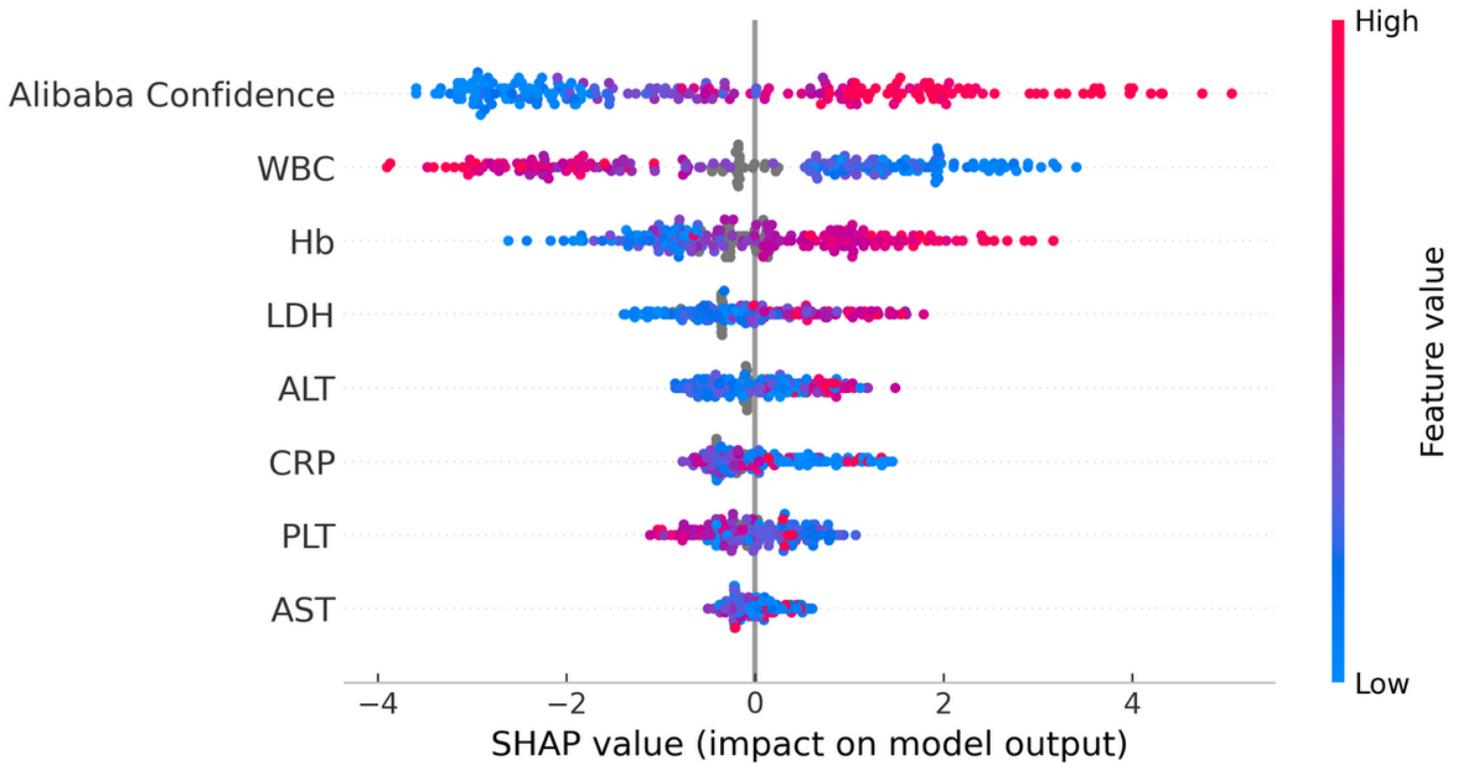


Figure 3

Variables that demonstrated the greatest association with COVID-19 real-time reverse transcription polymerase chain reaction (RT-PCR) positive in machine learning model using 8 variables (Ali-M3 confidence; White blood cell; Hemoglobin; Platelet; Aspartate aminotransferase; Alanine aminotransferase; Lactate dehydrogenase; C-reactive protein) in the test data. WBC: white blood cells, Hb: hemoglobin, LDH: lactate dehydrogenase, ALT: alanine aminotransferase, CRP: C-reactive protein, PLT: platelet, AST: aspartate aminotransferase, SHAP: shapley additive explanations

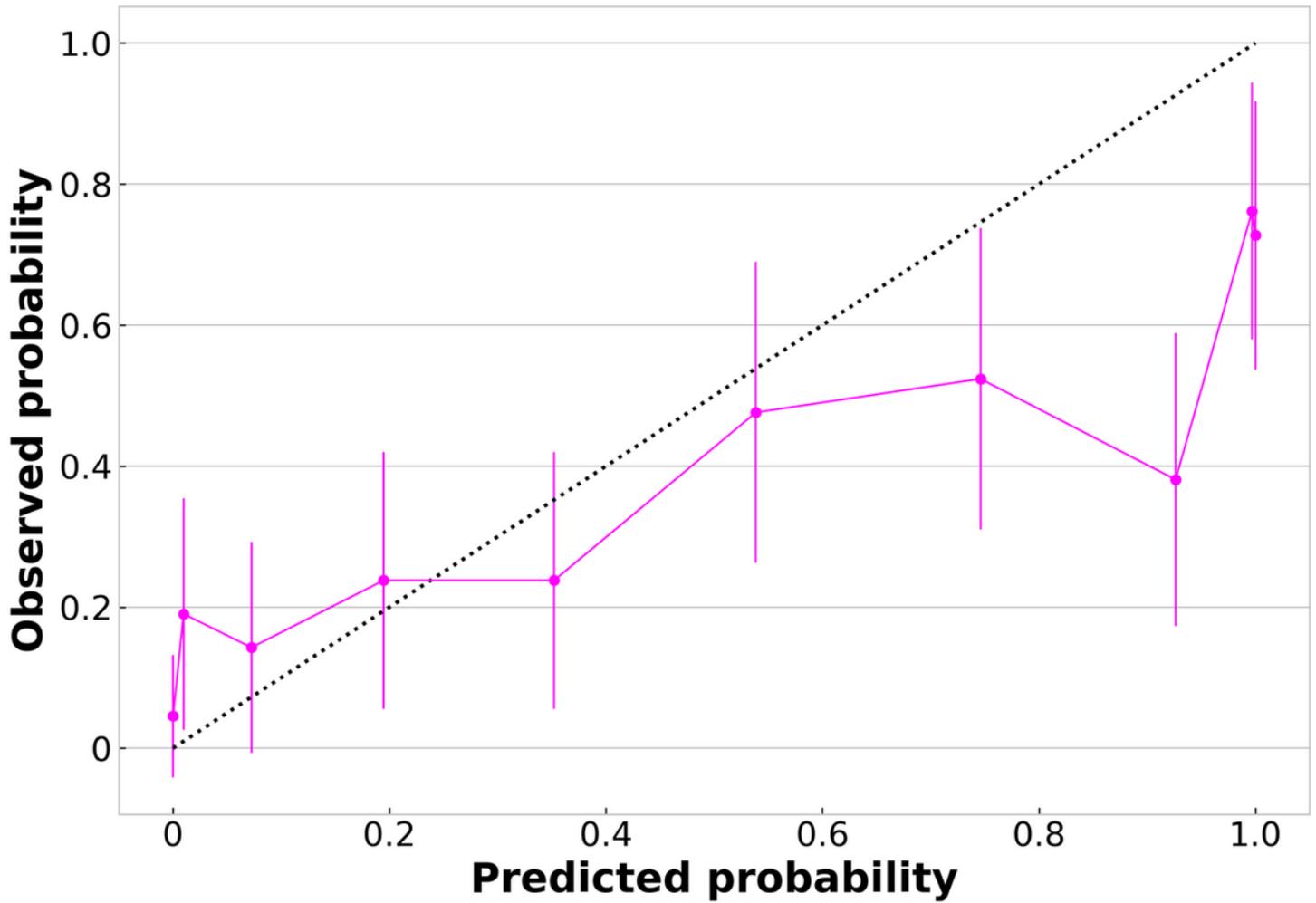


Figure 4

Predicted versus observed probability of COVID-19 real-time reverse transcription polymerase chain reaction (RT-PCR positive (calibration; pink line) for Ali-M3 Confidence.

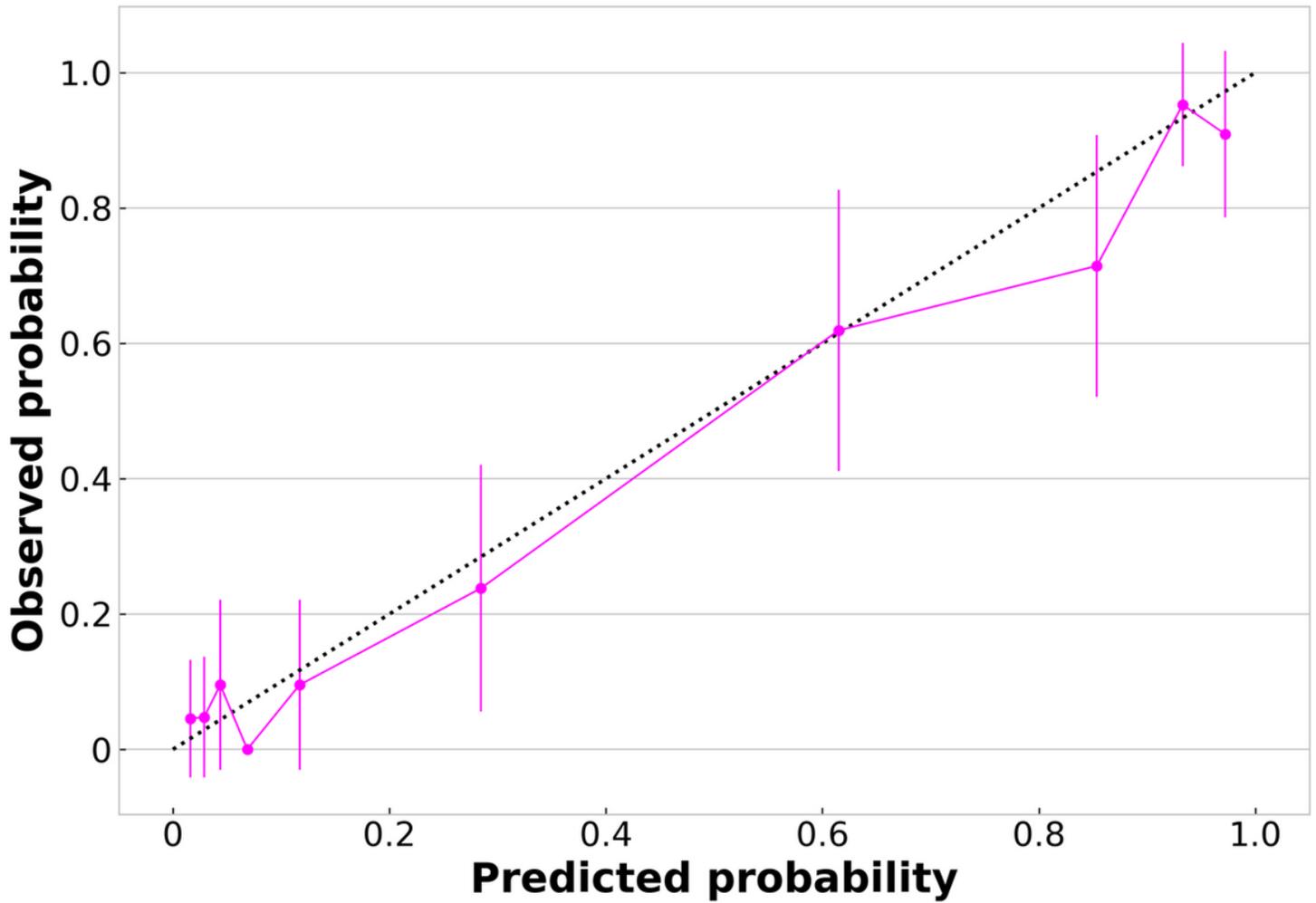


Figure 5

Predicted versus observed probability of COVID-19 real-time reverse transcription polymerase chain reaction (RT-PCR positive (calibration; pink line) for machine learning model using all variables.

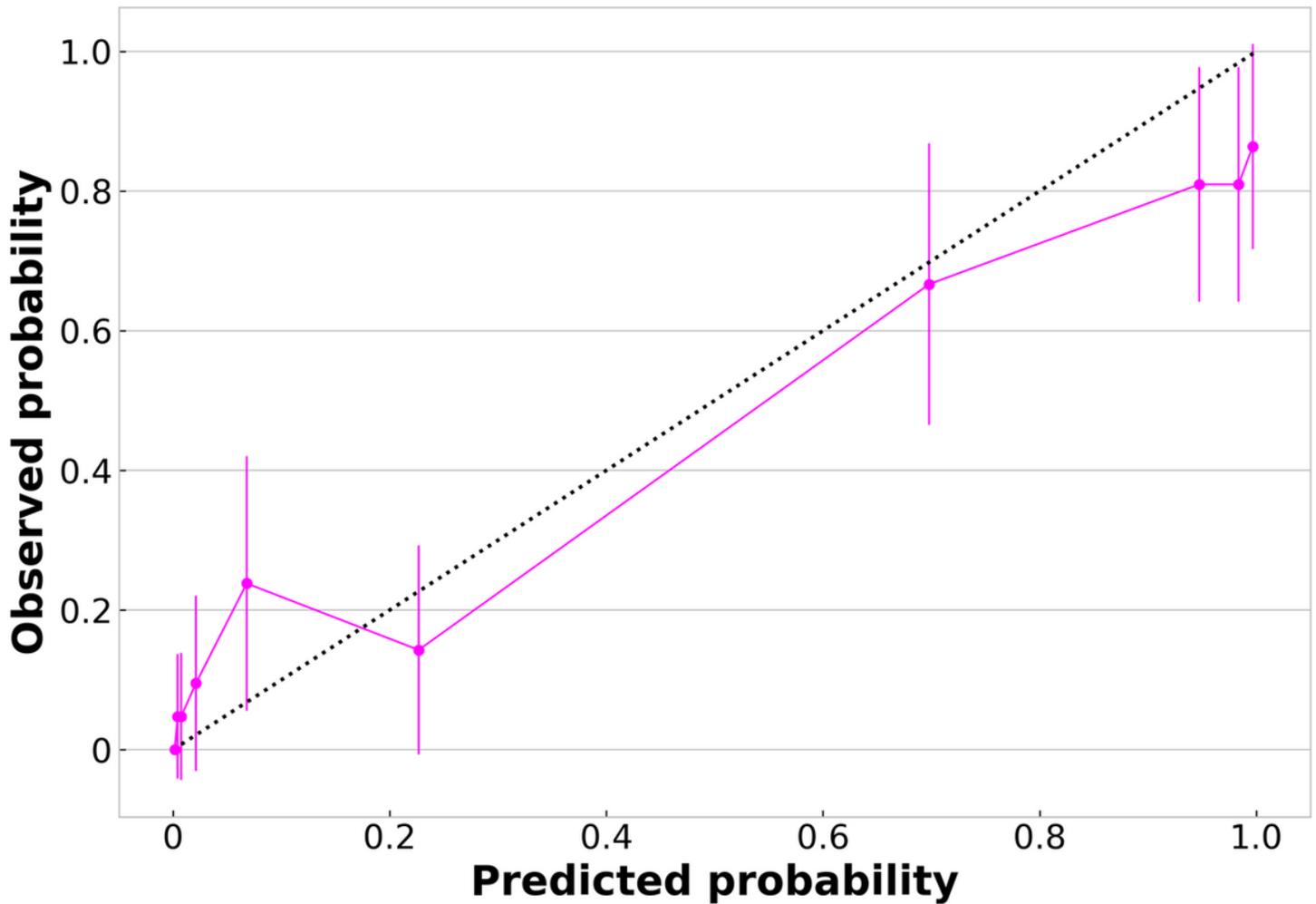


Figure 6

Predicted versus observed probability of COVID-19 real-time reverse transcription polymerase chain reaction (RT-PCR positive (calibration; pink line) for machine learning model using 8 variables (Ali-M3 confidence; White blood cell; Hemoglobin; Platelet; Aspartate aminotransferase; Alanine aminotransferase; Lactate dehydrogenase; C-reactive protein).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TripodChecklistPredictionModelDevelopmentandValidationWord.docx](#)