OXFORD

Precision Clinical Medicine, 4(1), 2021, 62–69

doi: 10.1093/pcmedi/pbaa040
Advance Access Publication Date: 28 December 2020
Clinical Practice

CLINICAL PRACTICE

Clinical longitudinal evaluation of COVID-19 patients and prediction of organ-specific recovery using artificial intelligence

Winston T. Wang[1,2], Charlotte L. Zhang[1], Kang Wei[3], Ye Sang[4], Jun Shen[5], Guangyu Wang[6,*] and Alexander X. Lozano[1,*]

[1]Department of Materials Science & Engineering, Stanford University, Stanford, CA 94305, USA
[2]Mayo Clinic Alix school of Medicine, Scottsdale, AZ 85259, USA
[3]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510623, China
[4]The First College of Clinical Medical Science, China Three Gorges University, Yichang 443000, China
[5]Department of Radiology, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, China
[6]School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

*Correspondence: Guangyu Wang, guangyu.wang24@gmail.com; Alexander X. Lozano, axlozano@stanford.edu

Abstract

Within COVID-19 there is an urgent unmet need to predict at the time of hospital admission which COVID-19 patients will recover from the disease, and how fast they recover to deliver personalized treatments and to properly allocate hospital resources so that healthcare systems do not become overwhelmed. To this end, we have combined clinically salient CT imaging data synergistically with laboratory testing data in an integrative machine learning model to predict organ-specific recovery of patients from COVID-19. We trained and validated our model in 285 patients on each separate major organ system impacted by COVID-19 including the renal, pulmonary, immune, cardiac, and hepatic systems. To greatly enhance the speed and utility of our model, we applied an artificial intelligence method to segment and classify regions on CT imaging, from which interpretable data could be directly fed into the predictive machine learning model for overall recovery. Across all organ systems we achieved validation set area under the receiver operator characteristic curve (AUC) values for organ-specific recovery ranging from 0.80 to 0.89, and significant overall recovery prediction in Kaplan-Meier analyses. This demonstrates that the synergistic use of an artificial intelligence (AI) framework applied to CT lung imaging and a machine learning model that integrates laboratory test data with imaging data can accurately predict the overall recovery of COVID-19 patients from baseline characteristics.

Key words: artificial intelligence; medical image; COVID-19; longitudinal evaluation

Received: 28 November 2020; **Revised:** 20 December 2020; **Accepted:** 21 December 2020

Introduction

The outbreak of the 2019 novel coronavirus (SARS-CoV-2) began in December 2019[1,2] and has since rapidly spread across the globe,[3] causing the World Health Organization (WHO) to officially declare COVID-19 a pandemic. Non-specific findings associated with COVID-19 can include fever, cough, and other flu-like symptoms.[4–6] A significant proportion of COVID-19 patients progress into severe acute respiratory syndrome, which is generally treated with mechanical ventilation and intensive care unit (ICU) admission.[7–9]

Both computed tomography (CT) images and laboratory tests have been applied to diagnose COVID-19 during the current pandemic.[10–12] Significant findings on chest CT scan and/or a positive viral RT-PCR test are used to indicate infection.[13,14] On CT scans, COVID-19 patients classically develop ground glass opacity (GGO) lesions or consolidation lesions, which can present clinically as shortness of breath and decreased oxygen saturations. Laboratory values such as creatinine and potassium are often informative and aid clinicians in their diagnoses, but independently do not capture the entire clinical picture.

It has been shown that COVID-19 patients can develop serious liver, kidney, and cardiac injuries[15] that may lead to further long-term damage and functional deterioration.[16,17] Moreover, little is known about the organ-specific recovery of COVID-19 symptoms after hospital discharge.[18] In general, lung lesion sizes on a CT scan of COVID-19 patients increase from their initial hospital admission to the beginning of remission,[19,20] but the longitudinal evolution of their lesions and clinical presentation remains unclear.

It would be of great clinical utility to characterize and predict recovery within each organ system impacted by COVID-19 to better personalize therapies, and to efficiently allocate hospital resources based on need. Moreover, the accurate prediction of organ-specific clinical outcomes is important to flag potential future complications, and thus plan relevant and appropriate longitudinal follow-up. To this end, we investigated the impact of COVID-19 patients' initial CT scans and blood draws on future organ function scores to establish a predictive model for organ-specific recovery from baseline characteristics.

Results and discussion

Cohort statistics

The general scheme of our study design and procedures are described in Fig. 1. We prospectively followed 285 COVID-19 patients from January to July of 2020. Patients' characteristics are presented in Table 1. These patients were followed up monthly after their initial discharge from the Yichang Central Hospital where they were first diagnosed with COVID-19. If they agreed to participate in this longitudinal study, they were enrolled and consented to use of their anonymized medical record data, including demographic information, lifestyle (including smoking and alcohol use), routine physical examination, and clinical laboratory data. We then analyzed these data and established an artificial intelligence (AI) model to predict organ functional recovery (Fig. 1). Demographics and clinical parameters of all the study subjects are provided in Table 1, and the association of each clinical parameter with critical illness is quantified with a t-test *P* value for each.

Prediction of future organ function

We trained five organ-specific machine learning models that synergistically used AI-analyzed baseline CT scan data combined with laboratory testing values to predict organ recovery for the renal, pulmonary, immune, hepatic, and cardiac organ systems, where baseline data are defined as the first laboratory and imaging data obtained after admission. The models predicted whether a patient would recover at any point during their follow-up, including after hospital discharge. To ensure that the models were not overfit and that they were generalizable, we then applied these models into a held–out validation set of 87 patients. We achieved strong AUCs of 0.89 (95% CI = 0.82–0.94) for the immune model and 0.83 (95% CI = 0.74–0.90) for the renal model on held–out validation set data (Fig. 2). The analogous results for the lung, liver, and coagulation models were also significant, with AUCs of 0.84 (95% CI = 0.75–0.90) for lung, 0.81 (95% CI = 0.72–0.89) for liver, and 0.81 (95% CI = 0.72–0.89) for coagulation (Fig. S1).

Kaplan-Meier analysis

To demonstrate the prognostic utility of the organ-specific recovery AI models, we applied our model scores to predict prognosis within the context of a Kaplan-Meier analysis, which is well-suited to the analysis of data with differing follow-up timepoints. We used the models to stratify the patient population into "high model score" and "low model score" groups for each of the organ systems characterized based upon the midpoint of the LightGBM prediction scale, in this case 0.5. The model is trained such that high model scores are associated with organ-specific recovery, while low model scores are associated with non-recovery, where the clinical endpoint of recovery with respect to a given organ is defined as the point in time when the organ function score crosses over from abnormal range to normal range (methods). Both the immune and kidney models have strong performance, with significant stratification of patients with log rank *P* values <0.0001 for each organ-specific overall recovery prediction when applied to the pooled training and validation cohorts (Fig. 3). It is clear from the Kaplan-Meier curves that the patients

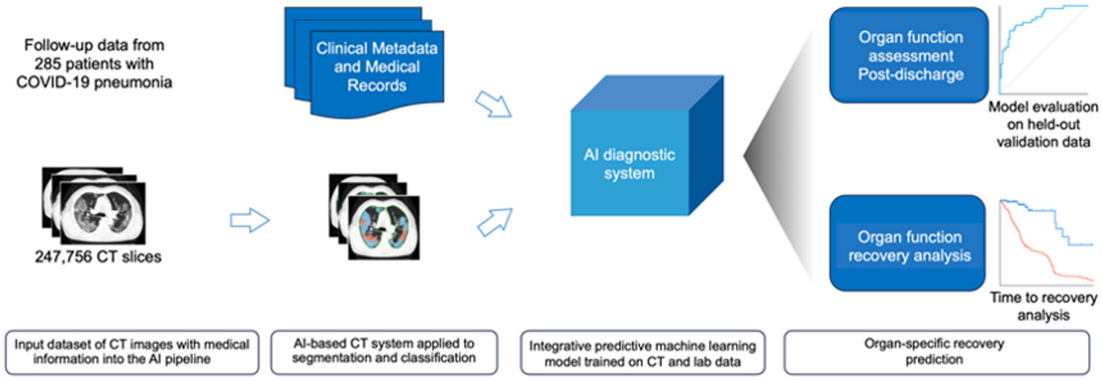

Figure 1. Proposed framework for NCP diagnosis and prognosis prediction. A large CT and metadata dataset was constructed from 285 patients (247 756 CT images from COVID-19 pneumonia). These NCP images were entered into an AI diagnostic system with patients' medical records to generate a quantitative report of lung lesions. We next analyzed lung-lesion features and clinical metadata to evaluate the changing of bodily functions after the initial hospital discharge using a LightGBM model and conduct a prognosis analysis using a Kaplan-Meier curve

who are predicted to recover by the model tend to recover earlier and more consistently than those predicted not to recover. The liver, lung, and coagulation models also demonstrate a statistically significant ability to stratify patients into recovering and non-recovering populations, with the most striking difference coming early in the patients' disease course (Fig. S2).

Cross-correlation matrix of clinical and imaging features

To characterize the relationships between each laboratory test and lesion size, we computed Spearman's rank correlation coefficient across the set of laboratory test values and CT lung imaging feature sizes as determined by the AI-imaging pipeline, and then applied a hierarchical clustering algorithm (hclust(), implemented in R using the complete linkage method) to the results (Fig. S3). We clustered the laboratory and CT feature values such that highly correlated features come together, and highly anti-correlated features move apart.

There are two large clusters that are formed. The first is located in the lower right portion of the figure and comprises features that are positively correlated with larger lung-lesion sizes, and generally indicate a heightened inflammatory state that portends worse clinical outcome. Higher levels of several key features from this data are known in the literature to be associated with severe COVID-19 disease including erythrocyte sedimentation rate,[21] C-reactive protein,[22] lactate dehydrogenase,[23] and blood urea nitrate.[24]

Conversely, the cluster in the top left of Fig. S3 is the feature for which lower values correspond to more favorable outcomes, and smaller lesion size on CT imaging. Several are described within the literature as anti-correlated with COVID-19 severity including sodium,[25] hemoglobin,[26] and albumin.[27]

Discussion

To our knowledge, this study is the first to demonstrate organ-specific recovery prediction from baseline imaging and laboratory values in COVID-19 patients. The integrative machine learning models were shown to generalize to a held-out validation dataset, achieving AUCs > 0.8 for prediction of recovery in each of the organ systems to which they were applied.

In the era of precision healthcare, prediction of recovery using only baseline data is highly valuable to clinicians, as it is critical to understand the longitudinal impacts of COVID-19 tailored to the characteristics of each patient and to their condition. Moreover, prediction of time to patient recovery would enable healthcare systems to better allocate resources ahead of time to better ensure that ICUs, personnel, and resources are distributed in such a way as to avoid the system from reaching capacity and restricting patient care. Finally, each of the organ systems characterized in the present study: renal, pulmonary, immune, hepatic, and cardiac, is essential to a patient's short- and long-term health, and hence insight into their recovery has key consequences for treatment as well as the wellbeing of the patient in the acute and chronic settings.

Given the success of our integrative machine learning model in this held-out validation cohort, the next step to bring it closer to translation to the clinical environment is to independently validate it at a separate medical center, and even in a population in another nation. Moreover, we expect the model to maintain predictive power and accuracy with longer follow-up data, and to this end intend to keep accumulating data.

In sum, COVID-19 can have critical impacts on key organ systems within the body, and there is an unmet need to predict organ-specific recovery from the impacts of this disease to better allocate healthcare resources and provide the patients with highly personalized medical treatments. We have trained and validated in a held-out test set of organ-specific integrative machine learning models that can predict recovery from baseline patient

Table 1. Demographic and clinical statistics of the cohort. Data are mean (range) and *n* (%). The cohort of these 285 patients as COVID-19 pneumonia is stratified into groups of critically ill and non-critically ill patients. *P* values indicated comparison of critically ill patients versus non-critically ill patients and were calculated by independent-samples t test via SPSS 22.0.

	Total		Non-critically ill			Critically ill			P value
Variables	sample size	mean (min-max)	sample size	%	mean (min-max)	sample size	%	mean (min-max)	t test
Age (Years)	285	49 (0–91)	101	48.8	48.1 (17–77)	184	51.2%	49.5 (0–91)	0.542
Sex									
Men	101	139 (48.8%)	51	50.5		88	47.8		
Women	184	146 (51.2%)	50	49.5		96	52.2		
Hemoglobin (g/L)	269	133.3 (69–195)	99	36.8	137.4 (89–177)	170	63.2	130.99 (69–195)	0.003**
Total protein (g/L)	218	63.5 (50.9–90.5)	62	28.4	64.4 (54.8–78.0)	156	71.6	63.2 (50.1–90.5)	0.132
Albumin (g/L)	218	36.1 (24.3–49.5)	62	28.4	37.8 (27.2–46.1)	156	71.6	35.5 (24.3–49.5)	0.001**
Creatinine (μmol/L)	282	75.9 (18–1017.9)	100	35.5	65.5 (39.8–138.4)	182	64.5	81.7 (18–1017)	0.033*
Lactate dehydrogenase (U/L)	278	235.7 (97–559)	100	36.0	204.6 (119–358)	178	64.0	253.1 (97–559)	0.000***
C-reactive protein (mg/L)	272	32.1 (0.1–333.1)	99	36.4	19.7 (0.1–130.5)	173	63.6	39.2 (0.1–333.1)	0.000***
Creatine kinase (IU/L)	278	146.1 (13–1849)	100	36.0	129.4 (13–871))	178	64.0	155.4 (20–1849)	0.283
Creatine kinase myocardial band (IU/L)	278	12.6 (3–165)	100	36.0	11.0 (4–19)	178	64.0	13.4 (3–165)	0.019*
Potassium (IU/L)	282	3.8 (2.6–5.8)	100	35.5	3.8 (2.6–4.6)	182	64.5	3.8 (2.9–5.8)	0.579
Sodium (IU/L)	282	137.9 (124.5–147.7)	100	35.5	139.0 (131.7–147.7)	182	64.5	137.3 (124.5–144.3)	0.000***
Erythrocyte sedimentation rate (mm/h)	267	31.2 (0.5–126)	99	37.1	25.9 (0.5–105)	168	62.9	34.3 (0.5–126)	0.012*
Globulin (g/L)	218	27.4 (17.7–49.4)	62	28.4	26.6 (19.6–36.9)	156	71.6	27.7 (17.7–49.4)	0.144
Blood urea nitrate (mmol/L)	282	4.6 (1.8–26.9)	100	35.5	4.0 (1.8–10.9)	182	64.5	4.9 (1.9–26.9)	0.000***
Kidney score	285	1.04 (0–4)	101	35.5	0.69 (0–3)	184	64.5	1.22 (0–4)	0.000***
Liver score	285	1.28 (0–5)	101	35.5	0.89 (0–4)	184	64.5	1.49 (0–5)	0.000***
Pulmonary score	285	1.93 (0–4)	101	35.5	1.42 (0–3)	184	64.5	2.21 (0–4)	0.000***
Immune score	285	2.23 (0–5)	101	35.5	1.96 (0–5)	184	64.5	2.38 (0–5)	0.011*
Coagulation score	285	1.75 (0–4)	101	35.5	1.50 (0–3)	184	64.5	1.88 (0–4)	0.005**

*https://www.graphpad.com/support/faq/what-is-the-meaning-of–or–or–in-reports-of-statistical-significance-from-prism-or-instat/#:~:text=NEJM%20(New%20England%20Journal%20of,than%200.001%20shown%20as%20%22%3C%20.&text=0001%22.-,P%20values%20less%20than%200.001%20are%20summarized%20with%20three%20asterisks,are%20summarized%20with%20four%20asterisks.

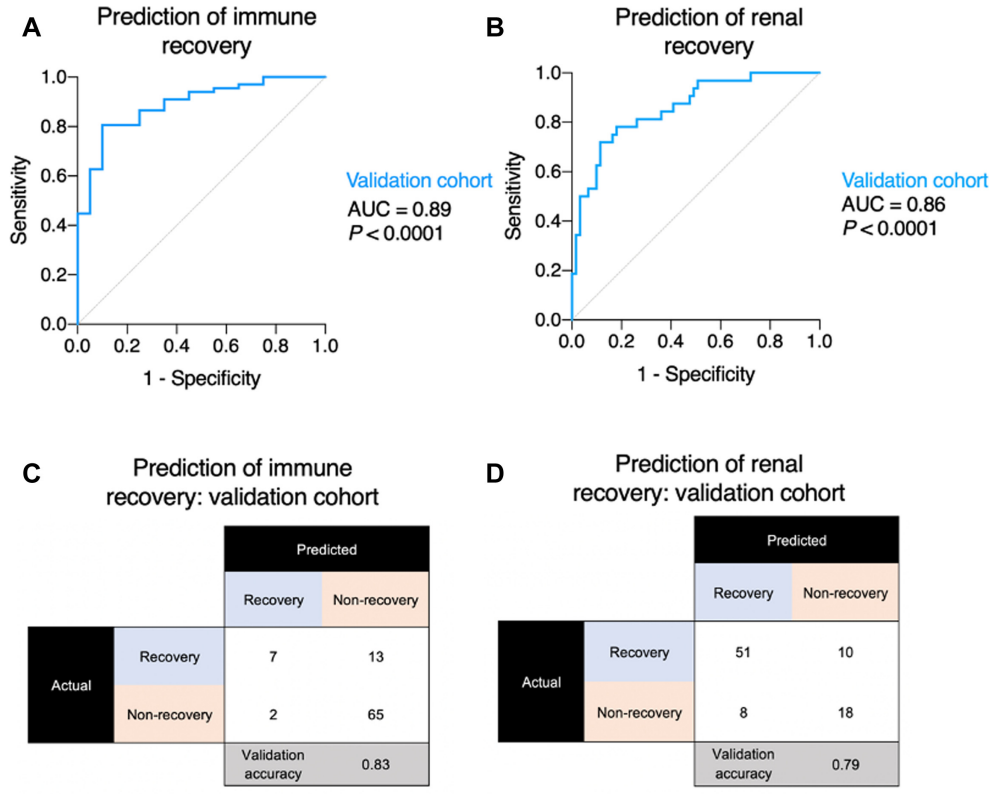

Figure 2. ROC curves and confusion matrices for immune and renal systems. (A) The immune system model resulted in an AUC of 0.89. (B) The renal system model resulted in an AUC of 0.83. (C) The confusion matrix for the immune system model on the validation set, with an accuracy of 0.83. (D) The confusion matrix for the renal system model on the validation set, with an accuracy of 0.79.

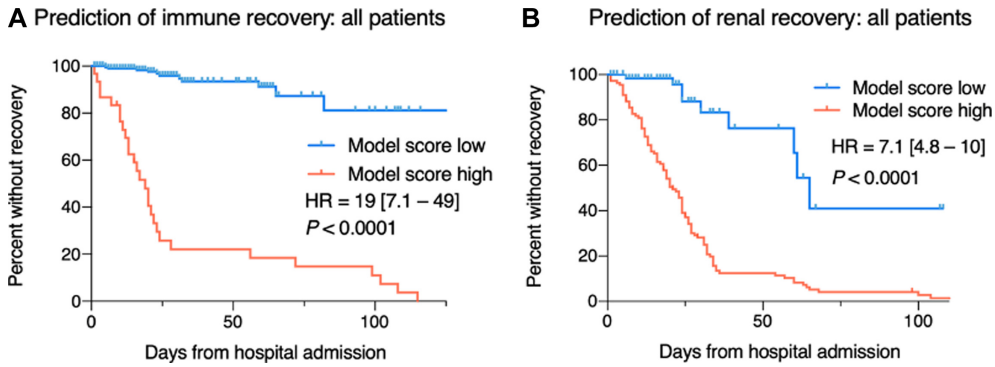

Figure 3. Kaplan-Meier curves to assess patient recovery. Kaplan-Meier curves for the (A) immune and (B) renal systems. The patient population was stratified into high and low model scores based on a cutpoint of 0.5, the general midpoint of the lightGBM model output range.

imaging and laboratory testing upon hospital admission and demonstrated AUCs > 0.8 across each organ system to which the model is applied.

Methods

Study design and population

This was a retrospective observational study performed at the Yichang Central Hospital in Yichang, Hubei, China. We obtained 247 756 deidentified individual CT image slices from 285 patients (Table 1). Patients were eligible for recruitment if they 1) were confirmed as novel coronavirus pneumonia (NCP) patients, 2) had one or more clinical visits to the Yichang Central Hospital, 3) had CT imaging completed and 4) had comprehensive blood tests completed. Patients were excluded if they did not meet any single criterion of recruitment above.

CT imaging data collection

CT scans from the 285 patients were collected longitudinally both before and after their hospital discharge (Fig. 1). Scans were acquired using a Siemens CT scanner with 2–3 mm thick slices. To ensure that these images were accurate and usable, first-pass screening was done on all images to filter out low-quality, unreadable, and artifact-heavy scans. These scans were then fed into

a segmentation and classification algorithm[19] to calculate the volume of different lung lesions such as ground glass opacity (GGO), consolidation, interstitial thickening, pleural effusion, and fibrosis lesions.

Patient metadata collection

Demographic data and a wide range of laboratory value data were collected from each patient. Blood tests were collected many times for each patient, although the specific timing and number of blood tests were variable. Laboratory values that were collected include albumin, C-reactive protein, lactate dehydrogenase, and other salient values (Table 1).[19]

Integrative machine learning model

We applied the LightGBM machine learning model[28–30] architecture, a tree-based model, to predict post-discharge outcomes in our patient cohort. Because it is based on decision trees, the LightGBM model is capable of dealing with missing values (NaN values) from missing tests. This allowed us to include the maximum number of laboratory tests without leaving out any patients because of missing data.

When training the LightGBM models, we used the Python function "random.shuffle" to split the dataset of 285 patients randomly into separate training and validation sets, with a train to validation ratio of 7:3 while preserving the original critically ill to non-critically ill ratio in both the training and validation sets. Each model was trained separately on the training data and had hyperparameters fine-tuned using the Optuna algorithm for 1000 iterations.[31]

Prognosis prediction

Patients were categorized by hospital physicians into critically ill and non-critically ill cohorts, with ICU admission or need for mechanical ventilation as the criteria for the critically ill cohort.

Organ function calculation

Each organ or system that we examined in this study (liver, lung, kidney, immune, and coagulation) was associated with a set of laboratory values (CT scan values for lung) that we observed, which were available to us in the study. For example, the set of laboratory values associated with the liver are blood urea nitrogen (BUN),[32–34] alanine transaminase (ALT), aspartate transaminase (AST), albumin, and direct bilirubin. For kidneys, we associated BUN, estimated glomerular filtration rate (eGFR), potassium, sodium, and bicarbonate. For immune, we used erythrocyte sedimentation rate (ESR), lactate dehydrogenase (LDH), C-reactive protein (CRP), white blood cell count, and neutrophil count. For coagulation, we used platelet count, partial thromboplastin time (PTT), prothrombin time (PT), and D-dimer. For pulmonary, we used the sizes of the five lesion types output by the segmentation algorithm: ground glass opacity (GGO), interstitial thickening, pleural effusion, pulmonary cavity, and fibrosis. We then created organ function scores for each of the organs using their associated laboratory values, modeled after peer-reviewed equations such as the multiple organ dysfunction (MODS) score,[35] SOFA score for sepsis,[36] and the MELD score for liver disease.[37] In all three clinical scoring systems, several markers of disease are aggregated linearly to result in a single score. Similarly in our case, if any of a patient's laboratory values were determined to be abnormal as per guidelines in the UpToDate clinical decision resource,[38] that patient's organ function score would increase by 1. Therefore, in the example above, a patient could have a minimum score of 0, and a maximum score of 5. As there was no strong rationale to treat organ markers differently, we weighted them equally within the organ score function. In addition, we did not have sufficient data to stratify organ function classification with more granularity than "normal" and "abnormal". We further considered a score of 0 to be normal or recovered, and a score of 1 or above to be abnormal. We then defined a patient as recovered for a given organ when the score for that organ reached zero, i.e. all the associated lab values for the organ have returned to normal. For example, patients would be considered "liver recovered" if they had follow-up lab tests with a liver score of 0. Additionally, we estimated the day that the patient recovered as the number of days after admission when the first score of 0 blood test was taken.

Statistics

To assess the model's performance for each classification task accuracy, receiver operator characteristic curve (ROC) analysis was applied. ROC curves are generated by plotting the true positive rate (sensitivity) versus the false-positive rate (1–specificity) across different cut points for the output machine learning model score. An AUC value of 1 indicates perfect performance, whereas an AUC approaching 0.5 indicates performance equivalent to random chance. Sensitivity, specificity, and accuracy were determined using a cutpoint model score of 0.5.

The organ-specific trained machine learning models were used to stratify patients into two populations, which were allocated into high- and low-risk groups in a Kaplan-Meier analysis for overall recovery. In this analysis the traditional Kaplan-Meier analysis method is applied; however, recovery is used instead of the canonical overall survival clinical end-point.

In addition, we applied Spearman's rank correlation coefficient to analyze the relationships between quantitative lesion features on CT imaging as ascertained by the AI image analysis workflow in addition to laboratory test values. We also provided descriptive statistics such as the results of two-tailed t-tests to quantify the difference between critically ill and non-critically ill patients at the time of hospital admission.

We applied the Python scikit-learn library and Prism version 8 to generate plots and conduct all other statistical analyses. The measurements of sensitivity, specificity, and accuracy were calculated by Python scikit-learn library. Unless noted otherwise, all experiments were performed with separate training and validation sets (7:3 train: validation ratio).

Study approval

All radiographic and laboratory tests were performed as a part of patients' routine clinical care, including CT images. Institutional Review Board (IRB)/Ethics Committee approvals were obtained (IRB number 2020-KY-010, Sun Yat-Sen University Memorial Hospital). The work was conducted in compliance with the United States Health Insurance Portability and Accountability Act (HIPAA), the tenets of the Declaration of Helsinki, the Chinese CDC policy on reportable infectious diseases, and Chinese Health and Quarantine Law.

Supplementary data

Supplementary data are available at *PCMEDI Journal* online.

Conflict of interest

None declared.

Funding

This study was funded by the National Natural Science Foundation of China (Grant No. 61906105). Alexander Lozano received funding from the Natural Sciences and Engineering Research Council of Canada (NSERC); however, this funding was for his graduate education at Stanford University and had no specific tie to this project or bearing on the project. Winston Wang received funding from the C.L. and Patty M. Pecchenino Family Scholarship Fund for Mayo Clinic Alix School of Medicine. This funding was for his graduate education at Mayo Clinic and has no bearing on this project.

Author contributions

WTW, CLZ, AXL, and GW conceived the study. CZ, KW, YS, JS, and GW assisted in curating and obtaining the data. WTW, KW, and AXL performed the analysis. All authors contributed to the manuscript.

Acknowledgements

The authors extend their gratitude to Prof. Shan X. Wang of Stanford University for many productive discussions about the project. Charlotte Zhang performed a summer internship at Stanford University during which she participated in this project.

References

1. Munster VJ, Koopmans M, van Doremalen N, *et al.* A novel coronavirus emerging in China — Key questions for impact assessment. *N Engl J Med* 2020;**382**:692–4. doi:10.1056/NEJMp2000929.
2. Wang C, Horby PW, Hayden FG, *et al.* A novel coronavirus outbreak of global health concern. *Lancet North Am Ed* 2020;**395**:470–3. doi:10.1016/S0140-6736(20)30185-9.
3. Phan LT, Nguyen TV, Luong QC, *et al.* Importation and Human-to-Human transmission of a novel coronavirus in Vietnam. *N Engl J Med* 2020;**382**:872–4. doi:10.1056/NEJMc2001272.
4. Chan JF-W, Yuan S, Kok K-H, *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet North Am Ed* 2020;**395**:514–23. doi:10.1016/S0140-6736(20)30154-9.
5. Zhu N, Zhang D, Wang W, *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;**382**:727–33. doi:10.1056/NEJMoa2001017.
6. Lechien JR, Chiesa-Estomba CM, De Siati DR, *et al.* Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *Eur Arch Otorhinolaryngol* 2020;**277**:2251–61. doi:10.1007/s00405-020-05965-1.
7. Guan W-j, Ni Z-y, Hu Y, *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;**382**:1708–20. doi:10.1056/NEJMoa2002032.
8. Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet North Am Ed* 2020;**395**:497–506. doi:10.1016/S0140-6736(20)30183-5.
9. Yang X, Yu Y, Xu J, *et al.* Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020;**8**:475–81. doi:10.1016/S2213-2600(20)30079-5.
10. Chen A, Karwoski RA, Gierada DS, *et al.* Quantitative CT analysis of diffuse lung disease. *Radiographics* 2020;**40**:28–43. doi:10.1148/rg.2020190099.
11. Kermany DS, Goldbaum M, Cai W, *et al.* Identifying medical diagnoses and treatable diseases by Image-Based deep learning. *Cell* 2018;**172**:1122–31. e9. doi:10.1016/j.cell.2018.02.010.
12. Huang P, Liu T, Huang L, *et al.* Use of Chest CT in Combination with Negative RT-PCR Assay for the 2019 Novel Coronavirus but High Clinical Suspicion. *Radiology* 2020;**295**:22–3. doi:10.1148/radiol.2020200330.
13. Carter LJ, Garner LV, Smoot JW, *et al.* Assay techniques and test development for COVID-19 diagnosis. *ACS Cent Sci* 2020;**6**:591–605. doi:10.1021/acscentsci.0c00501.
14. Xie X, Zhong Z, Zhao W, *et al.* Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: Relationship to negative RT-PCR testing. *Radiology* 2020;**296**:E41–E5. doi:10.1148/radiol.2020200343.
15. Puntmann VO, Carerj ML, Wieters I, *et al.* Outcomes of cardiovascular magnetic resonance imaging in patients recently recovered from coronavirus disease 2019 (COVID-19). *JAMA Cardiol* 2020;**5**:1265–73. doi:10.1001/jamacardio.2020.3557.
16. Chen N, Zhou M, Dong X, *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneu-

monia in Wuhan, China: a descriptive study. *Lancet North Am Ed* 2020;**395**:507–13. doi:10.1016/S0140-6736(20)30211-7.
17. Carfì A, Bernabei R, Landi F, *et al*. Persistent symptoms in patients after acute COVID-19. *JAMA* 2020;**324**:603–5. doi:10.1001/jama.2020.12603.
18. Del Rio C, Collins LF, Malani P. Long-term health consequences of COVID-19. *JAMA* 2020, (Online ahead of print). doi:10.1001/jama.2020.19719.
19. Zhang K, Liu X, Shen J, *et al*. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 2020;**181**:1423–33. e11. doi:10.1016/j.cell.2020.04.045.
20. Shi H, Han X, Jiang N, *et al*. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis* 2020;**20**:425–34. doi:10.1016/S1473-3099(20)30086-4.
21. Lapic I, Rogic D, Plebani M. Erythrocyte sedimentation rate is associated with severe coronavirus disease 2019 (COVID-19): a pooled analysis. *Clin Chem Lab Med* 2020;**58**:1146–8. doi:10.1515/cclm-2020-0620.
22. Huang I, Pranata R, Lim MA, *et al*. C-reactive protein, procalcitonin, D-dimer, and ferritin in severe coronavirus disease-2019: a meta-analysis. *Ther Adv Respir Dis* 2020;**14**:1753466620937175. doi:10.1177/1753466620937175.
23. Henry BM, Aggarwal G, Wong J, *et al*. Lactate dehydrogenase levels predict coronavirus disease 2019 (COVID-19) severity and mortality: A pooled analysis. *Am J Emerg Med* 2020;**38**:1722–6. doi:10.1016/j.ajem.2020.05.073.
24. Cheng A, Hu L, Wang Y, *et al*. Diagnostic performance of initial blood urea nitrogen combined with D-dimer levels for predicting in-hospital mortality in COVID-19 patients. *Int J Antimicrob Agents* 2020;**56**:106110. doi:10.1016/j.ijantimicag.2020.106110.
25. Post A, Dullaart RPF, Bakker SJL. Sodium status and kidney involvement during COVID-19 infection. *Virus Res* 2020;**286**:198034. doi:10.1016/j.virusres.2020.198034.
26. Hariyanto TI, Kurniawan A. Anemia is associated with severe coronavirus disease 2019 (COVID-19) infection. *Transfus Apher Sci* 2020;**59**:102926. doi:10.1016/j.transci.2020.102926.
27. Huang W, Li C, Wang Z, *et al*. Decreased serum albumin level indicates poor prognosis of COVID-19 patients: hepatic injury analysis from 2,623 hospitalized cases. *Sci China Life Sci* 2020;**63**:1678–87. doi:10.1007/s11427-020-1733-4.
28. Ke G, Meng Q, Finley T, *et al*. Lightgbm: A highly efficient gradient boosting decision tree. *Adv Neural Information Processing Syst* 2017:1279–87.
29. Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. *arXiv preprint* 2018, arXiv:1802.03888.
30. Zeng H, Yang C, Zhang H, *et al*. A LightGBM-Based EEG analysis method for driver mental states classification. *Comput Intell Neurosci* 2019;**2019**:3761203. doi:10.1155/2019/3761203.
31. Akiba T, Sano S, Yanase T, *et al*. Optuna: A Next-generation hyperparameter optimization framework. *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 2019:2623–31. doi:10.1145/3292500.3330701.
32. Alvarez Pérez JA, González JJ, Baldonedo RF, *et al*. Clinical course, treatment, and multivariate analysis of risk factors for pyogenic liver abscess. *Am J Surg* 2001;**181**:177–86. doi:10.1016/S0002-9610(00)00564-X.
33. Liu X, Zhang H, Liang J. Blood urea nitrogen is elevated in patients with non-alcoholic fatty liver disease. *Hepatogastroenterology* 2013;**60**(122):343–5. doi:10.5754/hge11216.
34. Tarhini AA, Zahoor H, Lin Y, *et al*. Baseline circulating IL-17 predicts toxicity while TGF-beta1 and IL-10 are prognostic of relapse in ipilimumab neoadjuvant therapy of melanoma. *J Immunother Cancer* 2015;**3**:39. doi:10.1186/s40425-015-0081-1.
35. Marshall JC, Cook DJ, Christou NV, *et al*. Multiple organ dysfunction score: A reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995;**23**(10):1638–52. doi:10.1097/00003246-199510000-00007.
36. Singer M, Deutschman CS, Seymour CW, *et al*. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;**315**:801–10. doi:10.1001/jama.2016.0287.
37. Wiesner R, Edwards E, Freeman R, *et al*. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003;**124**:91–6. doi:10.1053/gast.2003.50016.
38. Post TWed. *UpToDate*. Waltham, MA: UpToDate; 2020, URL: https://www.uptodate.com/contents/laboratory-test-reference-ranges-in-adults.