Radiology

Manuscript type: Original research

Model-based Prediction of Critical Illness in Hospitalized Patients with COVID-19

*S. Schalekamp[1,3], *M. Huisman[1], R. A. van Dijk[2], M.F. Boomsma[2], P.J. Freire Jorge[2], W.S de Boer[4], G.J.M. Herder[5], M. Bonarius[5], O.A. Groot[6], E. Jong[7], A. Schreuder[3], C.M. Schaefer-Prokop[1,3]

[1] Department of Radiology, Meander Medisch Centrum, Amersfoort, the Netherlands

[2] Department of Radiology, Isala, Zwolle, the Netherlands

[3] Department of Radiology, Nuclear Medicine, and Anatomy, Radboud University Medical Center, Nijmegen, the Netherlands

[4] Department of Internal Medicine, Pulmonology, Isala, Zwolle, the Netherlands

[5] Department of Internal Medicine, Pulmonology, Meander Medisch Centrum, Amersfoort, the Netherlands

[6] Department of Internal Medicine, Intensive Care Medicine, Meander Medisch Centrum, Amersfoort, the Netherlands

[7] Department of Internal Medicine, Infectiology, Meander Medisch Centrum, Amersfoort, the Netherlands

*S.S. and M.H. contributed equally to this work.

Corresponding author:

S. Schalekamp

Meander Medical Center

Maatweg 3

3813 TZ Amersfoort

The Netherlands

steven.schalekamp@gmail.com

Key Results

1. A model using baseline patient characteristics, laboratory markers, and chest radiography can predict short-term critical illness in hospitalized patients with COVID-19, with an internally validated AUC = 0.77.
2. At an example model risk threshold of 0.70, 71 of 356 patients would be predicted to develop critical illness of which 59 (83%) would be true-positives.
3. A risk calculator has been made available for download: Dutch COVID-19 risk model (https://docs.google.com/spreadsheets/d/1eFrdHxnOA-M_P-ijxnC2u30qk7IhMVV6YvHvJhrZ8Ws/edit#gid=0) (see Appendix E2).

Summary Statement

Chest radiographic findings, together with patient history and laboratory values, predicted critical illness in hospitalized patients with COVID-19.

Abbreviations:

AUC = area under the ROC curve

CI = confidence interval

COVID-19 = coronavirus disease 2019

ICU = intensive care unit

OR = odds ratio

ROC = receiver operating characteristics

RT-PCR = real-time reverse transcription-polymerase chain reaction

Abstract

Background

The prognosis of hospitalized patients with severe coronavirus disease 2019 (COVID-19) is difficult to predict, while the capacity of intensive care units (ICUs) is a limiting factor during the peak of the pandemic and generally dependent on a country's clinical resources.

Purpose

To determine the value of chest radiographic findings together with patient history and laboratory markers at admission to predict critical illness in hospitalized patients with COVID-19.

Material and Methods

In this retrospective study including patients from 7th March 2020 to 24th April 2020, a consecutive cohort of hospitalized patients with RT-PCR-confirmed COVID-19 from two large Dutch community hospitals was identified. After univariable analysis, a risk model to predict critical illness (i.e. death and/or ICU admission with invasive ventilation) was developed, using multivariable logistic regression including clinical, CXR and laboratory findings. Distribution and severity of lung involvement was visually assessed using an 8-point scale (chest radiography score). Internal validation was performed using bootstrapping**.** Performance is presented as an area under the receiver operating characteristic curve (AUC). Decision curve analysis was performed, and a risk calculator was derived.

Results

The cohort included 356 hospitalized patients (69 ±12 years, 237 male) of whom 168 (47%) developed critical illness. The final risk model's variables included gender, chronic obstructive lung disease, symptom duration, neutrophil count, C-reactive protein level, lactate dehydrogenase level, distribution of lung disease and chest radiography score at hospital presentation. The AUC of the

model was 0.77 (95% CI: 0.72-0.81, P < .001). A risk calculator was derived for individual risk assessment; Dutch COVID-19 risk model (see Appendix E2). At an example threshold of 0.70, 71 of 356 patients would be predicted to develop critical illness of which 59 (83%) would be true-positives.

Conclusion:

A risk model based on chest radiographic and laboratory findings obtained at admission was predictive of critical illness in hospitalized patients with coronavirus disease 2019. This risk calculator might be useful for triage of patients to the limited number of ICU beds/facilities.

Introduction

The worldwide spread of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), leading to coronavirus disease 2019 (COVID-19) caused more than 19 million confirmed cases and over 700,000 deaths up to August 9th, 2020 (1). The majority of patients with COVID-19 present without any symptoms or only mild respiratory symptoms, not requiring hospital admission. However, a proportion of patients are hospitalized because of systemic or severe respiratory symptoms (2-5) of which a subgroup may undergo fast clinical deterioration requiring invasive ventilation and intensive care unit (ICU) treatment, with lethal outcome in the worst-case scenario. This demand for ICU treatment exceeding the options of existing facilities lead to the mobilization of new ICUs, worldwide drastic actions of unprecedented social distancing and different degrees of lockdown to decrease the number of critically ill patients. A number of factors have been described that contribute to poor prognosis. Higher mortality rates occur with increasing age and presence of chronic comorbidities (6-8). A wide spectrum of laboratory values have been linked to poor prognosis and risk of ICU admission (2-4, 6, 7, 9-15).

Although some guidelines do not advise screening of patients suspected of COVID-19 with chest radiography or CT (16, 17), imaging remains important for the diagnosis and risk stratification at admission, as in other types of pneumonia (5, 18-20). In the previous severe acute respiratory syndrome epidemic in 2003 and Middle Eastern respiratory syndrome epidemic in 2012-2015, chest radiography findings were found to correlate with prognosis (21-25).

Publications exploring prognostic markers for COVID-19 thus far largely focused on laboratory findings alone, considered the presence of any abnormality on chest radiography to be a risk factor, or utilized the longitudinal development of disease on thoracic CT. The body of knowledge regarding the complementary information of clinical, laboratory and chest radiography findings is limited. Recently, a Chinese study was published including 1590 patients with COVID-19 of whom 131 (8%)

became critically ill, and therefore needed ICU admission or died. The authors presented a predictive risk-score based on 10 variables, showing good discrimination (AUC = 0.88) (26).

The purpose of this study was to determine the prognostic value of the combined information of patient history, readily available laboratory markers, and chest radiography at admission to predict critical illness, defined as ICU admission for invasive ventilation and/or death, in hospitalized patients with known RT-PCT positive COVID-19.

Materials and Methods

Study population and data collection

In this retrospective cohort-study, inclusion took place in two large community teaching hospitals (Meander Medical Center, Amersfoort, the Netherlands; Isala Hospitals, Zwolle, the Netherlands)during the peak of the SARS-CoV-2 pandemic. The study was approved by the local institutional review boards of both hospitals, and written informed consent was waived (20-036, and 200435).

Using our hospital's COVID-19 registry, all consecutive patients between March 7th and April 24th 2020 suspected of COVID-19 on admission at the emergency department were derived. Patients that did not have a positive real- time reverse transcription polymerase chain reaction (RT-PCR) proven COVID-19 or patients that were not hospitalized were excluded. First RT-PCR tests were taken within 24 hours of hospital admission. If the first test was negative but clinical suspicion remained subsequent tests were performed.

Other exclusion criteria were no chest radiography on admission, transferred patients with uncertain onset of symptoms, status after pneumonectomy, and children (<18 years).

Medical records were reviewed to collect information on patient gender, age, comorbidities (hypertension as diagnosed by a physician, diabetes (either type 1 or type 2 diabetes), history of any type of cancer (present or previous), chronic obstructive lung disease (i.e. COPD or asthma) and cardiac diseases (i.e. ischemic heart disease and/or pre-existent cardiomyopathy)), days of symptoms prior to admission, and temperature at admission (in Celsius). From the blood samples obtained at emergency ward presentation, the following markers were collected: white blood cell (WBC) count, lymphocyte count ($x10^9$/L), neutrophil granulocyte count ($x10^9$/L), C-reactive protein (CRP in mg/L), and lactate dehydrogenase (LDH in IU/L). Neutrophil-to-lymphocyte ratio (NLR) was calculated from the blood samples. Ferritin, procalcitonin, D-dimer and fibrinogen were not part of the standard laboratory work-up at admission in both centers and were not included. Subsequent laboratory tests were not taken into account.

Analysis of the chest radiographs

Chest radiographs obtained within 24 hours after emergency ward presentation were used for analysis. The collected radiographic features included a visual assessment of body habitus (evidently large vs. normal) and the distribution of lung abnormalities into predominantly central (i.e. perihilar), peripheral (i.e. attached to the pleural surface), or diffuse (i.e. both central and peripheral abnormalities). If other non-infectious extensive lung abnormalities, such as large tumors, extensive emphysema or fibrosis, were present in the image, these images were marked as images with pre-existing lung disease. Furthermore, the chest radiograph was divided into 4 zones: right upper lung, left upper lung, right lower lung, left lower lung, using simple anatomic landmarks (Figure 2). Then, the extent of abnormal lung parenchyma was visually scored as 0 = no involvement, 1 = mild/moderate involvement (estimated involvement 0-50% of lung parenchyma), and *2* = severe involvement (estimated involvement >50% of lung parenchyma) per zone resulting in a score between 0 and 8 (chest radiography score). Chest radiographs were independently analyzed by a

chest radiologists (SS and CSP in hospital A, 5 and 25 years of experience, RvD in hospital B, 7 years of experience), blinded to laboratory values and patient outcome. A training session with representative cases was done to ensure a common understanding of the reading task. Image interpretations were done blinded for the clinical outcome, and no double readings were performed.

Outcome

Critical illness was defined by admission to the ICU for invasive ventilation and/or death. Patients admitted to the COVID-19 cohort wards without the need for mechanical ventilation and alive at follow-up are referred to as non-critically ill in this study. The outcome was determined by reviewing the patient's records. Median follow-up time was 34 days (range, 12-53 days); no patients were lost to follow-up.

Statistical analysis

To achieve valid results for the chest radiography score and to reduce overfitting, a power analysis with $1-\beta = 0.8$ ($\alpha=.05$; 2-sided testing) resulted in a minimally required sample size of n=300. Continuous data are presented as means with standard deviations or medians with ranges as appropriate. Categorical data are presented as proportions. For cases with a missing value (i.e., CRP, LDH, lymphocytes, neutrophils, and days of symptoms, 21/356, 6%), data were imputed using multiple imputation (27). Prior to imputation, data was analyzed to ensure the assumption missingness at random was reasonable. Univariable analyses were performed by means of a Chi-square Test, Mann-Whitney U Test or Student T-Test as appropriate. Variables for the multivariable logistic regression were chosen based on univariable testing with $P<.1$ and clinical reasoning to prevent multiple testing.

Multivariable logistic regression (backward elimination with fractional polynomial transformation) was performed with $P<.2$ for the inclusion of covariables and $P<.05$ for fractional polynomial

transformation of variables. Fractional polynomial transformation was chosen because of the exploratory character of the study (28). Discrimination performance was determined by the area under the receiver operating characteristic curve (AUC). Calibration was assessed graphically (29). Internal validation was done by non-parametric bootstrapping (k=2000) for the AUC (30).

A decision curve for our model was included to aid clinical decision making based on a risk threshold preference (31). A decision curve is a plot of standardized net benefit against the risk probability threshold. The net benefit describes the difference between the benefit of true positive calls (correct prediction of critical illness) and the harm of false positives, the latter being adjusted to how much a false positive outcome weighs compared to a true positive call (e.g., if the harms of a false positive outcome is considered to weigh twice as much the benefits of a true positive call, the odds ratio is 2:1, i.e., 2/1 = 2). The standardized net benefit is net benefit divided by the prevalence of the outcome (the proportion of patients with critical illness in this study) (32).

Analysis was conducted in IBM SPSS Statistics for Windows (Version 26.0; IBM, Armonk, NY) and R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Australia). Results of the logistic regression are presented as adjusted odds ratios (ORs) with 95% confidence intervals (CIs). All tests were two-sided, a *P*-value < .05 was deemed statistically significant.

Results

Patient's characteristics

A total of 356 patients were included in the study (Figure 1), with 121/356 (34%) patients from hospital A and 235/345 (66%) patients from hospital B. Between the patients from the two participating centers, no significant differences were found for laboratory values, chest radiography score, or duration of symptoms before admission (Table 1). Mean patient age was 69 years (±12) and 237/356 (67%) were male. Median duration of symptoms was 7 days (range, 0-35 days). A full summary of patient characteristics per hospital is given in Tables E1 and E2.

Outcome

Of all included patients 168/356 (47%) developed critical illness (ICU admission and/or death). Of those 24/168 (14%) died at the ICU, 73/168 (44%) died at a COVID-19 cohort ward, and 71/168 (42%) patients were admitted to ICU but were alive at follow-up.

For all patients, the median duration of symptoms at admission was 7 days (range 0-35 days). Median time-to-intubation from admission was 4 days (range, 0-18 days) and 8 days (range, 1-35 days) since the onset of symptoms. Median survival in the deceased population was 4 days (range, 0-30 days) after admission and 16 days (range, 4-39 days) since the onset of symptoms.

Univariable analysis

Patients who developed critical illness were significantly older (mean age, 70 years ±11 vs. 67 years ±13, p=.03), more often male (124/168 [73%] vs. 114/188 [61%], p=.01), and more often had chronic obstructive lung disease (40/168 [24%] vs. 29/188 [14%], p=.06) (Table 1).

Regarding chest radiographic findings, patients who developed critical illness more often had higher chest radiography scores (mean 4.4±1.9 vs. 3.3±1.8, p<.001) and bilateral involvement (144/168

[86%] vs. 137/188 [73%]; p=.006) at admission. Regarding the distribution of abnormalities on chest radiography, predominantly central or diffuse opacifications were correlated with the development of critical illness (p=.001) (Table 2).

Patients who developed critical illness more often had leukopenia (37/168 [22%] vs. 36/188 [19%], p=.004) and lymphopenia (84/168 [51%] vs. 63/188 [34%], p=.002) and a higher NLR (median 7.0 [range, 0.5-66.5] vs. 5.75 [range, 0.09-47.0], p<.001) at admission. Additionally, they exhibited higher CRP (mean 139.1±100.5 vs. 94.6±74.6, p=.001) and LDH (421.2±251.4 vs. 317.1±139.9, p=.001). No significant difference was found for the duration of symptoms (p=.15), but the proportion of patients with a symptom duration >7 days was slightly smaller in the critical group (64/168 [38%] vs. 87/188 [48%]; p=.08). No significant difference was found for body habitus on chest radiography (p=.12), temperature (p=.43), and the other comorbidities (Table 2).

Model development and internal validation

After backward elimination, the following variables remained in the model: male gender(adjusted OR 1.5, 95% CI 0.9-2.7, P=.10), obstructive lung disease (adjusted OR 1.9, 95% CI 1.0-3.5, P=.045), symptom duration > 7 days (adjusted OR 0.5, 95% CI 0.3-0.8, P=.003), $(neutrophils/10)^3$ (adjusted OR 1.8, 95% CI 1.2-2.9, P=.01), $(CRP/100)^{-2}$ (adjusted OR 0.98, 95% CI 0.95-0.997, P=.08), LDH/1000 (adjusted OR 8.4, 95% CI 1.7-49, P=.01), diffuse distribution (vs. peripheral) (adjusted OR 1.9, 95% CI 1.1-3.3, P=.03), central distribution (vs. peripheral) (adjusted OR 3.6, 95% CI 1.1-9.3, P=.01), and chest radiography score (adjusted OR 1.2 per point increasing, 95% CI 1.1-1.5, P=.01). There were no significant interactions between variables. A summary including a description of variables entered in the full model is given in Table 3. The AUC for this model was 0.77 (95% CI 0.72-0.81; P < .001). The model showed good calibration, with a slope of 0.974 (Fig E1 [Appendix E1]). The average AUCs of the models derived from 2000 bootstrap samples on the bootstrap and original samples were 0.78 (95% CI 0.78-0.78) and 0.75 (95% CI 0.75-0.76), respectively (optimism = 0.780 - 0.754 = 0.026).

Risk score and risk calculator

The risk score was developed using the regression coefficients of the logistic regression model (β) resulting in the following formula: probability = 1/(1+exp(-(intercept + gender × β + obstructive lung disease × β + > 7 days of symptoms × β + (neutrophils/10)^3 × β + (CRP/100)^-2 × β + (LDH/1000) × β + distribution × β + CXR score × β))). The risk calculator can be accessed on line (hyperlink: Dutch COVID-19 risk model) (see Appendix E2).

Decision curve analysis

The decision curve (Figure 4) shows that our model (red line) provides a superior net benefit for risk thresholds of 0.5 or higher, equivalent to a cost-benefit ratio of 1:1 or greater in favor of patients who do not develop critical illness. For example, the risk score threshold of 0.70 (cost-benefit-ratio 7:3) is at the 80th percentile of the population, meaning that 20% of the population will get a positive test (i.e. positive for development of critical illness). In our study this is 71/356 (20%) patients. Of these 59/71 (83%) would be true positives and 12/71 (17%) false positives. The calculated net benefit is: ((59/356 - (12/356 x 0.7/ (1-0.7))) = 0.087. The standardized net benefit is: 0.087 / (168/356) = 0.18. This calculation would be equivalent to, out of 100 patients analyzed, 18 of the critical patients would be correctly predicted without any incorrect predictions among severe patients. Note that the risk probability threshold should not be selected based on the highest net benefit, but rather on the clinical scenario (32).

Discussion

In this study we developed a risk model to predict short -term critical illness defined as ICU admission with mechanical ventilation and / or death in hospitalized patients with coronavirus disease 2019 (COVID-19). The model was based on clinical, laboratory, and radiographic findings obtained at admission in 356 patients and showed an AUC of 0.77 (95% CI 0.72-0.81; P <.001).

We show a complementary effect of clinical, laboratory, and specific chest radiography findings for the prediction of short-term critical illness in hospitalized patients with COVID-19. Two recently published retrospective studies explored the prognostic value of chest radiographic findings (26, 33). Toussie et al. included patients under 50 years old with confirmed COVID-19 (n=338) presenting at the emergency ward, thus a younger population with fewer comorbidities than in our study. The authors looked at the predictive value of chest radiography for various outcomes, including hospital admission, and did not include laboratory markers in their models. In the study by Liang et al., a clinical risk score was developed on a large cohort of hospitalized patients with COVID-19 (n=1590) recruited from a national Chinese database. Clinical parameters, laboratory findings, and presence of any abnormality on chest radiography were used to predict critical illness. Their population was also relatively healthy: the mean age was under 50 and only 8% (131/1590) of patients developed critical illness. This also explains why age, presence of severe symptoms (e.g., unconsciousness), or number of comorbidities were very strong predictors in their model and contributed to a high discriminatory predictive ability (AUC = 0.88). Our study population however was different and consisted exclusively of patients with severe disease who required hospital admission. Our population had a large proportion that developed critical illness (168/356; 47%). It is likely that the discriminatory ability of our model would improve when also applied to outpatients, since those patients would show less disease findings than our study population (26, 33). Furthermore, Liang et al. included only one measure of abnormality on chest radiography (yes vs. no), despite it is known that the extent of lung involvement in viral pneumonia visible on chest radiography or CT may predict a worsened outcome (21-23, 25, 33, 34). Moreover, we showed that the distribution of lung abnormalities on chest radiography provides additional prognostic information as well. Patients with diffuse or central lung

abnormalities, had higher odds for critical illness in our risk model than only peripheral abnormalities.

Although we included several chest radiography characteristics and laboratory markers in our analysis, we aimed to keep our risk model simple so that it can be used in resource limited areas. We chose to use chest radiography instead of CT and the scoring of the extent of disease on chest radiography was a simple semi-quantitative visual assessment, since estimation of the precise pulmonary involvement on chest radiography is challenging, especially from portable bedside chest radiographs. Secondly, we did not include many of the commonly used laboratory markers such as ferritin, procalcitonin, D-dimer, fibrinogen (3, 4, 6, 7, 11, 26) in our analysis, since some laboratory tests are not widely available or are expensive.

Besides providing accuracy measures such as AUC, sensitivity, and specificity, we also performed decision curve analysis. This method has been increasingly used to assess diagnostic tests and prediction models. The advantage of decision curve analysis is that it incorporates the preference of the physicians, patients, and/or policy makers into analysis. It weighs the benefits and costs at a certain threshold, i.e., the minimum event risk probability which defines a positive test outcome. When ICU facilities are limited, one may prefer to predict critical illness with high specificity, i.e., a low false positive rate. We suggested a risk probability threshold of 0.70 using our model; this means that the early preparation for clinical deterioration would be justified if there were at least 2.3 true positives (0.70/[1–0.70]) per false positive. In our population, this would have resulted in 71 of 356 patients having a positive test (i.e., be predicted to develop critical illness) of which 59 (83%) would be true positives. This results in a standardized net benefit of 0.18. This concept is easiest understood by first considering the baseline where all patients are assumed to have a negative test, i.e., zero true positives and zero false positives. In comparison: By applying our model to 100 patients using the probability threshold of 0.70 for a positive test, an additional 18 true positives could be detected

while keeping the number of false positives at zero. A more detailed explanation about decision curve analysis is provided in Appendix E3.

There are a number of limitations to this work. Our prediction model has not been externally validated, and is based on a retrospective cohort of only 356 patients from two hospitals. We included known predictive parameters of worse outcome in patients with COVID-19 at the time of inclusion of our study population. However, we were not able to incorporate all known clinical parameters and laboratory values such as, body mass index, history of tobacco use, oxygenation levels at admission or pulmonary function, D-dimer, ferritin, mainly because they were not available for all patients during the peak of the pandemic.

Chest radiographs were independently analyzed by radiologists from their respective centers using a scoring system designed for this study, and heterogeneity among radiologists was unaccounted for. However, no significant differences in chest radiography scoring were seen between centers with similar source populations, suggesting satisfactory inter-rater agreement.

Lastly, this study aimed to identify predictors and develop a risk model for critical illness in a setting where the knowledge regarding risk factors for critical illness in COVID-19 is still limited. The risk model shows good calibration and robustness upon internal validation (optimism = 0.026) and will therefore very likely perform well in a similar population. However, the performance may differ in other source populations.

In summary, we found that basic laboratory findings and a simple assessment of parenchymal involvement on a chest radiography acquired at hospital admission may provide complimentary information for short-term prognosis in hospitalized patients with coronavirus disease. We show that a simple model composed of gender, chronic obstructive lung disease, neutrophil granulocytes, C-reactive protein, lactate dehydrogenase, distribution of lung abnormalities and a chest radiography

score were predictive of the need for mechanical ventilation/death among hospitalized patients with coronavirus disease. A risk calculator has been made available for download: Dutch COVID-19 risk model (Appendix E2).

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis [Internet]. 2020 May;20(5):533-4. Accessed August 9th 2020.

2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, china. Lancet [Internet]. 2020 Feb 15;395(10223):497-506.

3. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, china. JAMA [Internet]. 2020 Feb 7;323(11):1061-9.

4. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, china: A descriptive study. Lancet [Internet]. 2020 Feb 15;395(10223):507-13.

5. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, Liu L, Shan H, Lei CL, Hui DSC, Du B, Li LJ, Zeng G, Yuen KY, Chen RC, Tang CL, Wang T, Chen PY, Xiang J, Li SY, Wang JL, Liang ZJ, Peng YX, Wei L, Liu Y, Hu YH, Peng P, Wang JM, Liu JY, Chen Z, Li G, Zheng ZJ, Qiu SQ, Luo J, Ye CJ, Zhu SY, Zhong NS, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in china. N Engl J Med [Internet]. 2020 Apr 30;382(18):1708-20.

6. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, china: A retrospective cohort study. Lancet [Internet]. 2020 Mar 28;395(10229):1054-62.

7. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, Huang H, Zhang L, Zhou X, Du C, Zhang Y, Song J, Wang S, Chao Y, Yang Z, Xu J, Zhou X, Chen D, Xiong W, Xu L, Zhou F, Jiang J, Bai C, Zheng J, Song Y. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, china. JAMA Intern Med [Internet]. 2020 Mar 13

8. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, Li Q, Jiang C, Zhou Y, Liu S, Ye C, Zhang P, Xing Y, Guo H, Tang W. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. J Infect [Internet]. 2020 Apr 23

9. Li LQ, Huang T, Wang YQ, Wang ZP, Liang Y, Huang TB, Zhang HY, Sun W, Wang Y. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. J Med Virol [Internet]. 2020 March 12

10. Yang AP, Liu JP, Tao WQ, Li HM. The diagnostic and predictive role of NLR, d-NLR and PLR in COVID-19 patients. Int Immunopharmacol [Internet]. 2020 Apr 13;84:106504.

11. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, HLH Across Specialty Collaboration, UK. COVID-19: Consider cytokine storm syndromes and immunosuppression. Lancet [Internet]. 2020 March 28;395(10229):1033-4.

12. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, china. Intensive Care Med [Internet]. 2020 Mar 3

13. Gao L, Jiang D, Wen XS, Cheng XC, Sun M, He B, You LN, Lei P, Tan XW, Qin S, Cai GQ, Zhang DY. Prognostic value of NT-proBNP in patients with severe COVID-19. Respir Res [Internet]. 2020 Apr 15;21(1):83-w.

14. Zhang JJ, Dong X, Cao YY, Yuan YD, Yang YB, Yan YQ, Akdis CA, Gao YD. Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, china. Allergy [Internet]. 2020 February 19

15. Velavan TP, Meyer CG. Mild versus severe COVID-19: Laboratory markers. Int J Infect Dis [Internet]. 2020 Apr 25

16. Simpson S, Kay FU, Abbara S, Bhalla S, Chung JH, Chung M, Henry TS, Kanne JP, Kligerman S, Ko JP, Litt H. Radiological society of north America expert consensus statement on reporting chest CT findings related to COVID-19. Endorsed by the society of thoracic radiology, the American college of radiology, and RSNA. J Thorac Imaging [Internet]. 2020 April 28

17. World Health Organization. Clinical management of COVID-19. [Internet]. 2020 27 May(WHO/2019-nCoV/clinical/2020.5) Available from: https://www.who.int/publications-detail/clinical-management-of-covid-19

18. Kim H, Hong H, Yoon SH. Diagnostic performance of CT and reverse transcriptase-polymerase chain reaction for coronavirus disease 2019: A meta-analysis. Radiology [Internet]. 2020 April 17:201343.

19. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in china: A report of 1014 cases. Radiology [Internet]. 2020 Feb 26:200642.

20. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, Schluger NW, Volpi A, Yim JJ, Martin IBK, Anderson DJ, Kong C, Altes T, Bush A, Desai SR, Goldin J, Goo JM, Humbert M, Inoue Y, Kauczor HU, Luo F, Mazzone PJ, Prokop M, Remy-Jardin M, Richeldi L, Schaefer-Prokop CM, Tomiyama N, Wells AU, Leung AN. The role of chest imaging in patient management during the COVID-19 pandemic: A multinational consensus statement from the Fleischner society. Radiology [Internet]. 2020 April 07:201365.

21. Chau TN, Lee PO, Choi KW, Lee CM, Ma KF, Tsang TY, Tso YK, Chiu MC, Tong WL, Yu WC, Lai ST. Value of initial chest radiographs for predicting clinical outcomes in patients with severe acute respiratory syndrome. Am J Med [Internet]. 2004 August 15;117(4):249-54.

22. Antonio GE, Wong KT, Tsui EL, Chan DP, Hui DS, Ng AW, Shing KK, Yuen EH, Chan JC, Ahuja AT. Chest radiograph scores as potential prognostic indicators in severe acute respiratory syndrome (SARS). AJR Am J Roentgenol [Internet]. 2005 March 01;184(3):734-41.

23. Hui DS, Chan MC, Wu AK, Ng PC. Severe acute respiratory syndrome (SARS): Epidemiology and clinical features. Postgrad Med J [Internet]. 2004 July 01;80(945):373-81.

24. Cha MJ, Chung MJ, Kim K, Lee KS, Kim TJ, Kim TS. Clinical implication of radiographic scores in acute Middle East respiratory syndrome coronavirus pneumonia: Report from a single tertiary-referral center of South Korea. Eur J Radiol [Internet]. 2018 October 01; 107:196-202.

25. Das KM, Lee EY, Al Jawder SE, Enani MA, Singh R, Skakni L, Al-Nakshabandi N, AlDossari K, Larsson SG. Acute middle east respiratory syndrome coronavirus: Temporal lung changes observed on the chest radiographs of 55 patients. AJR Am J Roentgenol [Internet]. 2015 September 01;205(3):267.

26. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, Li Y, Guan W, Sang L, Lu J, Xu Y, Chen G, Guo H, Guo J, Chen Z, Zhao Y, Li S, Zhang N, Zhong N, He J, China Medical Treatment Expert Group for, COVID-19. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. JAMA Intern Med [Internet]. 2020 May 12

27. van Buuren S, Groothuis-Oudshoorn CGM. Mice: Multivariate imputation by chained equations in R. Journal of statistical software [Internet]. 2011;45(3) Available from: https://www.narcis.nl/publication/RecordID/oai:ris.utwente.nl:publications%2Ff332b769-85cd-4d13-818f-5799cb957366

28. Zhang Z, Rousson V, Lee WC, Ferdynus C, Chen M, Qian X, Guo Y, written on behalf of AME Big-Data Clinical Trial Collaborative Group. Decision curve analysis: A technical note. Ann Transl Med [Internet]. 2018 August 01;6(15):308.

29. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. Ann Intern Med [Internet]. 2019 January 01; 170(1):W1-W33.

30. Steyerberg, EW. Clinical prediction models: a practical approach to development, validation, and updating. [Internet]. New York, NY: Springer; 2009. 497 p

31. Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. Med Decis Making [Internet]. 2006 December 01;26(6):565-74.

32. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the clinical impact of risk prediction models with decision curves: Guidance for correct interpretation and appropriate use. J Clin Oncol [Internet]. 2016 July 20;34(21):2534-40.

33. Toussie D, Voutsinas N, Finkelstein M, Cedillo MA, Manna S, Maron SZ, Jacobi A, Chung M, Bernheim A, Eber C, Concepcion J, Fayad Z, Gupta YS. Clinical and chest radiography features determine patient outcomes in young and middle age adults with COVID-19. Radiology [Internet]. 2020 May 14:201754.

34. Tabatabaei SMH, Talari H, Moghaddas F, Rajebi H. Computed tomographic features and short-term prognosis of coronavirus disease 2019 (COVID-19) pneumonia: A single-center study from kashan, iran. Radiology: Cardiothoracic Imaging [Internet]. 2020 Apr 20, 02(02)

Tables

Table 1. Demographics and Clinical Findings per Disease Category.

Clinical characteristics at presentation to the hospital	Non-critical COVID-19 group (n=188)	Critical COVID-19 group* (n=168)	P-value
Age (years, mean ± SD)	67±13	70±11	.03
Male: Female (n, %)	114:74	124:44	.01
Large body habitus on CXR (n)	95 (51%)	89 (53%)	.12
Duration of symptoms prior to presentation (median days , range)	7 (0-30)	7 (0-35)	.15
> 7 days symptoms	87 (46%)	64 (38%)	.08
Absence of comorbidities (n,%)	62 (33%)	40 (24%)	.06
No. of comorbidities (median, range)	1 (0-5)	1 (0-4)	.03
Hypertension	65 (35%)	68 (41%)	.25
History of cancer	30 (16%)	21 (13%)	.35
Diabetes	31 (17%)	35 (21%)	.29
Cardiac disease	47 (25%)	52 (31%)	.21
Chronic obstructive lung disease	29 (15%)	40 (24%)	.046
Temperature degrees Celsius (mean, SD)	38.2±1.1	38.3±1.0	.43
Leucopenia (< $4.5x10^9$/L) **Normal (4.5-$11x10^9$/L)** **Leukocytosis (>$11x10^9$/L)**	36 (19%) 136 (72%) 16 (9%)	37 (22%) 98 (53%) 33 (20%)	.004
Lymphocytopenia (<$0.8x10^9$/L) (n,%)	63 (34%)	84 (51%)	.002
Neutrophilia (>$8.0x10^9$/L) (n,%)	18 (10%)	45 (27%)	<.001
Neutrophil-lymphocyte ratio (median, range)	5.8 (0.09-47)	7.0 (0.5-67)	<.001

CRP (mg/L, mean ± SD)	95±75	139±101	.001
LDH (IU/L, mean ± SD)	317±1340	421±251	.001
* Critical illness is defined by need for invasive ventilation and/or death. Abbreviations: CRP = C-reactive protein; LDH = lactate dehydrogenase.			

Demographics, clinical characteristics on admission per outcome.

Table 2 : Chest Radiographic Findings per Disease Category

Imaging characteristics at presentation to the hospital	Severe COVID-19 group (n=188)	Critical COVID-19 group * (n=168)	P-value
Anteroposterior image (n,%)	182 (97%)	164 (98%)	.64
Pre-existent abnormalities on chest radiography (n,%)	12 (6%)	23 (14%)	.02
Bilateral pneumonia (n,%)	137 (73%)	144 (86%)	.01
Distribution **Peripheral** **Diffuse** **Central** **Normal**	 119 (63%) 43 (23%) 10 (5%) 16 (9%)	 74(44%) 72(43%) 18(11%) 4 (2%)	.001
Chest radiography score (0-8) **median (range)** **mean (SD)**	 4 (0-8) 3.3 (1.8)	 4 (0-8) 4.4 (1.9)	<.001
Upper zones Involved (n,%)	134 (71%)	146 (87%)	<.001
Lower zones Involved (n,%)	170 (90%)	161 (96%)	<.05
Lung zones involved (median, range)	3.5 (0-4)	4 (0-4)	.003
Involvement per lung zone (n, %) **Left upper zone** normal	 74 (40%)	 43 (26%)	 .01

mild/moderate	104 (55%)	105(63%)	
severe	10 (5%)	20 (12%)	
Right upper zone			<.001
normal	72 (38%)	38 (23%)	
mild/moderate	101 (54%)	96 (57%)	
severe	15 (8%)	34 (20%)	
Left lower zone			<.001
normal	30 (16%)	14(8%)	
mild/moderate	120 (64%)	90 (54%)	
severe	38 (20%)	64 (38%)	
Right lower zone			<.001
normal	38 (20%)	13 (8%)	
mild/moderate	132 (70%)	104 (62%)	
severe	18 (10%)	51 (30%)	

* Critical illness is defined by need for invasive ventilation and/or death.

Radiographic findings on admission per outcome.

Table 3. Dutch COVID-19 Prognostic Model for Critical Illness in Known RT-PCR Positive and Hospitalized Patients.

Variables at Presentation*	Adjusted Odds Ratio	95% CI	P-value
Male gender (yes vs. no)	1.5	0.9-2.6	.10
Obstructive lung disease (yes vs. no)	1.9	1.0-3.5	.045
> 7 days symptoms (yes vs. no)	0.5	0.3-0.8	.003
Neutrophils (10^9/L)†	1.8	1.2-2.9	.01
C-reactive protein (mg/L)‡	0.98	0.95-0.997	.08
Lactate dehydrogenase (IU/L)§	8.4	1.7-49	.01
Distribution of Lung disease on CXR			
Peripheral	1		Reference group
Diffuse	1.9	1.1-3.3	.03
Central	3.6	1.4-9.3	.01
Normal	1.8	0.4-6.9	.41
Chest radiography score (per point) ǁ	1.2	1.1-1.5	.01

* Variables included in the full model: age, gender, obstructive lung disease (yes vs. no), number of comorbidities present, leucocytes (leukopenia, normal, leukocytosis), lymphocytes, neutrophil-lymphocyte-ratio, C-reactive protein, Lactate dehydrogenase, pre-existent abnormalities on chest radiography, bilateral involvement, distribution (peripheral, diffuse, central or normal), number of lung zones involved, chest radiography score. Variables not included in this table did not contribute significantly to the parsimonious final model (p>.2).
† Transformed by dividing by 10 and then exponentiating by the power 3.
‡ Transformed by dividing by 100 and then exponentiating by the power -2.
§ Transformed by dividing by 1000.
ǁ Treated as a continuous variable.

Multivariable logistic regression model for predicting short-term critical illness (ICU admission for invasive ventilation and/or death) in patients with coronavirus disease 2019 based on parameters obtained at admission. A calculator for this model is available for download (see Appendix E2).

Table 4: Critical Illness Prediction Accuracy at Different Model Risk Thresholds in 356 Known RT-PCR Positive and Hospitalized Patients.

Model risk percentile	Model risk probability	Cost:benefit ratio	Critical illness patients (%)	Non-critical illness patients (%)	Sensitivity, %	Specificity, %	PPV, %	NPV, %	sNB
≥0	.00	1:∞	168 (100)	188 (100)	100	0	100	0	1.00
≥50	.46	23:27	115 (68)	63 (34)	68	67	64	70	0.36
≥60	.51	51:49	100 (60)	42 (22)	60	78	70	68	0.34
≥70	.60	3:2	84 (50)	23 (12)	50	88	79	66	0.28
≥80	.70	7:3	59 (35)	12 (6)	35	94	83	62	0.18
≥90	.81	81:19	34 (20)	2 (1)	20	99	94	58	0.09

This table reports the model accuracy at various risk thresholds for predicting short-term critical illness (ICU admission for invasive ventilation and/or death) in hospitalized patients with coronavirus disease 2019. Fewer patients receive a positive test outcome at higher risk thresholds, though the ratio of true positives (i.e., critical ill patients) compared to false positives (i.e., non-critical ill patients) increases. A positive standardized net benefit (sNB) value indicates that applying the model at that threshold is superior to a scenario where all patients are assumed to not develop critical illness (Figure 4); this is not necessarily superior to a scenario where all patients are assumed to develop critical illness. Please refer to Appendix E3 for a more detailed explanation for interpreting sNB. Abbreviations: PPV, positive predictive value; NPV, negative predictive value; sNB, standardized net benefit.

Figures

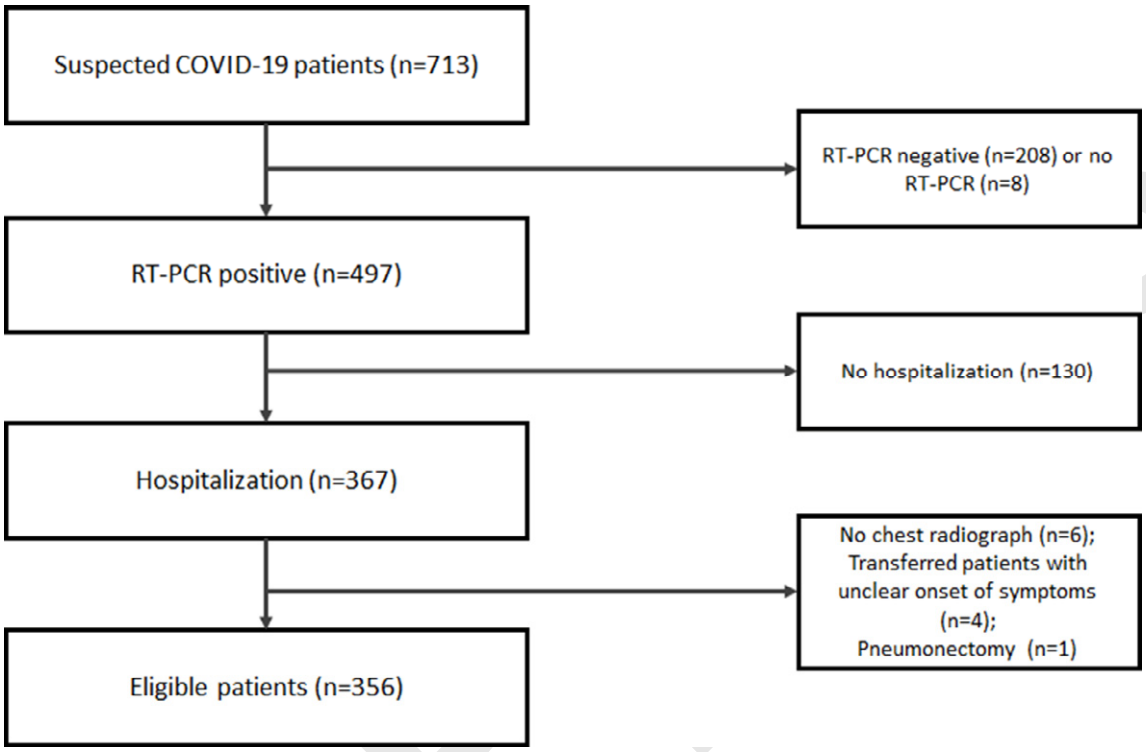

Figure 1: Flowchart of patient inclusion. Patients without or with a negative RT-PCR, patients who were not hospitalized, and a small number of patients who did not receive a chest radiograph or were transferred from another hospital without a clear onset of symptoms were excluded. A total of 356 patients were eligible for this study.

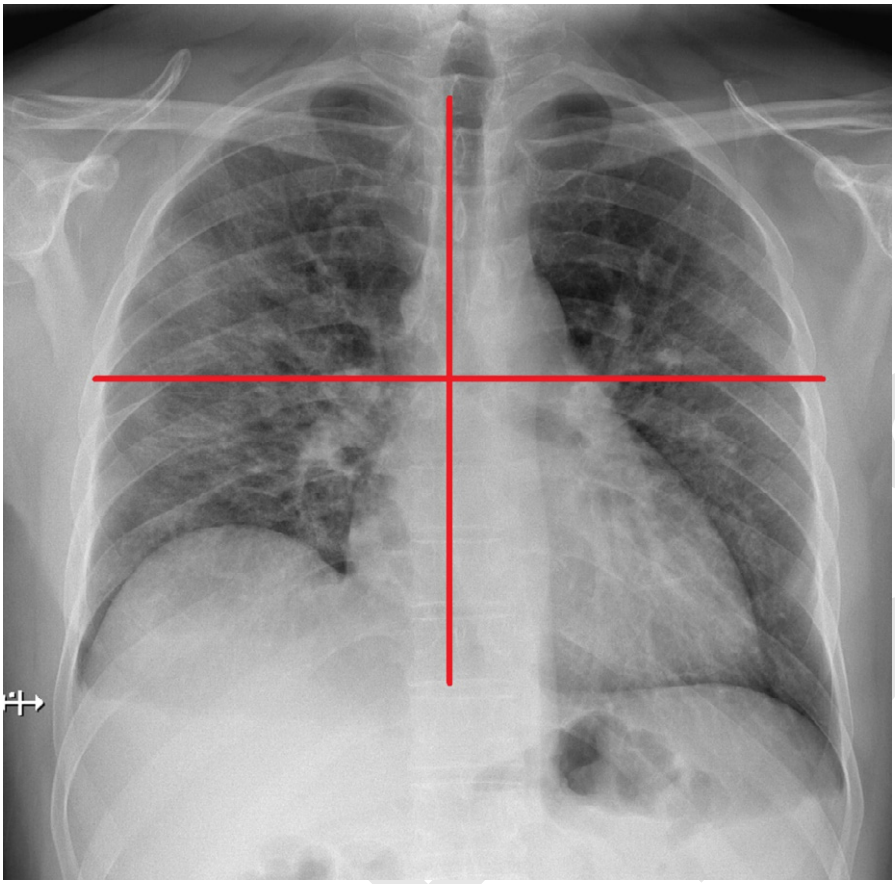

Figure 2: Chest radiography scoring: in a 50-year-old patient with RT-PCR proven coronavirus disease 2019 who was hospitalized but not admitted to the ICU. Chest radiography scoring was as follows: right upper lung zone mild/moderate involvement (1 point), the right lower lung zone mild/moderate involvement (1 points), as well as for the left upper and lower lung zones mild/moderate involvement (both 1 point), resulting in a cumulative score of 4.

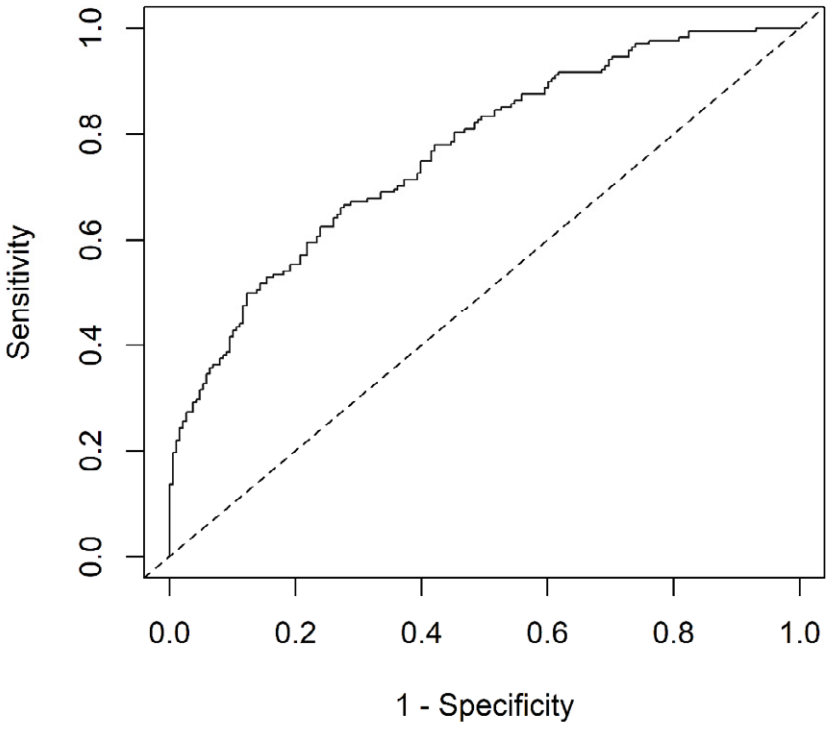

Figure 3: Receiver operating characteristics curve for the multivariable logistic regression model predicting critical illness in patients with COVID-19. The model including gender, clinical, laboratory, and imaging (chest radiography) parameters (n=356) reached an AUC of 0.77 for prediction of critical illness.

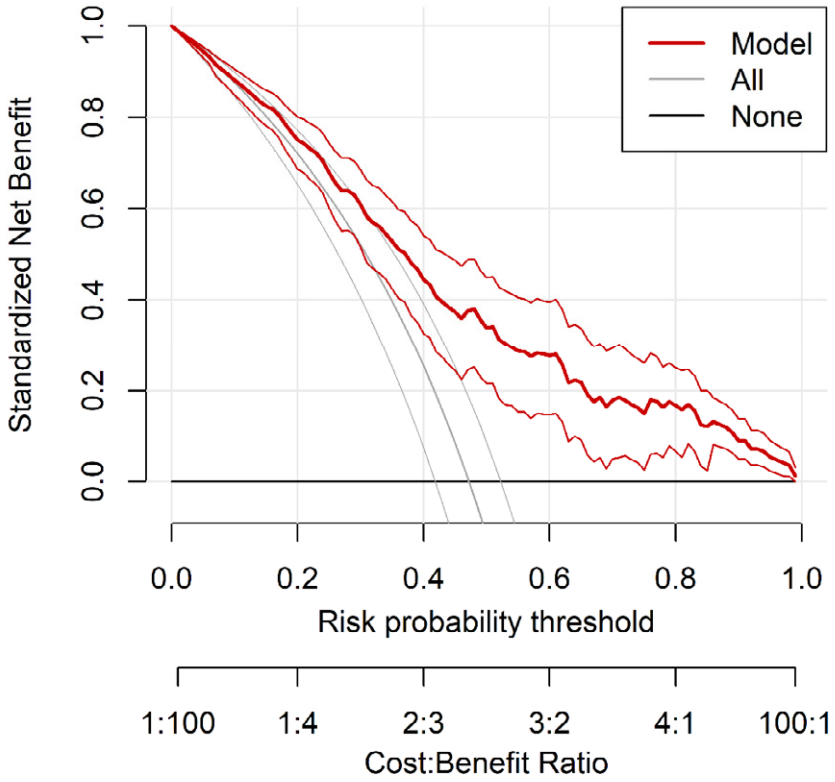

Figure 4: Decision curve for the Dutch COVID-19 risk model. The gray and black lines (horizontal) represent the scenarios where all or none of the hospitalized patients would be prospectively determined by the risk model, respectively. The red line demonstrates the net benefit of the risk model dependent at the chosen risk threshold. The accompanying thinner lines represent the 95% confidence intervals.

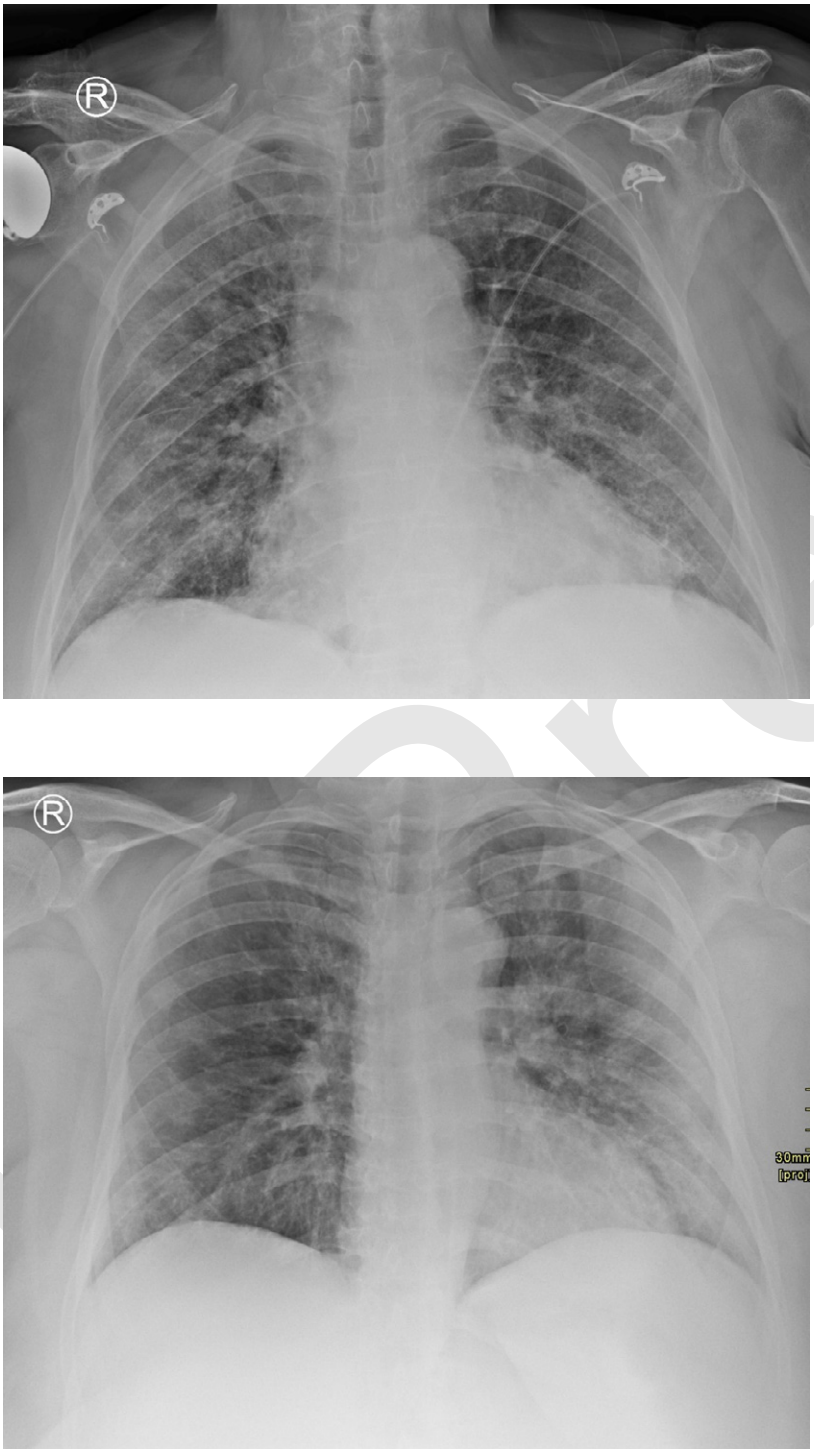

Figure 5: Examples of patients with coronavirus disease 2019 who did (i.e. required intubation and or died) and did not develop critical illness (i.e. did not need intubation and were discharged). *A,* Radiograph in an 80-year-old woman with 7 days of symptoms prior to emergency ward presentation. She had hypertension and chronic obstructive lung disease as comorbidity. Laboratory findings were as follows: neutrophil granulocytes $6.4x10^9$/L, C-reactive protein 282 mg/L, lactate

dehydrogenase 335 IU/L. On chest radiography she exhibited diffuse bilateral opacities and a chest radiography score of 4. The calculated risk score was 0.69. She developed critical illness and died 4 days after hospital admission. *B,* Radiograph in a 46-year-old woman with no comorbidities had 9 days of symptoms prior to emergency ward presentation. Laboratory findings were as follows: neutrophil granulocytes $3.8x10^9$/L, C-reactive protein 29 mg/L, lactate dehydrogenase 429 IU/L. Chest radiography showed bilateral peripheral opacities and a chest radiography score of 5. The calculated risk score was 0.24. She was not admitted to the ICU and successfully discharged after 2 days of hospitalization.

Appendix E1

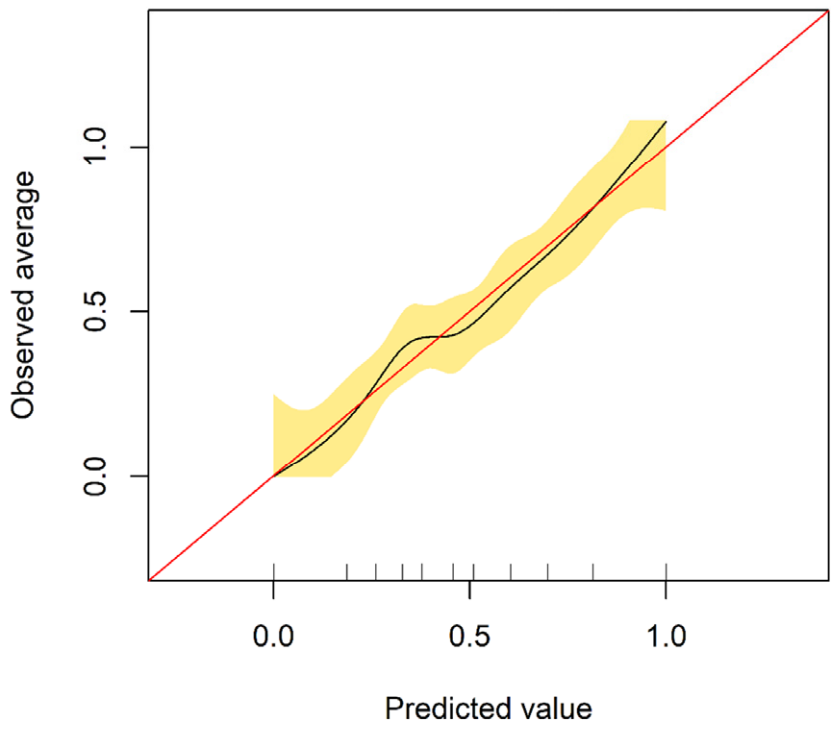

Figure E1: Calibration plot of the multivariable logistic regression model predicting short-term poor prognosis in patients with coronavirus disease 2019. The calibration plot refers to the agreement between the predicted outcome and observed outcome. A perfect calibration results in a calibration plot in which the observed values exactly match the predicted values. This would result in a slope of exactly 1. The slope of our model is 0.974, which is considered a good calibration slope, and implies that the model is reliable.

Appendix E2: Risk Calculator.

Variables	Input values below	Equivalent value	Description
Male sex	No	0	Male sex
Obstructive lung disease	No	0	Medical diagnosis of asthma or chronic obstructive pulmonary disease
> 7 days complaints	No	0	Self reported symptoms for more than 7 days on hospital admission
Neutrophils (10^9/L)	3.8	3.8	Blood neutrophil level
C-reactive protein (mg/L)	29	29	Blood C-reactive protein level
Lactate dehydrogenase (IU/L)	429	429	Blood Lactate dehydrogenase level
Distribution	Peripheral	0	Distribution pattern on the chest x-ray, either peripheral, difuse, central, or normal
Chest x-ray score	0	0	Value 0 to 8. Sum of score per quadrant: no involvement = 0, mild/moderate involvement =1, severe involvement (>50%) = 2 points

RISK OF ICU/DEATH (%)	18.1

Appendix E3: Short Guide to Decision Curve Analysis.

The output of risk prediction models is the probability that an event will happen given a set of predictor variables. Using the Dutch coronavirus disease 2019 (COVID-19) risk model as an example, the probability that hospitalized patients with COVID-19 will develop critical illness is based on the predictors given in Table 3. The standard reported measure of risk model performance is the area under the receiver operating characteristic curve (AUC). In this case, AUC should be interpreted as the probability that a patient which encounters critical illness has a greater risk probability than a patient which does not. Although simple to understand and interpret, AUC is not informative as to its clinical value (1). This is because AUC does not consider whether it is preferable to have false-negative or false-positive results. For example, in a scenario where there are limited intensive care unit (ICU) spaces, a patient who does not develop critical illness but who is immediately assigned to an ICU bed at admission (false positive) may be considered to be less desirable than a patient who does develop critical illness but for who an ICU bed was not immediately reserved. Note that this is a measure of subjective preference which varies among physicians and patients.

Decision curve analysis was introduced as a discriminatory metric for models which considers such clinical consequences (2–4). A key aspect of decision curves is the concept of net benefit. This is a calculation which involves determining the difference between the benefits (i.e., correctly predicting that a hospitalized patient with COVID-19 will develop critical illness [true positive rate]) and costs (i.e., incorrectly predicting that a patient will develop critical illness [false positive rate]) of the intervention. As indicated, the relative weighting of the costs and benefits this is a subjective measure of preference which is equivalent to the risk probability threshold. A higher probability threshold will reduce both the number of patients at "high risk" with and without the event, i.e., the true and false positive rates, respectively (Figure E2). For a model with some predictive ability, this would also result in proportionally more patients with high risk (Figure E3). We suggested a probability threshold of 0.70 for the Dutch COVID-19 risk model, meaning that only patients with a critical illness risk greater of 70% were predicted to develop critical illness. In this case, a false positive result weighs heavier than a true positive result with an odds of 7:3 ($\frac{Probability\ threshold}{1-Probability\ threshold}$). Finally, by adjusting the true positive rate and false positive rate with the prevalence of the event of interest (critical illness) is then possible to calculate the net benefit using the following equation:

$$Net\ benefit = True\ positive\ rate \times Event\ prevalence - \frac{Probability\ threshold}{1 - Proability\ threshold} \times False\ positive\ rate \times (1 - Event\ prevalence)$$

Note that net benefit may be difficult to interpret because the maximum possible value is the event prevalence (i.e., when true positive rate = 1 and false positive rate = 0). It is therefore commonplace to calculate the standardized net benefit by dividing net benefit by the event prevalence. This may simplify the interpretation because standardized net benefit always has a maximum possible value of 1. At a risk probability threshold of 0.70, a standardized net benefit of 0.18 was calculated (Table 4). This can be interpreted as follows: compared to no intervention (all hospitalized patients with COVID-19 are assumed to not develop critical illness), 18% more of the patients which develop critical illness would be correctly predicted (true positives) without additional incorrect predictions among the patients which do not (false positives).

Figure 4 shows the decision curve of the Dutch COVID-19 risk model, where the standardized net benefit is given on the y-axis and the probability threshold (i.e., cost:benefit odds ratio) on the x-axis. Besides the performance of the proposed model (red line), it is also typical to include the decision curves for a policy where all hospitalized patients with COVID-19 are assumed to develop critical illness (grey line) or none at all (horizontal black line). In short, the proposed model can be of added value in the clinic at probability thresholds where its standardized net benefit is greater than that of the two other lines.

For a more detailed guide to the correct use and interpretation of decision curve analysis, please refer to Kerr et al. (3) or Vickers et al. (4).

References

1. Vickers AJ, Cronin AM. Traditional Statistical Methods for Evaluating Prediction Models Are Uninformative as to Clinical Value: Towards a Decision Analytic Framework. *Semin Oncol*. 2010;37(1):31–8. doi:10.1053/j.seminoncol.2009.12.004.

2. Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. *Med Decis Mak*. 2006;26(6):565–74. doi:10.1177/0272989X06295361.

3. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *J Clin Oncol*. 2016;34(21):2534–40. doi:10.1200/JCO.2015.65.5654.

4. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic Progn Res*. 2019;3(1):18. doi:10.1186/s41512-019-0064-7.

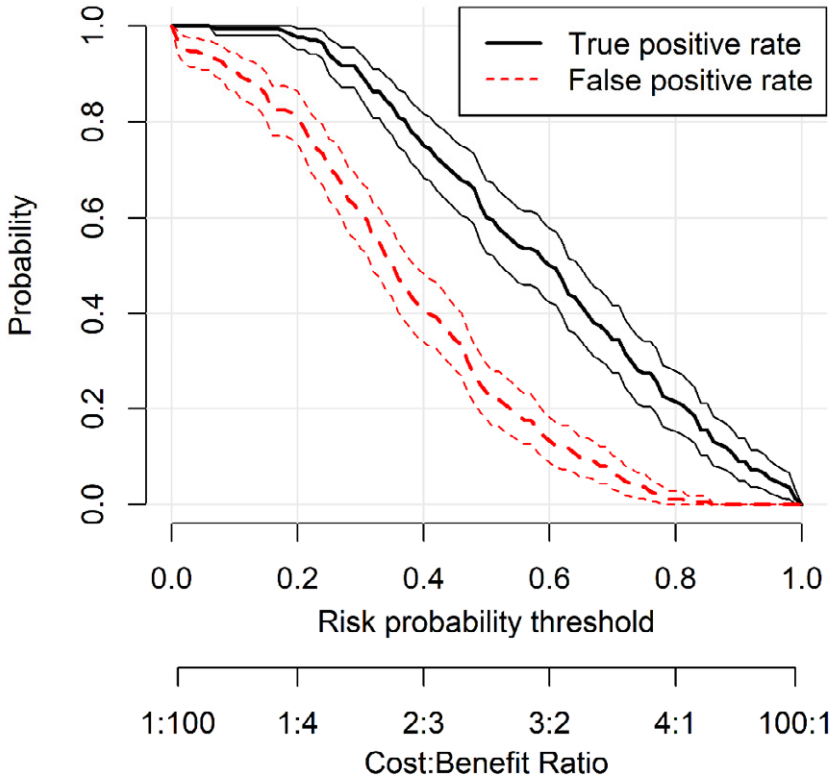

Figure E2: This figure shows the true positive rate (black line) and false positive rate (red line) and their accompanying 95% confidence intervals of the Dutch coronavirus disease 2019 risk model per risk probability threshold. Essentially these are the component of receiver operating characteristics analysis. It can be observed that at higher risk probability thresholds of greater than 0.8 almost no false positives are encountered.

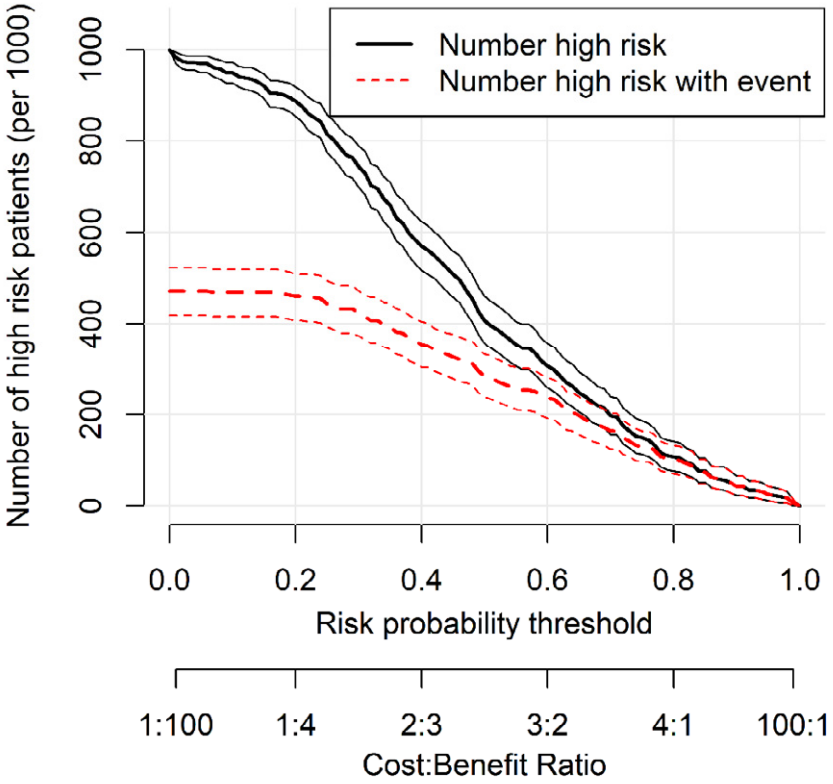

Figure E3: This figure shows the classification of a 1000 hypothetical subject. It gives the number of high risk patients and high risk patients with event (i.e. critical illness) per probability threshold of the risk model, accompanied by the 95% confidence intervals.

Table E1. Study Demographics and Clinical Data by Hospital.

Clinical characteristics at presentation to the hospital	Hospital A	Hospital B	All	P-value
Patients (n)	121	235	356	-
Age in years (mean, SD)	66±14	70±11	69 ± 12	.01
Male: Female	77:44	160:75	237:119	.23
Large body habitus on chest radiography (n, %)	63 (52%)	121 (52%)	184 (52%)	.50
Absence of comorbidities (n,%)	43 (36%)	59 (25%)	102 (29%)	.04
No. of comorbidities (median, range)	1 (0-4)	1 (0-5)	1 (0-5)	.06
Hypertension (n,%)	35 (29%)	98 (42%)	133 (37%)	.02
Cancer (n,%)	18 (15%)	33 (14%)	51 (14%)	.83
Diabetes (n,%)	18 (15%)	48 (20%)	66 (19%)	.20
Cardiac disease (n,%)	34 (28%)	65 (28%)	99 (28%)	.93
Chronic obstructive lung disease (n,%)	20 (17%)	49 (21%)	69 (19%)	.33
Duration of symptoms prior to presentation (median days, range)	8 (0-35)	7 (0-30)	7 (0-35)	.06
> 7 days symptoms (n,%)	64 (53%)	93 (40%)	151 (43%)	.02
Body temperature (°C) (mean, SD)	38.0±1.02	38.4±1.02	38.3 ± 1.04	.001
Leukopenia (< 4.5 x10^9/L) **Normal (4.5-11x10^9/L)** **Leukocytosis (>11x10^9/L)**	23 (19%) 80 (66%) 18 (15%)	50 (21%) 154 (66%) 31 (13%)	73 (20%) 234 (66%) 49 (14%)	.83
Lymphocytopenia (<0.8x10^9/L) (n,%)	51 (43%)	96 (42%)	147 (42%)	.97
Neutrophilia (>8.0x10^9/L) (n,%)	26 (22%)	37 (16%)	63 (18%)	.24

Neutrophil-lymphocyte ratio (median, range)	6.5 (0.6-67)	6.3 (0.1-47)	6.3 (0.1-67)	.63
CRP (mg/L, mean ±SD)	129±98	109±86	116 ±90	.04
LDH (IU/L, mean ± SD)	359±235	368±190	365±207	.70
Critical illness (ICU admission and/or death) (n,%)	61 (50%)	107 (46%)	168	.38
Abbreviations: ; CRP = C-reactive protein; LDH = lactate dehydrogenase; ICU = intensive care unit.				

Table E2: Chest Radiographic Findings on Admission per Center.

Imaging characteristics at presentation to the hospital	Hospital A	Hospital B	All	P-value
Anteroposterior image (n,%)	114 (94%)	232 (99%)	346 (97%)	.02
Pre-existent abnormalities on chest radiography (n,%)	16 (13%)	19 (8%)	35 (10%)	.12
Bilateral pneumonia (n, %)	93 (77%)	188 (80%)	281 (79%)	.55
Distribution				.51
Peripheral (n, %)	60 (50%)	133 (57%)	193 (54%)	
Diffuse (n, %)	41 (34%)	74 (32%)	115 (32%)	
Central (n, %)	11 (9%)	17 (7%)	28 (8%)	
No abnormalities (n, %)	9 (7%)	11 (5%)	20 (6%)	
Chest radiography score (0-8)				
median (range)	4 (0-8)	4 (0-8)	4 (0-8)	.06
mean (SD)	3.5 (1.9)	3.9 (1.9)	3.8 (1.9)	
Involvement per lung zone (n, %)				
Left upper zone				
normal	42 (35%)	75 (32%)	117 (33%)	.79
mild/moderate	68 (56%)	141 (60%)	209 (59%)	
severe	11 (9%)	19 (8%)	30 (8%)	
Right upper zone				.01
normal	45 (37%)	65 (28%)	110 (31%)	
mild/moderate	68 (56%)	129 (55%)	197 (55%)	
severe	8 (7%)	41 (17%)	49 (14%)	
Left lower zone				.02
normal	20 (17%)	24 (10%)	44 (12%)	
mild/moderate	77 (64%)	133 (57%)	210 (59%)	
severe	24 (20%)	78 (33%)	102 (29%)	
Right lower zone				.09
normal	21 (17%)	30 (13%)	51 (14%)	
mild/moderate	71 (59%)	165 (70%)	236 (66%)	
severe	29 (24%)	40 (17%)	69 (19%)	
Upper lung zone involvement (n,%)	92 (76%)	188 (80%)	280 (79%)	.39
Lower lung zone involvement (n,%)	110 (91%)	221 (94%)	331 (93%)	.27