External validation of prognostic scores for Covid-19: a multicenter cohort study of patients hospitalized in Greater Paris University Hospitals

Yannis Lombardi, Loris Azoyan, Piotr Szychowiak, Ali Bellamine, Guillaume Lemaitre, Mélodie Bernaux, Christel Daniel, Judith Leblanc, Quentin Riller, Olivier Steichen

TITLE

External validation of prognostic scores for Covid-19: a multicentre cohort study of patients hospitalized in Greater Paris University Hospitals.

AUTHORS

Yannis Lombardi[1], Loris Azoyan[1,*], Piotr Szychowiak[2,*], Ali Bellamine, MD[3], Guillaume Lemaitre, PhD[4], Mélodie Bernaux, PhD[5], Christel Daniel, MD PhD[6], Judith Leblanc, RN PhD[7], Quentin Riller, MD[1], Olivier Steichen, MD PhD[8,†], and the AP-HP/Universities/INSERM COVID-19 research collaboration and AP-HP COVID CDR Initiative[‡].

[1]Sorbonne Université, Faculty of Medicine, AP-HP, Paris, France

[2]Médecine Intensive-Réanimation, Centre Hospitalier Régional Universitaire de Tours, Tours, France; Université de Tours, Tours, France

[3]TAL Group, WIND Department, AP-HP, Paris, France

[4]Paris Saclay University, INRIA, CEA, Palaiseau, France

[5]Strategy and transformation department, AP-HP, Paris, France

[6]WIND Department, AP-HP, Paris, France

[7]Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique, Paris, France; Clinical Research Platform, Saint Antoine Hospital, AP-HP, Paris, France

[8]Internal Medicine Department, Tenon Hospital, AP-HP, Sorbonne Université, Paris, France

*Equal contributions

[†]Corresponding author (Address: Pr Olivier Steichen, Service de Médecine Interne, Hôpital Tenon, 4 rue de la Chine, 75020 Paris, France; Phone: + 33 1 56 01 78 31; E-mail: olivier.steichen@aphp.fr

[‡]Collaborators are listed in **Appendix 1**

ACKNOWLEDGMENTS

Data used in preparation of this article were obtained from the AP-HP Covid CDW Initiative database. A complete listing of the members can be found at https://eds.aphp.fr/covid-19. We thank the four anonymous reviewers for their insightful comments and constructive suggestions.

ABSTRACT

Purpose

The Coronavirus disease 2019 (Covid-19) has led to an unparalleled influx of patients. Prognostic scores could help optimizing healthcare delivery, but most of them have not been comprehensively validated. We aim to externally validate existing prognostic scores for Covid-19.

Methods

We used “Covid-19 EvidenceAlerts” (McMaster University) to retrieve high-quality prognostic scores predicting death or intensive care unit (ICU) transfer from routinely collected data. We studied their accuracy in a retrospective multicentre cohort of adult patients hospitalized for Covid-19 from January 2020 to April 2021 in the Greater Paris University Hospitals. Areas under the receiver operating characteristic curves (AUC) were computed for the prediction of the original outcome, 30-day in-hospital mortality and the composite of 30-day in-hospital mortality or ICU transfer.

Results

We included 14,343 consecutive patients, 2,583 (18%) died and 5,067 (35%) died or were transferred to the ICU. We examined 274 studies and found 32 scores meeting the inclusion criteria: 19 had a significantly lower AUC in our cohort than in previously published validation studies for the original outcome; 25 performed better to predict in-hospital mortality than the composite of in-hospital mortality or ICU transfer; 7 had an AUC >0.75 to predict in-hospital mortality; 2 had an AUC >0.70 to predict the composite outcome.

Conclusion

Seven prognostic scores were fairly accurate to predict death in hospitalized Covid-19 patients. The 4C Mortality Score and the ABCS stand out because they performed as well in our cohort and their initial validation cohort, during the first epidemic wave and subsequent waves, and in younger and older patients.

KEYWORDS

Covid-19; SARS-CoV2; Prognosis; Intensive Care Units; Mortality; Cohort Studies

INTRODUCTION

Since the end of 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) has spread worldwide [1]. At the end of May 2021, there were over 167 million confirmed cases and over 3.4 million deaths from the coronavirus disease 2019 (Covid-19) around the world [2]. Hospital facilities have thus faced an unparalleled influx of patients. The evolution of hospitalized patients varies widely, from those necessitating no or low level of oxygen to those evolving to acute respiratory or hemodynamic failure requiring admission to intensive care units (ICU) [3, 4]. Accurate outcome prediction with scores based on patient characteristics (age, sex, comorbidities, clinical state, laboratory and imaging results...) help optimizing healthcare delivery in a limited medical resources context [5]. They can also be used to select patients with a homogeneous risk for a given outcome for inclusion in clinical studies.

Various scores have been developed since the beginning of the outbreak and older ones, routinely used in community acquired pneumonia and other conditions, have also been tested in the setting of Covid-19. A systematic review updated in July 2020 found 39 published prognostic scores estimating mortality risk in Covid-19 patients and 28 aimed to predict progression to severe or critical disease. All scores were rated at high or unclear risk of bias. Only a few had undergone external validation, with shortcomings including unrepresentative patient sets, small sizes of the derivation samples and insufficient numbers of outcome events [6]. Moreover, the worldwide applicability of these prediction scores remains an open question: healthcare systems and patient profiles may differ between countries [7] and may impact these scores' performances.

The aim of this study was to evaluate the accuracy of published scores to predict in-hospital mortality or ICU admission in SARS-CoV2-infected patients, using a large multicentre cohort from the Greater Paris University Hospitals (GPUH).

METHODS

Study reporting

Our manuscript complies with the relevant reporting guidelines, namely the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement [8] and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [9]. Completed checklists are available in **Appendix 2**.

Study design and setting

We conducted a retrospective cohort study using the GPUH's Clinical Data Warehouse (CDW), an automatically filled database containing data collected during routine clinical care in the GPUH. GPUH is a public institution and count 39 hospitals (22,474 beds) spread across Paris and its region, accounting for 1.5 million hospitalizations each year (10% of all hospitalizations in France). The data of patients hospitalized for Covid-19 in GPUH was used to evaluate the accuracy of published prognostic scores for Covid-19. Final data extraction was performed on May 8th, 2021. The GPUH's CDW Scientific and Ethics Committee (IRB00011591) granted access to the CDW for the purpose of this study and no linkage was made with other databases.

Inclusion and exclusion criteria

Patients' selection process is summarized in **Figure 1**. All patients with a result found in the database for reverse transcriptase-polymerase chain reaction (PCR) for SARS-CoV2 in a respiratory sample were screened. Patients were included in the study if they met both following criteria:

- A hospital stay with an International Classification of Diseases, 10th edition (ICD-10) code for Covid-19 (U07.1),
- At least one positive respiratory PCR for SARS-CoV2 from 10 days before to 3 days after hospital admission.

Patients were excluded from the study if they met at least one of the following criteria:

- PCR result considered unreliable (i.e., time of validation by the biologist before the time of PCR sample collection, or more than 20 days after the time of sample collection),
- Asymptomatic positive PCR result during a COVID-unrelated hospitalization or COVID considered as hospital-acquired (i.e., a first positive PCR sample collected more than 3 days after hospital admission),

- Direct ICU admission (i.e., time between recorded hospital admission and recorded ICU admission less than 2 hours and no visit in another GPUH hospital in the preceding 24 hours),
- Age <18, not recorded or unknown,
- Hospitalization in Georges Pompidou European hospital, one of the 39 GPUH hospitals (as all biological and clinical data from this hospital were missing, due to interoperability issues with the CDW).

To have a follow-up of 30 days or more for all hospitalized patients, only patients with a PCR performed before March 30th were considered.

Data collection

The reference date used for baseline characteristics was the date of hospital admission for Covid-19. The following data were collected:

- Demographic data and data on hospital admission,
- Medical history (based on ICD-10 codes for current or previous hospital visits; the list of codes used is based on a previously published work [10]),
- Vital signs and biological values (the first value found in the database from 24 hours before to 48 hours after hospital admission was retrieved for each patient, as a delay can exist for logistical reasons between true and recorded admission date; values obtained in ICU were not considered),
- Outcomes (in-hospital mortality, ICU admission and invasive mechanical ventilation within 30 days from admission).

Of note, invasive mechanical ventilation is always performed in ICU in France.

Selection of published scores

The selection of high quality published scores was performed using "Covid-19 Evidence Alerts" (https://plus.mcmaster.ca/Covid-19/), a service provided by the McMaster University, in which evidence reports on Covid-19 published in all journals included in MEDLINE are critically appraised for scientific merit based on prespecified criteria (see https://hiru.mcmaster.ca/hiru/InclusionCriteria.html). All studies identified by the "Clinical Prediction

Guide" filter were systematically screened by two independent investigators (L.A. and P.S.), and discrepancies were adjudicated by a third investigator (Y.L.). Studies were included if they met all the following criteria:

- studies on prognostic scores predicting ICU transfer or in-hospital mortality for patients hospitalized for Covid-19, including scores primarily developed for other purposes prior to the pandemic,
- meeting all the prespecified criteria for "higher quality" (i.e. generated in one or more sets of real patients; validated in another set of real patients; study providing information on how to apply the prediction guide); or studies excluded from this category only due to the lack of an independent validation cohort, but in which derivation and validation were performed in different samples from the same cohort (split validation),
- computable with the data collected in the CDW.

The last search in "Covid-19 Evidence Alerts" was performed on April 3rd, 2021. The process for scores' selection and reasons for exclusion are detailed in **Appendix 3** and **Figure S1,** and information on scores included in the study in **Table S1** and **S2**.

Statistical analysis

Aberrant values for biological tests and vital signs were treated as described in **Table S3**. Missing data were treated by multiple imputations (*mice* function of the mice package, 50 imputed datasets with 15 iterations, predictive means matching method for quantitative variables, after log or square-root transformation when needed to get a more normalised dataset), under the missing-at-random hypothesis. Outcome variables were included in the dataset used for imputation. Rubin's rule was used to pool estimates obtained in each imputed dataset. Variables used for multiple imputations are detailed in **Table S4**.

For each score included in the analysis and each outcome, discrimination was assessed by drawing a receiver operating characteristics (ROC) curve and computing the corresponding area under the curve (AUC). DeLong's method [11] was used to estimate the variance in each dataset, results were pooled with Rubin's rule and used to compute pooled 95% confidence intervals.

First, we assessed the performance of each score to predict the available outcome closest to the one used in the original study, with the required adaptations to be computed with the available data. AUC in our cohort and in previously published studies were compared using a Z-test for independent samples. Second, we assessed the performance of each score to predict 30-day in-hospital mortality and the composite of 30-day in-hospital mortality or ICU transfer. Third, we used a Z-test for paired data following DeLong's method [11] to compare the accuracy of scores with an AUC >0.75 to predict 30-day hospital mortality. Sensitivity analyses were conducted on subgroups of age (≤65 or >65 years old) or wave of admission (before or after June 15th, 2020, a graphically determined threshold), considering only complete cases (only patients with all data available to compute a given score), and considering the area under the precision-recall curve instead of under the ROC curve (*pr.curve* function of the PRROC package). Heterogeneity of AUC between subgroups was assessed using an interaction term between the score and the grouping variable in a logistic regression model predicting the outcome.

Post hoc analyses were performed to further characterize the best scores at predicting 30-day in-hospital mortality (AUC > 0.75). Calibration curves were drawn by plotting the observed mortality rate in each class as a function of the predicted probability of mortality, with patients grouped by deciles of predicted probability. For each score, a logistic regression model was built to predict 30-day in-hospital mortality with its predictors and fitted on our data. Variable importance was determined using the absolute value of the t-statistic for each predictor in this model (*varImp* function of the caret package). Calibration curves were drawn using probabilities predicted by the revised logistic regression models fitted on our data.

All tests are two-sided, and a p-value <0.05 was considered significant. Continuous variables are reported as mean (standard deviation) for normally distributed variables, and median [interquartile range] for non-normally distributed variables. Binary variables are reported as number of patients with a positive result (percentage of patients with a positive result). Analyses were performed using the R freeware version 4 (packages mice, pROC, psfmi, Amelia, PRROC, caret).

RESULTS

Baseline characteristics and outcomes of patients included in the study

We included 14,343 patients in the validation cohort (**Figure 1**). First hospital admission for Covid-19 was on January 29th, 2020 and last on April 6th, 2021. Patients' baseline characteristics are summarized in **Table 1** and outcomes are summarized in **Table 2**. Baseline characteristics appeared similar during the first wave and subsequent waves (**Table S5**). Initial care site appeared to be an important factor for vital signs or biological values to be missing (**Table S6**). Multiple imputations were therefore stratified by centre. In-hospital mortality at day 30 was 18% overall, significantly lower during the first wave than in the subsequent waves, and significantly higher in patients older than 65 years old (**Figure S2**, $p<0.001$ for Log-Rank test).

Selected scores and their performance to predict the original outcome

Thirty-two scores [12–37] were included in the study: 23 were specifically derived in Covid-19 patients and 9 were pre-existing scores developed for other purposes and tested in Covid-19 patients (**Table 3**, **Table S1** and **S2**, **Appendix 3**). Among 27 scores with available 95% CI to estimate AUC variance in previous reports, 19 (70%) had an AUC significantly lower in our cohort (**Table 3**). The 4C Mortality Score was the only one with an AUC significantly higher in our cohort compared to the previously published value ($p<0.001$).

Performance to predict 30-day in-hospital mortality and the composite of 30-day in-hospital mortality or ICU admission

Results are summarized in **Table S7**, and **Figure S3** shows the ROC curves of the three most accurate scores for each outcome. None of the included scores had a very high accuracy to predict 30-day in-hospital mortality alone, or the composite of 30-day in-hospital mortality or ICU admission (all AUC <0.8). AUC was higher to predict 30-day in-hospital mortality alone than 30-day in-hospital mortality or ICU admission for 25/32 scores (78%).

Seven scores had an AUC >0.75 to predict 30-day in-hospital mortality (**Table 4**). The 4C Mortality and the ABCS scores had the highest AUC to predict 30-day in-hospital mortality (4C Mortality score: 0.793, 95% CI: 0.783 to 0.803; ABCS score: 0.790, 95% CI: 0.780 to 0.801). Their AUC did not differ significantly from each other ($p=0.61$) but were significantly higher than that of the following scores ($p<0.01$ for all comparisons). The CORONATION-TR score had the highest AUC to predict 30-day in-hospital mortality or ICU admission (AUC 0.724, 95% CI: 0.714 to 0.733). **Table S8** provides the sensitivities and specificities for these scores to predict in-hospital mortality using cut-off values from previous reports, and **Figure S4** shows the Kaplan-Meier curves for in-hospital mortality of the three scores that performed best to predict in-hospital mortality.

Sensitivity and post hoc analyses

Among the seven scores with an AUC >0.75 to predict 30-day in-hospital mortality: accuracy was not significantly altered by wave of admission for any of them (**Table S9**); accuracy was significantly lower in the subgroup of patients >65 years-old for two of them (RISE-UP and COVID-19 SEIMC; **Table S10**); AUC was <0.75 in the analysis using complete cases for one of them (CORONATION-TR; **Table S7**); the 4C Mortality Score ranked first to predict in-hospital mortality in analyses using multiple imputed data and analyses using complete cases (**Table S7**).

Main results were unchanged when using the area under the precision-recall curve instead of under the ROC curve to measure discriminative ability: the 4C Mortality score and the ABCS ranked first and second to predict 30-day in-hospital mortality, and the CORONATION-TR score ranked first to predict 30-day in-hospital mortality or ICU transfer (**Table S11**).

As shown by calibration curves (**Figure S5**), the risk of 30-day in-hospital mortality was overestimated for 6/7 scores (all but the CORONATION-TR), and most notably so for the COVID-GRAM and ANDC scores. Overestimation was overall less important during the first epidemic wave than subsequent waves (**Figure S5**) and was corrected after logistic coefficients revision (**Figure S6**).

In variable importance analysis, age was the most important factor to predict 30-day in-hospital mortality in 5 scores (4C Mortality, ANDC, CORONATION-TR, COVID-GRAM, RISE UP), troponin positivity in 1 score (ABCS) and low estimated glomerular filtration rate in 1 score (COVID-19 SEIMC) (**Figure S7**).

DISCUSSION

Key results

Most scores (19/27 with available data for comparison) had a significantly lower accuracy in our study compared to previously published studies, and most scores (25/32) had a lower accuracy to predict the composite outcome of 30-day in-hospital mortality or ICU admission, compared to 30-day in-hospital mortality alone. Seven scores had a high accuracy (AUC >0.75) for the prediction of 30-day in-hospital mortality: the 4C Mortality and ABCS scores had significantly higher AUC values compared to the other scores; the CORONATION-TR score was the most accurate to predict in-hospital mortality or ICU admission; the RISE-UP and COVID-19 SEIMC scores were less accurate in the subgroup of patients >65 years-old. The discriminative performance of these scores was not altered by wave of admission despite changes in clinical care such as larger use of corticosteroids and lower use of invasive ventilation during the subsequent waves. On the opposite, calibration was poorer during the second and subsequent waves than in the first wave.

Limitations and strengths

We conducted a large, multicentre, independent study to validate systematically selected prognostic scores for Covid-19, using routine clinical care data. Selection criteria were chosen to identify the most promising scores, although many of them had not yet been externally validated or had been validated in small cohorts only. Outcomes used in our study (in-hospital mortality, ICU admission and invasive mechanical ventilation) are of high clinical importance, objective and reliably collected in the CDW.

The main limitations of our study are consequences of its retrospective design, with a risk for selection and information bias. Selection bias was controlled using objective and reproducible inclusion and exclusion criteria, based on both administrative (ICD-10 codes for Covid-19) and microbiological information (PCR for SARS-CoV2). This information is exhaustively recorded in the database, as ICD-10 codes for all hospital stays are independently assessed by a trained physician or technician before transmission to the national health insurance service for billing. Information bias for comorbidities and medical history was controlled by collecting ICD-10 codes for both index and previous visits, using a systematic procedure that was independently validated in a medico-administrative database whose structure is similar to ours [10]. Missing physiological values, such as oxygen saturation, respiratory rate, are explained by several templates available to record them in electronic health records. Only a limited number of these templates are used to gather and aggregate these data in the CDW. Missing biological values, such as D-dimers, CRP or ferritin, are explained by unstandardized practices across GPUH hospitals. As a result, the rate of missing values varied across centres for physiological and biological values (see **Table S6**), and was high for several important variables such as the Glasgow coma scale. To control these biases, we used multiple imputations under the missing-at-random hypothesis [38], taking centres into account, and we performed a confirmatory sensitivity analysis using complete data.

Several scores, based on machine- or deep-learning algorithms, or using data rarely collected for initial evaluation of patients in clinical practice (such as myoglobin or interleukins) could not be computed in our cohort (see **Appendix 3**). Although for many of them discriminative performance seemed high in previous studies, their use in clinical practice is more difficult, as they would require changing protocols for patients' initial evaluation to add costly biological tests, and, for machine- or deep-learning based algorithms, to set an automatic system for computation. Further prospective pragmatic studies are needed on these matters.

Interpretation and generalisability

Our cohort includes patients from Paris and its suburbs, with various ethnicities and socioeconomic backgrounds [39]. Patients are treated in various hospitals, each of them having different resources and practices. Our validation study is strengthened by the number and diversity of included patients and settings, and by the independence from all cohorts used for the derivation and first validation of investigated prognostic scores. Patients were consecutively recruited, and the number of outcome events was very large, overcoming two major shortcomings of previous validation studies. For example, several included scores were previously validated in less than 100 patients (Table 3). The waste of time and money on inappropriately designing or validating Covid 19 prognostic scores have been stressed in a living systematic review [6].

Using a cut-off value of 0.75 for AUC to predict in-hospital death, seven scores were identified in as having a high accuracy. They differ in characteristics that may influence their choice for a given use in a given clinical context. For example, some scores use costly biological tests and are not adapted for countries with limited resources; some use many variables and may be hard to compute at the bedside; some are less accurate in older patients; some are more accurate to predict ICU admission and therefore more suitable to predict the demand on healthcare systems. For the seven fairly accurate scores identified, we provide detailed characteristics that can help clinicians choose the best suited to their needs (**Table 4**). The 4C Mortality and ABCS scores appear to be the most promising ones, as they use a limited number of variables that are available in routine clinical care, had a fair accuracy in our external validation study, and performed equally well during the first epidemic wave and subsequent waves, and in younger and older patients.

The risk of 30-day in-hospital mortality was overestimated by 6/7 scores (all but the CORONATION-TR), and more so during the second and subsequent waves. This can be explained by overall better outcomes during these waves, as seen in our study and in other ones [40]. Many published scores were derived and validated on first wave data. Revising the scores using local and current data is necessary if accurate estimations of the mortality risk are needed. Likewise, the thresholds indicating a high risk of poor outcome should be locally defined.

In variable importance analysis, age was the most influential factor in 5/7 scores, even in those including many clinical and biological variables (for example, the CORONATION-TR score), underlining the importance of age in driving severity among hospitalized Covid-19 patients. Elevated baseline troponin was the most important factor in the ABCS, which discriminated and calibrated well in our cohort. Troponin has been previously shown to be independently associated with mortality in both non-ICU [41] and ICU [42] patients, stressing its potential relevance for risk stratification at bedside.

The place these scores could have to guide therapeutic strategies is yet to be determined. Their most promising use may be as a tool to guide hospital admission, in the context of a pandemic with a high demand and a low offer for hospital beds, especially in low-income countries [43, 44]. Further studies should be conducted on this important issue.

Scores specifically derived for Covid-19 outperformed generic scores for infectious pneumonia or for sepsis. This highlights the specificity of Covid-19 in comparison to other forms of pneumonia, with a key role for the inflammatory and pro-thrombotic status to drive severity [45–47]. However, given their simplicity of use and their good performance to predict in-hospital mortality in our cohort, scores such as the CURB-65 or A-DROP scores could still be considered for risk stratification in Covid-19 patients. On the opposite, scores used in sepsis such as qSOFA or SIRS seemed to offer no clear benefit for risk stratification. Low specificity can be explained by a limited number of factors used for initial evaluation, as many patients present with abnormal vital signs or white blood cells counts, and those factors alone are insufficient to identify patients at high risk for critical illness. Low sensitivity can be explained as patients truly at risk for critical illness (particularly the elderly or patients with many comorbidities) may initially appear clinically stable before suddenly and dramatically worsening.

Accuracy was lower in our cohort to predict ICU admission compared to in-hospital mortality, even for scores specifically aimed at predicting this endpoint. This could partly be explained by the complexity of ICU admission criteria, which may differ across countries according to local guidelines and demography, and may vary with time given the pressure on ICU beds [48]. In France for example, during the first wave of the pandemic, some patients with invasive mechanical ventilation urgently initiated in the emergency room or in general wards could not be transferred to the hospital-related ICU due to shortage of beds, and were transferred to other hospitals, either in the Paris region or in other regions [49].

In conclusion, several scores using routinely collected clinical and biological data have a fair accuracy to predict in-hospital death. The 4C Mortality Score and the ABCS stand out because they performed as well in our cohort and their initial validation cohort, during the first epidemic wave and subsequent waves, and in younger and older patients. Their use to guide appropriate clinical care and resource utilization should be evaluated in future studies.

DECLARATIONS

Funding

None.

Conflicts of interest

None for any of the authors.

Availability of data and material

Raw data cannot be transmitted to non-GPUH staff without specific authorization from the GPUH CDW Scientific and Ethics Committee.

Code availability

R scripts are available at request to the corresponding author.

Authorship statement

YL: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Writing-Original Draft, Writing-Review & Editing, Visualization; LA, PS: Validation, Methodology, Writing-Review & Editing; GL, MB: Software, Writing-Review & Editing; JL, AB: Software, Methodology, Writing-Review & Editing; QR: Methodology, Writing-Original Draft, Writing-Review & Editing; OS: Conceptualization, Methodology, Formal Analysis, Writing-Original Draft, Writing-Review & Editing, Supervision, Project Administration.

Ethics approval

The GPUH's CDW Scientific and Ethics Committee (IRB00011591) granted access to the CDW for the purpose of this study (authorization n°200063).

Take home message

In this retrospective cohort study of 14,343 patients, seven out of thirty-two previously published prognostic scores were able to fairly predict 30-day in-hospital mortality using routinely collected clinical and biological data (area under the ROC curve > 0.75). The 4C Mortality Score and the ABCS stand out because they performed as well in our cohort and their initial validation cohort, during the first and subsequent epidemic waves, in younger and older patients, and showed satisfactory calibration. Their ability to guide clinical management decisions and appropriate resource allocation should now be evaluated in future studies.

Tweet

The 4C Mortality Score and the ABCS predicted death as well in a cohort of 14,343 hospitalized COVID patients than in their original study.

REFERENCES

1. WHO | Novel Coronavirus – China. In: WHO. http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/. Accessed 21 Feb 2021

2. Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20:533–534. https://doi.org/10.1016/S1473-3099(20)30120-1

3. Wiersinga WJ, Rhodes A, Cheng AC, et al (2020) Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA 324:782. https://doi.org/10.1001/jama.2020.12839

4. (2020) Clinical characteristics and day-90 outcomes of 4244 critically ill adults with COVID-19: a prospective cohort study. Intensive Care Med 1–14. https://doi.org/10.1007/s00134-020-06294-x

5. Steinberg E, Balakrishna A, Habboushe J, et al (2020) Calculated decisions: COVID-19 calculators during extreme resource-limited situations. Emerg Med Pract 22:CD1–CD5

6. Wynants L, Calster BV, Bonten MMJ, et al (2020) Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ 369:. https://doi.org/10.1136/bmj.m1328

7. Balmford B, Annan JD, Hargreaves JC, et al (2020) Cross-Country Comparisons of Covid-19: Policy, Politics and the Price of Life. Environ Resour Econ 76:525–551. https://doi.org/10.1007/s10640-020-00466-5

8. Benchimol EI, Smeeth L, Guttmann A, et al (2015) The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Med 12:e1001885. https://doi.org/10.1371/journal.pmed.1001885

9. Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 350:g7594. https://doi.org/10.1136/bmj.g7594

10. Bannay A, Chaignot C, Blotière P-O, et al (2016) The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. Med Care 54:188–194. https://doi.org/10.1097/MLR.0000000000000471

11. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845

12. Knight SR, Ho A, Pius R, et al (2020) Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 370:m3339. https://doi.org/10.1136/bmj.m3339

13. Mejía-Vilet JM, Córdova-Sánchez BM, Fernández-Camargo DA, et al (2020) A risk score to predict admission to the intensive care unit in patients with Covid-19: the ABC-GOALS score. Salud Publica Mex. https://doi.org/10.21149/11684

14. Jiang M, Li C, Zheng L, et al (2021) A biomarker-based age, biomarkers, clinical history, sex (ABCS)-mortality risk score for patients with coronavirus disease 2019. Ann Transl Med 9:. https://doi.org/10.21037/atm-20-6205

15. Weng Z, Chen Q, Li S, et al (2020) ANDC: an early warning score to predict mortality risk for patients with Coronavirus Disease 2019. J Transl Med 18:328. https://doi.org/10.1186/s12967-020-02505-7

16. Bennouar S, Bachir Cherif A, Kessira A, et al (2021) Development and validation of a laboratory risk score for the early prediction of COVID-19 severity and in-hospital mortality. Intensive Crit Care Nurs 103012. https://doi.org/10.1016/j.iccn.2021.103012

17. Ruocco G, McCullough PA, Tecson KM, et al (2020) Mortality Risk Assessment Using CHA(2)DS(2)-VASc Scores in Patients Hospitalized With Coronavirus Disease 2019 Infection. Am J Cardiol 137:111–117. https://doi.org/10.1016/j.amjcard.2020.09.029

18. Cho S-Y, Park S-S, Song M-K, et al (2021) Prognosis Score System to Predict Survival for COVID-19 Cases: a Korean Nationwide Cohort Study. J Med Internet Res 23:e26257. https://doi.org/10.2196/26257

19. Tanboğa IH, Canpolat U, Çetin EHÖ, et al (2021) Development and validation of clinical prediction model to estimate the probability of death in hospitalized patients with COVID-19: Insights from a nationwide database. J Med Virol 93:3015–3022. https://doi.org/10.1002/jmv.26844

20. Berenguer J, Borobia AM, Ryan P, et al (2021) Development and validation of a prediction model for 30-day mortality in hospitalised patients with COVID-19: the COVID-19 SEIMC score. Thorax. https://doi.org/10.1136/thoraxjnl-2020-216001

21. Hajifathalian K, Sharaiha RZ, Kumar S, et al (2020) Development and external validation of a prediction risk model for short-term mortality among hospitalized U.S. COVID-19 patients: A proposal for the COVID-AID risk tool. PLOS ONE 15:e0239536. https://doi.org/10.1371/journal.pone.0239536

22. Liang W, Liang H, Ou L, et al (2020) Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Intern Med 180:1081–1089. https://doi.org/10.1001/jamainternmed.2020.2033

23. Ebell MH, Cai X, Lennon R, et al (2021) Development and Validation of the COVID-NoLab and COVID-SimpleLab Risk Scores for Prognosis in 6 US Health Systems. J Am Board Fam Med JABFM 34:S127–S135. https://doi.org/10.3122/jabfm.2021.S1.200464

24. Hachim MY, Hachim IY, Naeem KB, et al (2020) D-dimer, Troponin, and Urea Level at Presentation With COVID-19 can Predict ICU Admission: A Single Centered Study. Front Med 7:. https://doi.org/10.3389/fmed.2020.585003

25. Hu H, Yao N, Qiu Y (2020) Comparing rapid scoring systems in mortality prediction of critical ill patients with novel coronavirus disease. Acad Emerg Med Off J Soc Acad Emerg Med. https://doi.org/10.1111/acem.13992

26. Jamal MH, Doi SA, AlYouha S, et al (2020) A Biomarker Based Severity Progression Indicator for COVID-19: The Kuwait Prognosis Indicator Score. Biomark Biochem Indic Expo Response Susceptibility Chem 1–21. https://doi.org/10.1080/1354750X.2020.1841296

27. Soto-Mota A, Marfil-Garza BA, Rodríguez EM, et al (2020) The low-harm score for predicting mortality in patients diagnosed with COVID-19: A multicentric validation study. J Am Coll Emerg Physicians Open 1:1436–1443. https://doi.org/10.1002/emp2.12259

28. Mei Y, Weinberg SE, Zhao L, et al (2020) Risk stratification of hospitalized COVID-19 patients through comparative studies of laboratory results with influenza. EClinicalMedicine 26:100475. https://doi.org/10.1016/j.eclinm.2020.100475

29. Myrstad M, Ihle-Hansen H, Tveita AA, et al (2020) National Early Warning Score 2 (NEWS2) on admission predicts severe disease and in-hospital mortality from Covid-19 - a prospective cohort study. Scand J Trauma Resusc Emerg Med 28:66. https://doi.org/10.1186/s13049-020-00764-3

30. Li J, Chen Y, Chen S, et al (2020) Derivation and validation of a prognostic model for predicting in-hospital mortality in patients admitted with COVID-19 in Wuhan, China: the PLANS (platelet lymphocyte age neutrophil sex) model. BMC Infect Dis 20:959. https://doi.org/10.1186/s12879-020-05688-y

31. Bartoletti M, Giannella M, Scudeller L, et al (2020) Development and validation of a prediction model for severe respiratory failure in hospitalized patients with SARS-CoV-2 infection: a multicentre cohort study (PREDI-CO study). Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis 26:1545–1553. https://doi.org/10.1016/j.cmi.2020.08.003

32. Saberian P, Tavakoli N, Hasani-Sharamin P, et al (2020) Accuracy of the pre-hospital triage tools (qSOFA, NEWS, and PRESEP) in predicting probable COVID-19 patients' outcomes transferred by Emergency Medical Services. Casp J Intern Med 11:536–543. https://doi.org/10.22088/cjim.11.0.536

33. van Dam PM, Zelis N, Stassen P, et al (2021) Validating the RISE UP score for predicting prognosis in patients with COVID-19 in the emergency department: a retrospective study. BMJ Open 11:e045141. https://doi.org/10.1136/bmjopen-2020-045141

34. Ageno W, Cogliati C, Perego M, et al (2021) Clinical risk scores for the early prediction of severe outcomes in patients hospitalized for COVID-19. Intern Emerg Med. https://doi.org/10.1007/s11739-020-02617-4

35. Holten AR, Nore KG, Tveiten CEVWK, et al (2020) Predicting severe COVID-19 in the Emergency Department. Resusc Plus 4:100042. https://doi.org/10.1016/j.resplu.2020.100042

36. Demir MC, Ilhan B (2021) Performance of the Pandemic Medical Early Warning Score (PMEWS), Simple Triage Scoring System (STSS) and Confusion, Uremia, Respiratory rate, Blood pressure and age ≥ 65 (CURB-65) score among patients with COVID-19 pneumonia in an emergency department triage setting: a retrospective study. Sao Paulo Med J Rev Paul Med 139:170–177. https://doi.org/10.1590/1516-3180.2020.0649.R1.10122020

37. Wang K, Zuo P, Liu Y, et al (2020) Clinical and Laboratory Predictors of In-hospital Mortality in Patients With Coronavirus Disease-2019: A Cohort Study in Wuhan, China. Clin Infect Dis 71:2079–2088. https://doi.org/10.1093/cid/ciaa538

38. Sterne JAC, White IR, Carlin JB, et al (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 338:b2393. https://doi.org/10.1136/bmj.b2393

39. Jannot A-S, Coutouris H, Burgun A, et al (2020) COVID-19, a social disease in Paris: a socio-economic wide association study on hospitalized patients highlights low-income neighbourhood as a key determinant of severe COVID-19 incidence during the first wave of the epidemic. medRxiv 2020.10.30.20222901. https://doi.org/10.1101/2020.10.30.20222901

40. Kurtz P, Bastos LSL, Dantas LF, et al (2021) Evolving changes in mortality of 13,301 critically ill adult patients with COVID-19 over 8 months. Intensive Care Med 47:538–548. https://doi.org/10.1007/s00134-021-06388-0

41. Du R-H, Liang L-R, Yang C-Q, et al (2020) Predictors of Mortality for Patients with COVID-19 Pneumonia Caused by SARS-CoV-2: A Prospective Cohort Study. Eur Respir J. https://doi.org/10.1183/13993003.00524-2020

42. Azoulay E, Fartoukh M, Darmon M, et al (2020) Increased mortality in patients with severe SARS-CoV-2 infection admitted within seven days of disease onset. Intensive Care Med 46:1714–1722. https://doi.org/10.1007/s00134-020-06202-3

43. Ma X, Vervoort D (2020) Critical care capacity during the COVID-19 pandemic: Global availability of intensive care beds. J Crit Care 58:96–97. https://doi.org/10.1016/j.jcrc.2020.04.012

44. Sen-Crowe B, Sutherland M, McKenney M, Elkbuli A (2021) A Closer Look Into Global Hospital Beds Capacity and Resource Shortages During the COVID-19 Pandemic. J Surg Res 260:56–63. https://doi.org/10.1016/j.jss.2020.11.062

45. McElvaney OJ, McEvoy NL, McElvaney OF, et al (2020) Characterization of the Inflammatory Response to Severe COVID-19 Illness. Am J Respir Crit Care Med 202:812–821. https://doi.org/10.1164/rccm.202005-1583OC

46. Goligher EC, Ranieri VM, Slutsky AS (2020) Is severe COVID-19 pneumonia a typical or atypical form of ARDS? And does it matter? Intensive Care Med 1–3. https://doi.org/10.1007/s00134-020-06320-y

47. Helms J, Tacquard C, Severac F, et al (2020) High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study. Intensive Care Med 46:1089–1098. https://doi.org/10.1007/s00134-020-06062-x

48. Sprung CL, Joynt GM, Christian MD, et al (2020) Adult ICU Triage During the Coronavirus Disease 2019 Pandemic: Who Will Live and Who Will Die? Recommendations to Improve Survival. Crit Care Med. https://doi.org/10.1097/CCM.0000000000004410

49. Painvin B, Messet H, Rodriguez M, et al (2021) Inter-hospital transport of critically ill patients to manage the intensive care unit surge during the COVID-19 pandemic in France. Ann Intensive Care 11:54. https://doi.org/10.1186/s13613-021-00841-5

LEGENDS

Table 1. Baseline characteristics of patients included in the study.

Table 2. Outcomes of patients included in the study.

Table 3. Summary of scores included in the study and comparison to previously published data.

Table 4. Detailed characteristics of scores with an AUROC > 0.75 to predict 30-day in-hospital mortality in the analysis using multiple imputed data.

Figure 1. Flow chart of included patients.

TABLES

Variable		No death within 30 days† (n = 11760)		Death within 30 days (n = 2583)		All patients (n = 14343)
Demographic data	*Missing*		*Missing*		*Missing*	
Female sex, n (%)		5175 (44)		1014 (39.3)		6189 (43.1)
Age, years		66 (SD 17.6)		79.2 (SD 12)		68.4 (SD 17.5)
Diagnosis of Covid-19						
Admission during « first wave », n (%)		4863 (41.4)		1279 (49.5)		6142 (42.8)
Time between PCR and admission, days		-0.1 [-0.1, 0]		0 [-0.1, 0]		-0.1 [-0.1, 0]
Medical history, n (%)						
Modified Charlson comorbidity index, pts		0 [0, 2]		2 [0, 4]		1 [0, 2]
Congestive heart failure		1228 (10.4)		637 (24.7)		1865 (13)
Myocardial infarction		666 (5.7)		297 (11.5)		963 (6.7)
Peripheral vascular disease		620 (5.3)		264 (10.2)		884 (6.2)
Cerebrovascular disease		985 (8.4)		376 (14.6)		1361 (9.5)
Hemiplegia		442 (3.8)		157 (6.1)		599 (4.2)
Dementia		1364 (11.6)		638 (24.7)		2002 (14)
Arterial hypertension		4723 (40.2)		1403 (54.3)		6126 (42.7)
Diabetes		2699 (23)		716 (27.7)		3415 (23.8)
Diabetes with end-organ damage		1480 (12.6)		542 (21)		2022 (14.1)
Chronic pulmonary disease		1366 (11.6)		397 (15.4)		1763 (12.3)
Moderate or severe renal disease		1536 (13.1)		660 (25.6)		2196 (15.3)
Moderate or severe liver disease		127 (1.1)		33 (1.3)		160 (1.1)
Any tumor		1064 (9)		480 (18.6)		1544 (10.8)
Metastatic solid tumor		261 (2.2)		150 (5.8)		411 (2.9)
Connective tissue disease		241 (2)		64 (2.5)		305 (2.1)
HIV infection		218 (1.9)		20 (0.8)		238 (1.7)
Obesity (ICD-10 codes only)		2289 (19.5)		426 (16.5)		2715 (18.9)
Vital signs on admission						
Heart rate, beats per minute	*2729 (23.2)*	88.7 (SD 17.5)	*615 (23.8)*	87.5 (SD 18.5)	*3344 (23.3)*	88.5 (SD 17.7)
Respiratory rate, cycles per minute	*4623 (39.3)*	24.4 (SD 7.3)	*992 (38.4)*	27 (SD 8.1)	*5615 (39.1)*	24.9 (SD 7.5)
Altered consciousness, n (%)	*7008 (59.6)*	133 (2.8)	*1573 (60.9)*	112 (11.1)	*8581 (59.8)*	245 (4.3)
Diastolic blood pressure, mmHg	*4934 (42)*	75.3 (SD 14.5)	*1046 (40.5)*	72.4 (SD 17.1)	*5980 (41.7)*	74.8 (SD 15.1)
Mean blood pressure, mmHg	*5201 (44.2)*	94.4 (SD 15.2)	*1276 (49.4)*	91.3 (SD 17.6)	*6477 (45.2)*	93.9 (SD 15.6)
Systolic blood pressure, mmHg	*4932 (41.9)*	131.4 (SD 21.3)	*1044 (40.4)*	130.7 (SD 24.8)	*5976 (41.7)*	131.2 (SD 22)
Pulse saturometry, %	*3767 (32)*	96 [93, 98]	*784 (30.4)*	94 [90, 97]	*4551 (31.7)*	96 [93, 98]
Temperature, °C	*2759 (23.5)*	37.5 (SD 0.9)	*615 (23.8)*	37.5 (SD 1)	*3374 (23.5)*	37.5 (SD 1)
Body mass index (BMI), kg/m²	*4227 (35.9)*	27.2 (SD 6.4)	*1208 (46.8)*	26.6 (SD 7.1)	*5435 (37.9)*	27.1 (SD 6.5)
Biological values on admission						
Haemoglobin, g/dl	*1376 (11.7)*	13.1 (SD 1.9)	*383 (14.8)*	12.7 (SD 2.2)	*1759 (12.3)*	13 (SD 2)
Leukocytes, G/l	*1378 (11.7)*	7 (SD 3.7)	*384 (14.9)*	8 (SD 5.1)	*1762 (12.3)*	7.2 (SD 4)
Neutrophils, G/l	*1574 (13.4)*	5.3 (SD 3.1)	*416 (16.1)*	6.4 (SD 4.1)	*1990 (13.9)*	5.5 (SD 3.4)
Lymphocytes, G/l	*1597 (13.6)*	1 [0.7, 1.4]	*423 (16.4)*	0.8 [0.5, 1.1]	*2020 (14.1)*	0.9 [0.7, 1.3]
Platelets count, G/l	*1385 (11.8)*	223.5 (SD 93)	*384 (14.9)*	201.9 (SD 92.9)	*1769 (12.3)*	219.7 (SD 93.4)
Sodium, mmol/l	*467 (4)*	135.9 (SD 4.3)	*132 (5.1)*	136.6 (SD 6.2)	*599 (4.2)*	136 (SD 4.7)
Potassium, mmol/l	*652 (5.5)*	4.1 (SD 0.6)	*196 (7.6)*	4.2 (SD 0.7)	*848 (5.9)*	4.1 (SD 0.6)
Bicarbonates, mmol/l	*5361 (45.6)*	24.4 (SD 3.7)	*1196 (46.3)*	23 (SD 4.4)	*6557 (45.7)*	24.2 (SD 3.9)
Proteins, g/l	*796 (6.8)*	71.8 (SD 7.1)	*186 (7.2)*	69.8 (SD 8.1)	*982 (6.8)*	71.5 (SD 7.3)
Urea, mmol/l	*663 (5.6)*	6 [4.3, 8.8]	*168 (6.5)*	10 [6.6, 15.3]	*831 (5.8)*	6.5 [4.6, 9.9]
Serum creatinine, µmol/l	*436 (3.7)*	80 [64, 103]	*124 (4.8)*	103 [77, 152]	*560 (3.9)*	82.4 [66, 110]
Alanine aminotransferase, IU/l	*1995 (17)*	30 [20, 47.5]	*482 (18.7)*	28 [18.6, 45]	*2477 (17.3)*	29.5 [20, 47]
Asparate aminotransferase, IU/l	*2366 (20.1)*	41 [29, 60]	*560 (21.7)*	51 [34, 78]	*2926 (20.4)*	42 [29.2, 63]
Total bilirubin, µmol/l	*1959 (16.7)*	8 [6, 11.5]	*468 (18.1)*	9 [6, 13]	*2427 (16.9)*	8 [6, 12]
Lactate dehydrogenase, IU/l	*5688 (48.4)*	352 [267, 477]	*1273 (49.3)*	430 [322, 581]	*6961 (48.5)*	362 [275, 499]
Creatinine phosphokinase, IU/l	*5470 (46.5)*	123 [64, 276]	*1200 (46.5)*	186 [85 480]	*6670 (46.5)*	132 [67, 300]
Troponine, ng/l	*6149 (52.3)*	15 [9, 24]	*1283 (49.7)*	34 [18, 76.1]	*7432 (51.8)*	15 [10, 31]
Activated partial thromboplastin time	*2555 (21.7)*	1.2 (SD 0.3)	*605 (23.4)*	1.3 (SD 0.4)	*3160 (22)*	1.2 (SD 0.3)
Prothrombin time, %	*2238 (19)*	87 [76, 98]	*535 (20.7)*	82 [69, 93]	*2773 (19.3)*	87 [75, 97]
Fibrinogen, g/l	*4248 (36.1)*	5.8 (SD 1.6)	*952 (36.9)*	5.8 (SD 1.6)	*5200 (36.3)*	5.8 (SD 1.6)
D-dimers, µg/l	*4918 (41.8)*	900 [557, 1560]	*1287 (49.8)*	1375 [828, 2560]	*6205 (43.3)*	964 [585, 1690]
C-reactive protein, mg/l	*1104 (9.4)*	65 [26, 121]	*261 (10.1)*	96 [49.1, 163.9]	*1365 (9.5)*	70 [30, 129]
Procalcitonin, µg/l	*5973 (50.8)*	0.1 [0.1, 0.3]	*1263 (48.9)*	0.3 [0.2, 1]	*7236 (50.4)*	0.2 [0.1, 0.4]
Albumin, g/l	*7792 (66.3)*	32.7 (SD 5.4)	*1659 (64.2)*	30.9 (SD 5.4)	*9451 (65.9)*	32.4 (SD 5.5)

†Either patients discharged alive before day 30 (n=8459), or patients still in hospital and alive at day 30 (n=3301). SD: standard deviation. Continuous variables are reported as mean (SD) for normally distributed variables and median [interquartile range] for non-normally distributed variables.

Table 1. Baseline characteristics of patients included in the study.

Outcome	All patients (n = 14343)
In-hospital mortality†, n (%)	2583 (18)
Time between hospital admission and death, days	8.1 [4.2, 13.7]
ICU admission†, n (%)	3289 (22.9)
Time between hospital and ICU admission, days	1.0 [0.2, 2.8]
Invasive mechanical ventilation‡, n (%)	1634 (11.4)
In-hospital mortality or ICU admission, n (%)	5067 (35.3)

†Only deaths or ICU admissions within 30 days following hospital admission were considered linked to Covid-19. ‡All patients requiring invasive mechanical ventilation were admitted in ICU in GPUH's hospitals. Time delays are reported as median [interquartile range].

Table 2. Outcomes of patients included in the study.

Score name	Data from previously published studies			Current study		
	Sample size for validation	Outcome	AUROC [95% CI]	Outcome used for comparison	AUROC [95% CI]	P-value
4C Mortality Score [12]	22361	Death (in-hospital)	0.767 [0.760-0.773]	Death (in-hospital)	0.785 [0.775-0.795]	0.003
ABC-GOALSc [13],*	240	ICU admission	0.770 [0.710-0.830]	ICU admission	0.628 [0.616-0.640]	<0.001
ABCS [14]	188	Death (30 days)	0.838 [0.777-0.899]	Death (in-hospital, 30 days)	0.790 [0.780-0.801]	0.128
A-DROP [12],*	15572	Death (in-hospital)	0.736 [0.728-0.744]	Death (in-hospital)	0.730 [0.718-0.741]	0.415
ANDC [15]	125	Death	0.975 [0.947-1.000]	Death (in-hospital, 30 days)	0.751 [0.741-0.761]	**<0.001**
Bennouar et al. [16]	247	Death (28 days)	0.900 [0.870-0.940]	Death (in-hospital, 28 days)	0.724 [0.713-0.736]	**<0.001**
CHA(2)DS(2)-VASc [17]	864	Death	0.690 [0.650-0.730]	Death (in-hospital)	0.687 [0.677-0.697]	0.887
COPS [18],*	1865	Death (28 days)	0.896 [0.872-0.911]	Death (in-hospital, 28 days)	0.745 [0.734-0.756]	**<0.001**
CORONATION-TR [19],*	37377	Death (30 days)	0.896 [0.890-0.902]	Death (in-hospital, 30 days)	0.769 [0.757-0.780]	**<0.001**
COVID-19 SEIMC [20],*	2126	Death (in-hospital, 30 days)	0.831 [0.806-0.856]	Death (in-hospital, 30 days)	0.752 [0.743-0.762]	**<0.001**
COVID-AID [21],*	265	Death (7 days)	0.851 [0.781-0.921]	Death (in-hospital, 7 days)	0.775 [0.762-0.788]	0.036
COVID-GRAM [22],*	710	Composite: Death, ICU admission, invasive mechanical ventilation	0.880 [0.840-0.930]	Composite: Death (in-hospital), ICU admission, invasive mechanical ventilation	0.700 [0.690-0.711]	<0.001
COVID-NoLab [23]	537	Death (in-hospital)	0.803 [Unknown]	Death (in-hospital)	0.693 [0.683-0.704]	NA
COVID-SimpleLab [23]	295	Death (in-hospital)	0.833 [Unknown]	Death (in-hospital)	0.707 [0.696-0.718]	NA
CURB-65 [12]	15560	Death (in-hospital)	0.720 [0.713-0.728]	Death (in-hospital)	0.724 [0.711-0.736]	0.595
Hachim et al. [24]	289	ICU admission	Unknown [Unknown]	ICU admission	0.514 [0.503-0.526]	NA
Hu et al. [25]	64	Death	0.881 [Unknown]	Death (in-hospital)	0.724 [0.713-0.735]	NA
KPI Score [26]	309	Composite: Death (in-hospital), ICU, invasive mechanical ventilation, NIV, oxygen, steroids, IVIg, ECMO, CRRT, dyspnea, X-ray consolidation	0.888 [0.854-0.922]	Composite: Death (in-hospital), ICU admission, invasive mechanical ventilation	0.597 [0.588-0.606]	**<0.001**
LOW-HARM Score [27],*	400	Death (in-hospital)	0.960 [0.940-0.980]	Death (in-hospital)	0.603 [0.588-0.618]	**<0.001**
Mei et al. (Full) [28],*	276	Death (60 days)	0.970 [0.960-0.980]	Death (in-hospital, 60 days)	0.730 [0.719-0.741]	**<0.001**
Mei et al. (Simple) [28]	276	Death (60 days)	0.880 [0.800-0.960]	Death (in-hospital, 60 days)	0.717 [0.706-0.729]	**<0.001**
NEWS2 [29],*	66	Composite: Death or ICU admission	0.822 [0.690-0.953]	Composite: Death (in-hospital), ICU admission	0.639 [0.626-0.651]	0.006
PLANS [30]	1031	Death (in-hospital)	0.870 [0.850-0.890]	Death (in-hospital)	0.739 [0.729-0.750]	**<0.001**
PREDI-CO [31]	526	Composite: Invasive mechanical ventilation, NIV, oxygen saturation <93% with FiO2 = 1	0.850 [0.810-0.880]	ICU admission, invasive mechanical ventilation	0.646 [0.635-0.657]	**<0.001**

PRESEP [32]	557	Death (60 days)	0.607 [0.555-0.652]	Death (in-hospital, 60 days)	0.586 [0.571-0.600]	0.447
qSOFA [12]	19361	Death (in-hospital)	0.622 [0.615-0.630]	Death (in-hospital)	0.583 [0.566-0.601]	**<0.001**
RISE UP [33]	642	Death (30 days)	0.770 [0.680-0.760]	Death (in-hospital, 30 days)	0.770 [0.759-0.782]	1.000
SIMI [34]	275	Composite: Death, NIV, invasive mechanical ventilation	0.800 [Unknown]	Composite: Death (in-hospital), ICU admission, invasive mechanical ventilation	0.664 [0.655-0.674]	NA
SIRS [35]	175	Death (in-hospital)	0.700 [0.610-0.800]	Death (in-hospital)	0.538 [0.526-0.551]	**<0.001**
STSS [36]	100	Death (30 days)	0.962 [0.903-0.990]	Death (in-hospital, 30 days)	0.697 [0.683-0.712]	**<0.001**
Wang et al. (Clinical) [37]	44	Death	0.830 [0.680-0.930]	Death (in-hospital)	0.729 [0.720-0.738]	0.188
Wang et al. (Laboratory) [37]	44	Death	0.880 [0.750-0.960]	Death (in-hospital)	0.628 [0.616-0.640]	**<0.001**

*Alterations were used to compute these scores. Previously published values used are those from the validation cohorts of the initial studies (external if available, otherwise internal). Z-test was used to compare previously published values and values in our cohort. AUROC: area under the receiver operating characteristic curve; CI: confidence interval; IVIg: intravenous immunoglobulins; NIV: non-invasive ventilation; CRRT: continuous renal replacement therapy; ECMO: extracorporeal membrane oxygenation.

Table 3. Summary of scores included in the study and comparison to previously published data.

Score name	Information needed to compute the score				AUROC [95% CI]		Accuracy to predict in-hospital mortality		
	Patient's characteristics	Medical history	Initial presentation	Biology	In-hospital mortality	In-hospital mortality or ICU admission	Performed as well or better than in the first published validation cohort	Performed equally well in patients <65 years old	Performed equally well in all epidemic waves
4C Mortality Score	Age, sex	Chronic cardiac disease, chronic respiratory disease (excluding asthma), chronic renal disease, mild to severe liver disease, dementia, chronic neurological conditions, connective tissue disease, diabetes mellitus, HIV infection, malignancy	Respiratory rate, oxygen saturation, consciousness	Urea, CRP	0.793† [0.783-0.803]	0.659 [0.649-0.670]	Yes	Yes	Yes
ABCS	Age, sex	COPD	-	CRP, white blood cells, lymphocytes, D-dimer, AST, Troponin I, procalcitonin	0.790† [0.780-0.801]	0.682 [0.672-0.692]	Yes	Yes	Yes
COVID-GRAM*	Age	COPD, hypertension, diabetes, coronary artery disease, chronic kidney disease, cancer, cerebrovascular disease, hepatitis B, immunodeficiency	Abnormalities on chest radiography, haemoptysis, dyspnoea, consciousness	Neutrophils, lymphocytes, LDH, bilirubin	0.771 [0.760-0.783]	0.688 [0.677-0.699]	No	Yes	Yes
RISE UP	Age	-	Heart rate, mean blood pressure, respiratory rate, oxygen saturation, temperature, Glasgow coma scale	Albumin, urea, LDH, bilirubin	0.770 [0.759-0.782]	0.660 [0.650-0.671]	Yes	No	Yes
CORONATION-TR*	Age	Heart failure, diabetes, coronary artery disease, peripheral artery disease, collagen tissue disorders, malignancy, lymphoma, heart failure, COPD, cerebrovascular disease, hypertension, diabetes mellitus, valvular heart disease, chronic liver disease	Pneumonia on chest tomography	Neutrophils, lymphocytes, platelets, D-dimer, LDH, CRP, haemoglobin, creatinine, albumin	0.769 [0.757-0.780]	0.724 [0.714-0.733]	No	Yes	Yes
ANDC	Age	-	-	Neutrophils, lymphocytes, D-dimer, CRP	0.759 [0.748-0.769]	0.642 [0.632-0.652]	No	Yes	Yes
COVID-19 SEIMC*	Age, sex	-	Dyspnoea, oxygen saturation	Neutrophils, lymphocytes, eGFR	0.752 [0.743-0.762]	0.587 [0.578-0.597]	No	No	Yes

Scores are ordered by performance to predict in-hospital mortality.*Alterations were used to compute these scores. †p<0.01 for AUC comparison between these scores and the other scores. AUROC: area under the receiver operating characteristic curve; CI: confidence interval; CRP: C-reactive protein; LDH: lactate dehydrogenase; AST: aspartate transaminase; COPD: chronic obstructive pulmonary disease; eGFR: estimated glomerular filtration rate.

Table 4. Detailed characteristics of scores with an AUROC >0.75 to predict 30-day in-hospital mortality in the analysis using multiple imputed data.

Appendix 1. Collaborators of the the AP-HP/Universities/INSERM COVID-19 research collaboration and AP-HP COVID CDR Initiative.

Name	Affiliation	Contribution
ANCEL Pierre-Yves	APHP Paris University Center	Local CDW coordinator
BAUCHET Alain	APHP Saclay University	Local CDW coordinator
BEEKER Nathanael	APHP Paris University Center	Data scientist
BENOIT Vincent	WIND Department APHP Greater Paris University Hospital	Data engineer
BEY Romain	WIND Department APHP Greater Paris University Hospital	Data engineer, data scientist, regulatory assessment
BOURMAUD Aurélie	APHP Paris University North	Local CDW coordinator
BRÉANT Stéphane	WIND Department APHP Greater Paris University Hospital	Coordination of clinical research informatics
BURGUN Anita	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Medical & scientific coordination
CARRAT Fabrice	APHP Sorbonne University	Local CDW coordinator
CAUCHETEUX Charlotte	Université Paris-Saclay, Inria, CEA	Data integration and analysis
CHAMP Julien	INRIA Sophia-Antipolis – ZENITH team, LIRMM, Montpellier, France	Data integration and analysis
CORMONT Sylvie	WIND Department APHP Greater Paris University Hospital	Data standardisation
DUBIEL Julien	WIND Department APHP Greater Paris University Hospital	Data engineer
DUCLOS Catherine	APHP Paris Seine Saint Denis Universitary Hospital	Local CDW coordinator
ESTEVE Loic	SED/SIERRA, Inria Centre de Paris	Data engineer, data scientist
FRANK Marie	APHP Saclay University	Local CDW coordinator
GARCELON Nicolas	Imagine Institute	Data engineer, data scientist
GRAMFORT Alexandre	Université Paris-Saclay, Inria, CEA	Data engineer, data scientist
GRIFFON Nicolas	"WIND Department APHP Greater Paris University Hospital UMRS1142 INSERM"	Data standardisation
GRISEL Olivier	Université Paris-Saclay, Inria, CEA	Data engineer, data scientist

Name	Affiliation	Contribution
GUILBAUD Martin	WIND Department APHP Greater Paris University Hospital	Data engineer
HASSEN-KHODJA Claire	Direction of the Clinical Research and Innovation, AP-HP	Medical coordination of data-driven research
HEMERY François	APHP Henri Mondor University Hospital	Local CDW coordinator
HILKA Martin	WIND Department APHP Greater Paris University Hospital	Director of Big data platform
JANNOT Anne Sophie	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Biostatistician, local CDW coordonator
LAMBERT Jerome	APHP Paris University North	Local CDW coordinator
LAYESE Richard	APHP Henri Mondor University Hospital	Data scientist
LEBOUTER Léo	WIND Department APHP Greater Paris University Hospital	Data engineer
LEPROVOST Damien	Clevy.io	Data engineer, data scientist
LERNER Ivan	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Data engineer, data scientist
LEVI SALLAH Kankoe	APHP Paris University North	Data scientist
MAIRE Aurélien	WIND Department APHP Greater Paris University Hospital	Data engineer
MAMZER Marie-France	President of the AP-HP IRB	President of the AP-HP IRB
MARTEL Patricia	APHP Saclay University	Data scientist
MENSCH Arthur	ENS, PSL University	Data engineer, data scientist
MOREAU Thomas	Université Paris-Saclay, Inria, CEA	Data engineer, data scientist
NEURAZ Antoine	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Data engineer, data scientist
ORLOVA Nina	WIND Department APHP Greater Paris University Hospital	Data engineer
PARIS Nicolas	WIND Department APHP Greater Paris University Hospital	Data engineer, data scientist
RANCE Bastien	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Data engineer, data scientist

Name	Affiliation	Contribution
RAVERA Hélène	WIND Department APHP Greater Paris University Hospital	Data engineer
ROZES Antoine	APHP Sorbonne University	Data scientist
RUFAT Pierre	APHP Sorbonne University	Local CDW coordinator
SALAMANCA Elisa	WIND Department APHP Greater Paris University Hospital	Director of the Data & Innovation department
SANDRIN Arnaud	WIND Department APHP Greater Paris University Hospital	Director of the National Rare Diseases Database
SERRE Patricia	WIND Department APHP Greater Paris University Hospital	Data engineer, data standardisation
TANNIER Xavier	Sorbonne University	Data engineer, data scientist
TRELUYER Jean-Marc	APHP Paris University Center	Local CDW coordinator
VAN GYSEL Damien	APHP Paris University North	Local CDW coordinator
VAROQUAUX Gael	Université Paris-Saclay, Inria, CEA, Montréal Neurological Institute, McGill University	Data engineer, data scientist
VIE Jill-Jênn	SequeL, Inria Lille	Data engineer, data scientist
WACK Maxime	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Data engineer, data scientist
WAJSBURT Perceval	Sorbonne University	Data engineer, data scientist
WASSERMANN Demian	Université Paris-Saclay, Inria, CEA	Data engineer, data scientist
ZAPLETAL Eric	Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital	Data engineer

Appendix 2. RECORD and TRIPOD checklists

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Title Title NA
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction		
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction		
Methods					
Study Design	4	Present key elements of study design early in the paper	Study Design		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Study Design		

Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Inclusion and exclusion criteria, Data collection Data collection NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Data collection
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Data collection, Supplementary data		
Bias	9	Describe any efforts to address	Statistical analysis,		

		potential sources of bias	discussion		
Study size	10	Explain how the study size was arrived at	NA		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Statistical analysis, Supplementary data		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Statistical analysis		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Study design and setting, Other information Supplementary

					data
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Study design and setting
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Inclusion and exclusion criteria, Figure 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)			Table 1
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report			Table 2

		numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			Table 3, Table 4, Supplementary data
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			Supplementary data (notably Table S9 and S10)
Discussion					
Key results	18	Summarise key results with reference to study objectives	Key results		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Limitations and strengths
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Interpretation and generalisability		

Generalisability	21	Discuss the generalisability (external validity) of the study results	Interpretation and generalisability		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Funding		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Accessibility of protocol, raw data, and programming code

TRIPOD Checklist: Prediction Model Validation

Section/Topic		Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	3
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	4
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	4,5,8
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	Describe eligibility criteria for participants.	4,5
	5c	Give details of treatments received, if relevant.	NA
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	5,6,7
	6b	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Sup.
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	Explain how the study size was arrived at.	NA
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	6, Sup.
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	6,7, Sup
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	6,7
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Sup.
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Fig. 1
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Table 1.
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Table 1. and Sup
Model performance	16	Report performance measures (with CIs) for the prediction model.	Table 3 and 4, Sup
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	10,11
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	11,12
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	11,12
Implications	20	Discuss the potential clinical use of the model and implications for future research.	10,11, 12
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	13
Funding	22	Give the source of funding and the role of the funders for the present study.	13

Score name	Specific for Covid-19	Main outcome	Predictors	Sample size for validation	AUROC [95% CI]	Low risk cut-off value (discriminative performance)	High risk cut-off value (discriminative performance)
4C Mortality Score	Yes	Death (in-hospital)	Age, sex, number of comorbidities, respiratory rate, oxygen saturation, consciousness, urea, CRP	22361	0.770 [0.760 - 0.770]	3 (Se=0.997; NPV=0.998)	15 (PPV=0.615)
ABC-GOALSc	Yes	ICU admission	Gender, SBP, dyspnea, respiratory rate, Charlson index, obesity	240	0.770 [0.710-0.830]	NA	NA
ABCS	Yes	Death (30 days)	Age, hs-CRP, WBC, D-dimer, Sex, COPD, AST, hs-Tni, lymphocyte, procalcitonin	188	0.838 [0.777-0.899]	2%	9%
A-DROP	No	Death (in-hospital)	Age, urea, oxygen saturation, oxygen arterial pressure, confusion, SBP	15572	0.736 [0.728-0.744]	NA	NA
ANDC	Yes	Death	Age, neutrophils, lymphocytes, D-dimers, CRP	125	0.975 [0.947-1.000]	59	101
Bennouar et al.	Yes	Death (28 days)	Age, sodium, urea, CRP, NLR, LDH, albumin	247	0.900 [0.870-0.940]	NA	4 (Se=0.91; Sp=0.70)
CHA(2)DS(2)-VASc	No	Death	Age, gender, hypertension, diabetes, stroke, CAD, heart failure	864	0.690 [0.650-0.730]	NA	NA
COPS	Yes	Death (28 days)	Age, mental disturbance, dyspnea, chronic renal failure, dementia, lymphocyte count	1865	0.896 [0.872-0.911]	2	5
CORONATION-TR	Yes	Death (30 days)	Age, neutrophils, lymphocytes, D-dimer, LDH, CRP, haemoglobin, platelets, creatinine, creatinine, albumin, pneumonia on CT, heart failure, diabetes, coronary artery disease, peripheral artery disease, collagen tissue disorders, malignancy, lymphoma, heart failure, COPD, cerebrovascular disease, hypertension, diabetes mellitus, valvular heart disease, chronic liver disease	37377	0.896 [0.890-0.902]	NA	NA
COVID-19 SEIMC	Yes	Death (in-hospital, 30 days)	Age, oxygen saturation, neutrophil, lymphocytes, eGFR, dyspnea, sex	2126	0.831 [0.806-0.856]	2 (Se=1;Sp=0.081; PPV=0.159;NPV=1)	9 (Se=0.862;Sp=0.685; PPV=0.322;0.966)
COVID-AID	Yes	Death (7 days)	Age, mean arterial pressure, severe hypoxia (oxygen therapy, mechanical ventilation, NIV, oxygen saturation), SCr	265	0.851 [0.781-0.921]	NA	NA
COVID-GRAM	Yes	Composite: Death, ICU admission, mechanical ventilation	Age, number of comorbidities (COPD, hypertension, diabetes, CAD, CKD, cancer, cerebrovascular disease, hepatitis B, immunodeficiency), cancer history, neutrophils, lymphocytes, LDH, bilirubin, chest radiography abnormalities, hemoptysis, dyspnea, unconsciousness	710	0.880 [0.840-0.930]	NA	NA
COVID-NoLab	Yes	Death (in-hospital)	Age, respiratory rate, oxygen saturation	537	0.803 [Unknown]	1	6
COVID-SimpleLab	Yes	Death (in-hospital)	CRP, respiratory rate, oxygen saturation, age, asthma, WBC, creatinine	295	0.833 [Unknown]	7	12
CURB-65	No	Death (in-hospital)	Confusion, urea, respiratory rate, SBP, DBP, age	15560	0.720 [0.713-0.728]	NA	NA
Hachim et al.	Yes	ICU admission	D dimers, urea, troponin	289	NA	1 (Se=0.854;Sp=0.460)	3 (Se=0.302;Sp=0.931)
Hu et al.	Yes	Death	Age, hsCRP, lymphocytes, D-dimers	64	0.881	NA	0 (Se=0.839;Sp=0.794)
KPI Score	Yes	Composite: Death (in-hospital), ICU, MV, NIV, O2, CTC, IvIg, ECMO, CRRT, dyspnea, X-ray consolidation	Age, CRP, PCT, lymphocytes (%), monocytes (%), albumin	309	0.888 [0.854-0.922]	-7 (Se=0.9; NLR=0.225)	15 (Sp=0.9; PLR=5.334)
LOW-HARM Score	Yes	Death (in-hospital)	Hypertension, oxygen saturation, WBC, lymphocytes, SCr, CPK, troponin, myoglobin	400	0.960 [0.940-0.980]	NA	25 (Se=0.915;Sp=0.89; PPV=0.9;NPV=0.91)
Mei et al. (full)	Yes	Death (60 days)	Age, respiratory failure, WBC, lymphocytes, platelets, D-dimer and LDH	276	0.970 [0.960-0.980]	NA	"30% risk" (Se=0.742;Sp=0.972; PPV=0.717;NPV=0.975)

Score name	Specific for Covid-19	Main outcome	Predictors	Sample size for validation	AUROC [95% CI]	Low risk cut-off value (discriminative performance)	High risk cut-off value (discriminative performance)
Mei et al. (simple)	Yes	Death (60 days)	Age, respiratory failure, CAD, renal failure and heart failure	276	0.880 [0.800-0.960]	NA	NA
NEWS2	No	Composite: Death, ICU admission	Respiratory rate, oxygen saturation, systolic blood pressure, heart rate, temperature, oxygen therapy, counsciousness	66	0.822 [0.690-0.953]	NA	6 (Se=0.800;Sp=0.843; PPV=0.60;NPV=0.935)
PLANS	Yes	Death (in-hospital)	Age, sex, neutrophils, lymphocytes, platelets	1031	0.870 [0.850-0.890]	NA	NA
PREDI-CO	Yes	Composite: Mechanical ventilation, NIV, oxygen saturation <93% with FiO2=1	Age, obesity, temperature, respiratory rate, lymphocytes, CRP, LDH	526	0.850 [0,810-0,880]	NA	3 (Se=0.80;Sp=0.76; PPV=0.69;NPV=0.85)
PRESEP	No	Death (60 days)	Temperature, oxygen saturation, respiratory rate, heart rate, systolic blood pressure, glasgow coma scale	557	0.607 [0.555-0.652]	NA	1 (Se=0.6226;Sp=0.5655; PPV=0.175;NPV=0.91)
qSOFA	No	Death (in-hospital)	Respiratory rate, Glasgow coma scale, systolic blood pressure	19361	0.622 [0.615-0.630]	NA	NA
RISE UP	No	Death (30 days)	Age, heart rate, MBP, respiratory rate, oxygen saturation, temperature, Glasgow coma scale, albumin, urea, LDH, bilirubin	642	0.770 [0.680-0.760]	0.05 (Se=1;Sp=0.089; PPV=0.278;NPV=1)	0.5 (Se=0.217;Sp=0.915; PPV=0.473;NPV=0.770)
SIMI	Yes	Composite: NIV, mechanical ventilation, death	Age, coronary heart disease, CRP, AST, D-dimer, neutrophils, lymphocytes	175	0.800 [Unknown]	NA	7 (Se=0.93;Sp=0.34; PPV=0.59;NPV=0.82)
SIRS	No	Death (in-hospital)	Temperature, heart rate, respiratory rate, WBC	175	0.700 [0.610-0.800]	NA	2 (Se=76%;Sp=52%; PPV=32%;NPV=90%)
STSS	No	Death (30 days)	Respiratory rate, heart rate, SBP, oxygen saturation, Glasgow coma scale, age	100	0.962 [0.903-0.990]	NA	1 (Se=0.833;Sp=0.936; PPV=0.455;NPV=0.989)
Wang et al. (Clinical)	Yes	Death	Age, hypertension, CAD	44	0.830 [0.680-0.930]	NA	-1.798 (Se=0.643;Sp=0.933; PPV=0.818;NPV=0.849)
Wang et al. (Laboratory)	Yes	Death	Age, lymphocytes, hsCRP, D-dimer, AST, eGFR	44	0.880 [0.750-0.960]	NA	-3.829 (Se=1.00;Sp=0.70; PPV=0.609;NPV=1.00)

Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; PLR: positive likelihood ratio; NLR: negative likelihood ratio.

Table S1. General information on scores included in the study.

Score name	Unavailable variables?	Variable with the highest rate of missing data	Can the score be computed without alterations?	Alterations used to compute the score?	Sample size with complete data
4C Mortality Score	No	Glasgow coma scale	Yes	NA	3277
ABC-GOALSc	Yes (dyspnea)	NA	No	Dyspnea: defined as RR > 24/min and/or oxygen saturation < 92%	3784
ABCS	No	Troponin	Yes	NA	2411
A-DROP	Yes (oxygen arterial pressure)	NA	No	Oxygen arterial pressure: ignored, respiratory failure is defined using arterial oxygen saturation	3974
ANDC	No	D-dimers	Yes	NA	7137
Bennouar et al.	No	Albumin	Yes	NA	3395
CHA(2)DS(2)-VASc	No	None	Yes	NA	14343
COPS	Yes (dyspnea)	NA	No	Dyspnea: defined as RR > 24/min and/or oxygen saturation < 92%	4882
CORONATION-TR	Yes (pneumonia on CT)	NA	No	Pneumonia on CT: considered true for patients with ICD-10 codes for respiratory Covid-19, otherwise false	2572
COVID-19 SEIMC	Yes (dyspnea)	NA	No	Dyspnea: defined as RR > 24/min and/or oxygen saturation < 92%	7079
COVID-AID	Yes (oxygen therapy, mechanical or non-invasive ventilation)	NA	No	Oxygen therapy, mechanical or non-invasive ventilation: ignored, severe hypoxia is defined as oxygen saturation < 90%	6565
COVID-GRAM	Yes (chest radiography abnormalities, hemoptysis, direct bilirubin)	NA	No	Dyspnea: defined as RR > 24/min and/or oxygen saturation < 92% Chest radiography abnormalities: considered true for patients with ICD-10 codes for respiratory Covid-19, otherwise false Hemoptysis: ignored, rare event Direct bilirubin: estimated as 0.6 x total bilirubin	2667
COVID-NoLab	No	Respiratory rate	Yes	NA	8109
COVID-SimpleLab	No	Respiratory rate	Yes	NA	6640
CURB-65	No	Glasgow coma scale	Yes	NA	5300
Hachim et al.	No	D-dimers	Yes	NA	4920
Hu et al.	No	D-dimers	Yes	NA	7137
KPI Score	No	Albumin	Yes	NA	4703
LOW-HARM Score	Yes (myoglobin)	NA	No	Myoglobin: ignored, cardiac injury is defined as either CPK or troponin elevation	1957
Mei et al. (full)	Yes (respiratory failure)	NA	No	Respiratory failure: defined as RR ≥ 30/min and/or oxygen saturation < 90%	3071
Mei et al. (simple)	No	None	Yes	NA	8123
NEWS2	Yes (oxygen therapy)	NA	No	Oxygen therapy: considered true for patients with ICD-10 codes for respiratory Covid-19, otherwise false	3704
PLANS	No	Lymphocytes	Yes	NA	12307
PREDI-CO	No	LDH	Yes	NA	2898
PRESEP	No	Glasgow coma scale	Yes	NA	3704
qSOFA	No	Glasgow coma scale	Yes	NA	3718
RISE UP	No	Albumin	Yes	NA	1015

Score name	Unavailable variables?	Variable with the highest rate of missing data	Can the score be computed without alterations?	Alterations used to compute the score?	Sample size with complete data
SIMI	No	D-dimer	Yes	NA	6230
SIRS	No	Respiratory rate	Yes	NA	7688
STSS	No	Glasgow coma scale	Yes	NA	3707
Wang et al. (Clinical)	No	None	Yes	NA	14343
Wang et al. (Laboratory)	No	D-dimers	Yes	NA	4266

Table S2. Systematic evaluation for scores included in the study.

Variable	Cut-offs for aberrant or extreme values	Way to treat out-of-range values
Vital signs on admission		
Heart rate, beats per minute	NA	Out of range values were ignored (e.g. for a diastolic blood pressure of 7, the patient was considered as having a missing value for this variable; no control for user input exists in most electronic medical records used in GPUH's hospitals)
Respiratory rate, cycles per minute	8-80	
Altered consciousness (i.e. Glasgow Coma Scale < 15)	NA	
Diastolic blood pressure, mmHg	10-	
Mean blood pressure, mmHg	10-	
Systolic blood pressure, mmHg	40-250	
Arterial oxygen saturation, %	50-100	
Temperature, °C	30-42	
Body mass index (BMI), kg/m²	NA	
Biological values on admission		
Haemoglobin, g/dl	NA	Out-of-range values were modified to the closest in-range value (e.g. for a lactate dehydrogenase value of >1200 UI/l given by the laboratory, the patient was considered as having a value of 1200 UI/l)
Leukocytes, G/l	0-60	
Neutrophils, G/l	0-60	
Lymphocytes, G/l	0-40	
Eosinophils, G/l	NA	
Monocytes, G/l	0-10	
Basophils, G/l	NA	
Platelets count, G/l	0-2000	
Sodium, mmol/l	110-170	
Potassium, mmol/l	NA	
Bicarbonates, mmol/l	NA	
Proteins, g/l	25-	
Calcium, mmol/l	0-4	
Urea, mmol/l	0-100	
Serum creatinine, µmol/l	4.4-4000	
Alanine aminotransferase, IU/l	3-500	
Asparate aminotransferase, IU/l	3-1000	
Total bilirubin, µmol/l	0-500	
Lactate dehydrogenase, IU/l	0-1200	
Creatinine phosphokinase, IU/l	0-10000	
Troponine, ng/l	2.3-5000	
Activated partial thromboplastin time	0-7	
Prothrombin time, %	10-100	
Fibrinogen, g/l	NA	
D-dimers, µg/l	270-10000	
C-reactive protein, mg/l	0.2-4800	
Procalcitonin, µg/l	0-25	

Table S3. Lower and upper limits for aberrant or extreme values.

Variable	Variable class (transformation used)	Missing data, n (%)
Demographic data		
Sex	Binary	
Age	Continuous	
Department of residence	Factor (departments outside Paris region were regrouped)	
Diagnosis of Covid-19		
Admission during « first wave »	Binary	
Time between PCR sample and admission	Continuous (logarithmic transformation)	
Initial care site	Factor	
Medical history		
ICD-10 codes available for previous visits	Binary	
Modified Charlson comorbidity index	Ordered factor (classes: 0, 1, (1-2], (2-3], (3-5], (5-23])	
Cardiac disease		
Congestive heart failure		
Myocardial infarction		
Valvular heart disease		
Peripheral vascular disease		
Cerebrovascular disease		
Ischemic stroke		
Dementia		
Hemiplegia		0 (0)
Arterial hypertension		
Diabetes		
Diabetes with end-organ damage		
Chronic pulmonary disease		
Asthma	Binary	
Chronic obstructive pulmonary disease		
Moderate or severe renal disease		
Mild liver disease		
Moderate or severe liver disease		
Any tumor		
Metastatic solid tumor		
Lymphoma		
Connective tissue disease		
Ulcer disease		
HIV infection		
Obesity (ICD-10 codes only)		
Extreme obesity (ICD-10 codes only)		
Vital signs on admission		
Heart rate, beats per minute	Continuous	3344 (23.3)
Respiratory rate, cycles per minute	Continuous	5615 (39.1)
Altered consciousness (i.e. Glasgow Coma Scale < 15)	Binary	8581 (59.8)
Diastolic blood pressure, mmHg	Continuous	5980 (41.7)
Mean blood pressure, mmHg	Continuous	6477 (45.2)
Systolic blood pressure, mmHg	Continuous	5976 (41.7)
Pulse saturometry, %	Continuous (square root transformation)	4551 (31.7)
Temperature, °C	Continuous	3374 (23.5)
Body mass index (BMI), kg/m²	Continuous	5435 (37.9)
Biological values on admission		
Haemoglobin, g/dl	Continuous	1759 (12.3)
Leukocytes, G/l	Continuous	1762 (12.3)
Neutrophils, G/l	Continuous	1990 (13.9)
Lymphocytes, G/l	Continuous (logarithmic transformation)	2020 (14.1)
Eosinophils, G/l	Continuous (logarithmic transformation)	2026 (14.1)
Monocytes, G/l	Continuous (logarithmic transformation)	2019 (14.1)
Basophils, G/l	Continuous (logarithmic transformation)	2027 (14.1)
Platelets count, G/l	Continuous	1769 (12.3)
Sodium, mmol/l	Continuous	599 (4.2)
Potassium, mmol/l	Continuous	848 (5.9)
Bicarbonates, mmol/l	Continuous	6557 (45.7)
Proteins, g/l	Continuous	982 (6.8)
Calcium, mmol/l	Continuous	5842 (40.7)
Urea, mmol/l	Continuous (logarithmic transformation)	831 (5.8)
Serum creatinine, µmol/l	Continuous (logarithmic transformation)	560 (3.9)
Alanine aminotransferase, IU/l	Continuous (logarithmic transformation)	2477 (17.3)
Asparate aminotransferase, IU/l	Continuous (logarithmic transformation)	2926 (20.4)
Total bilirubin, µmol/l	Continuous (logarithmic transformation)	2427 (16.9)
Lactate dehydrogenase, IU/l	Continuous (logarithmic transformation)	6961 (48.5)
Creatinine phosphokinase, IU/l	Continuous (logarithmic transformation)	6670 (46.5)
Troponin, ng/l	Continuous (logarithmic transformation)	7432 (51.8)
Activated partial thromboplastin time	Continuous	3160 (22)
Prothrombin time, %	Continuous (logarithmic transformation)	2773 (19.3)
Fibrinogen, g/l	Continuous	5200 (36.3)
D-dimers, µg/l	Continuous (logarithmic transformation)	6205 (43.3)
C-reactive protein, mg/l	Continuous (logarithmic transformation)	1365 (9.5)
Procalcitonin, µg/l	Continuous (logarithmic transformation)	7236 (50.4)
Albumin, g/l	Continuous	9451 (65.9)
Outcomes		
Death	Binary	
ICU admission	Binary	0 (0)
Mechanical ventilation	Binary	

Table S4. Summary of variables used for multiple imputations.

Variable	First wave of admission (n = 6142)		Subsequent waves of admission (n = 8201)		All patients (n = 14343)	
Demographic data	*Missing*		*Missing*		*Missing*	
Female sex, n (%)		2553 (41.6)		3636 (44.3)		6189 (43.1)
Age, years		68.9 (SD 17.1)		68 (SD 17.8)		68.4 (SD 17.5)
Diagnosis of Covid-19						
Time between PCR and admission, days		0 [-0.1, 0]		-0.1 [-0.2, 0]		-0.1 [-0.1, 0]
Medical history, n (%)						
Modified Charlson comorbidity index, pts		1 [0, 2]		0 [0, 2]		1 [0, 2]
Congestive heart failure		817 (13.3)		1048 (12.8)		1865 (13)
Myocardial infarction		419 (6.8)		544 (6.6)		963 (6.7)
Peripheral vascular disease		388 (6.3)		496 (6)		884 (6.2)
Cerebrovascular disease		626 (10.2)		735 (9)		1361 (9.5)
Hemiplegia		297 (4.8)		302 (3.7)		599 (4.2)
Dementia		996 (16.2)		1006 (12.3)		2002 (14)
Arterial hypertension		2681 (43.7)		3445 (42)		6126 (42.7)
Diabetes		1455 (23.7)		1960 (23.9)		3415 (23.8)
Diabetes with end-organ damage		839 (13.7)		1183 (14.4)		2022 (14.1)
Chronic pulmonary disease		733 (11.9)		1030 (12.6)		1763 (12.3)
Moderate or severe renal disease		976 (15.9)		1220 (14.9)		2196 (15.3)
Moderate or severe liver disease		66 (1.1)		94 (1.1)		160 (1.1)
Any tumor		624 (10.2)		920 (11.2)		1544 (10.8)
Metastatic solid tumor		145 (2.4)		266 (3.2)		411 (2.9)
Connective tissue disease		93 (1.5)		212 (2.6)		305 (2.1)
HIV infection		114 (1.9)		124 (1.5)		238 (1.7)
Obesity (ICD-10 codes only)		1067 (17.4)		1648 (20.1)		2715 (18.9)
Vital signs on admission						
Heart rate, beats per minute	*1486 (24.2)*	88.9 (SD 17.9)	*1858 (22.7)*	88.2 (SD 17.6)	*3344 (23.3)*	88.5 (SD 17.7)
Respiratory rate, cycles per minute	*2466 (40.1)*	25.5 (SD 7.7)	*3149 (38.4)*	24.4 (SD 7.4)	*5615 (39.1)*	24.9 (SD 7.5)
Altered consciousness (i.e. GCS < 15), n (%)	*4095 (66.7)*	110 (5.4)	*4486 (54.7)*	135 (3.6)	*8581 (59.8)*	245 (4.3)
Diastolic blood pressure, mmHg	*2538 (41.3)*	75 (SD 15.1)	*3442 (42)*	74.6 (SD 15)	*5980 (41.7)*	74.8 (SD 15.1)
Mean blood pressure, mmHg	*3127 (50.9)*	94.2 (SD 15.7)	*3350 (40.8)*	93.7 (SD 15.6)	*6477 (45.2)*	93.9 (SD 15.6)
Systolic blood pressure, mmHg	*2545 (41.4)*	131.6 (SD 21.9)	*3431 (41.8)*	130.9 (SD 22)	*5976 (41.7)*	131.2 (SD 22)
Pulse saturometry, %	*1930 (31.4)*	96 [93, 98]	*2621 (32)*	95 [92, 97]	*4551 (31.7)*	96 [93, 98]
Temperature, °C	*1483 (24.1)*	37.5 (SD 1)	*1891 (23.1)*	37.4 (SD 0.9)	*3374 (23.5)*	37.5 (SD 1)
Body mass index (BMI), kg/m²	*2638 (43)*	27 (SD 6.5)	*2797 (34.1)*	27.1 (SD 6.5)	*5435 (37.9)*	27.1 (SD 6.5)
Biological values on admission						
Haemoglobin, g/dl	*765 (12.5)*	13.1 (SD 2)	*994 (12.1)*	13 (SD 2)	*1759 (12.3)*	13 (SD 2)
Leukocytes, G/l	*767 (12.5)*	7.3 (SD 4)	*995 (12.1)*	7.1 (SD 4)	*1762 (12.3)*	7.2 (SD 4)
Neutrophils, G/l	*873 (14.2)*	5.6 (SD 3.4)	*1117 (13.6)*	5.4 (SD 3.3)	*1990 (13.9)*	5.5 (SD 3.4)
Lymphocytes, G/l	*883 (14.4)*	1 [0.7, 1.3]	*1137 (13.9)*	0.9 [0.7, 1.3]	*2020 (14.1)*	0.9 [0.7, 1.3]
Platelets count, G/l	*773 (12.6)*	219.7 (SD 95)	*996 (12.1)*	219.7 (SD 92.1)	*1769 (12.3)*	219.7 (SD 93.4)
Sodium, mmol/l	*309 (5)*	136.4 (SD 5)	*290 (3.5)*	135.8 (SD 4.4)	*599 (4.2)*	136 (SD 4.7)
Potassium, mmol/l	*415 (6.8)*	4.1 (SD 0.6)	*433 (5.3)*	4.1 (SD 0.6)	*848 (5.9)*	4.1 (SD 0.6)
Bicarbonates, mmol/l	*3000 (48.8)*	23.8 (SD 3.9)	*3557 (43.4)*	24.4 (SD 3.9)	*6557 (45.7)*	24.2 (SD 3.9)
Proteins, g/l	*532 (8.7)*	71.8 (SD 7.4)	*450 (5.5)*	71.2 (SD 7.3)	*982 (6.8)*	71.5 (SD 7.3)
Urea, mmol/l	*398 (6.5)*	6.4 [4.5, 10.4]	*433 (5.3)*	6.5 [4.6, 9.7]	*831 (5.8)*	6.5 [4.6, 9.9]
Serum creatinine, µmol/l	*269 (4.4)*	82 [66, 111]	*291 (3.5)*	83 [65.5, 109]	*560 (3.9)*	82.4 [66, 110]
Alanine aminotransferase, IU/l	*1155 (18.8)*	29.5 [20, 48]	*1322 (16.1)*	29.5 [19.2, 46.8]	*2477 (17.3)*	29.5 [20, 47]
Asparate aminotransferase, IU/l	*1283 (20.9)*	43 [30, 64.4]	*1643 (20)*	41.5 [29, 62]	*2926 (20.4)*	42 [29.2, 63]
Total bilirubin, µmol/l	*1136 (18.5)*	8 [6, 12]	*1291 (15.7)*	8 [6, 12]	*2427 (16.9)*	8 [6, 12]
Lactate dehydrogenase, IU/l	*2694 (43.9)*	366 [277, 503]	*4267 (52)*	359 [273, 494]	*6961 (48.5)*	362 [275, 499]
Creatinine phosphokinase, IU/l	*2673 (43.5)*	136 [69.6, 325]	*3997 (48.7)*	127 [65, 282]	*6670 (46.5)*	132 [67, 300]
Troponine, ng/l	*2901 (47.2)*	15 [10, 31.4]	*4531 (55.2)*	15 [9.5, 31]	*7432 (51.8)*	15 [10, 31]
Activated partial thromboplastin time	*1611 (26.2)*	1.2 (SD 0.3)	*1549 (18.9)*	1.2 (SD 0.3)	*3160 (22)*	1.2 (SD 0.3)
Prothrombin time, %	*1453 (23.7)*	86 [75, 96]	*1320 (16.1)*	87 [75, 98]	*2773 (19.3)*	87 [75, 97]
Fibrinogen, g/l	*2614 (42.6)*	5.9 (SD 1.6)	*2586 (31.5)*	5.7 (SD 1.6)	*5200 (36.3)*	5.8 (SD 1.6)
D-dimers, µg/l	*3616 (58.9)*	1014 [593, 1780]	*2589 (31.6)*	950 [580, 1661]	*6205 (43.3)*	964 [585, 1690]
C-reactive protein, mg/l	*638 (10.4)*	77 [34.2, 137.1]	*727 (8.9)*	65.9 [27, 121]	*1365 (9.5)*	70 [30, 129]
Procalcitonin, µg/l	*3057 (49.8)*	0.2 [0.1, 0.4]	*4179 (51)*	0.2 [0.1, 0.4]	*7236 (50.4)*	0.2 [0.1, 0.4]
Albumin, g/l	*4235 (69)*	32.1 (SD 5.7)	*5216 (63.6)*	32.5 (SD 5.4)	*9451 (65.9)*	32.4 (SD 5.5)
Outcomes						
Death		1279 (20.8)		1304 (15.9)		2583 (18)
ICU admission		1326 (21.6)		1963 (23.9)		3289 (22.9)
Mechanical ventilation		695 (11.3)		939 (11.4)		1634 (11.4)

Continuous variables are reported as mean (standard deviation (SD)) for normally distributed variables and median [interquartile range] for non-normally distributed variables. GCS: Glasgow Coma Scale.

Table S5. Baseline characteristics and outcomes according to wave of admission.

	Centre 1 (n=1742)	Centre 2 (n=1538)	Centre 3 (n=1283)	Centre 4 (n=1129)	Centre 5 (n=1123)	Centre 6 (n=1076)	Centre 7 (n=957)	Centre 8 (n=711)	Centre 9 (n=636)	Centre 10 (n=613)	Centre 11 (n=605)	Centre 12 (n=563)	Centre 13 (n=361)	Centre 14 (n=323)	Centre 15 (n=243)	Centre 16 (n=160)	Centres 17-28† (n=1203)	Centres 29-33‡ (n=77)
Demographic data																		
Female sex, n (%)	675 (38.7)	660 (42.9)	539 (42)	442 (39.1)	436 (38.8)	444 (41.3)	388 (40.5)	306 (43)	276 (43.4)	288 (47)	269 (44.5)	214 (38)	148 (41)	120 (37.2)	101 (41.6)	62 (38.8)	776 (64.5)	45 (58.4)
Age, years	65.7 (SD 16.9)	66.3 (SD 18)	65.8 (SD 16.6)	67.4 (SD 17)	65.2 (SD 16.5)	68.3 (SD 16.2)	69.1 (SD 16.3)	68.6 (SD 17.5)	64.1 (SD 17.5)	69.2 (SD 19.2)	73 (SD 17.3)	67.3 (SD 16.6)	66.2 (SD 17.3)	63.3 (SD 16.5)	64.9 (SD 17.8)	61.1 (SD 15.7)	85.7 (SD 8.4)	49.3 (SD 20.2)
Diagnosis of Covid-19																		
Time between PCR and admission, days	-0.1 [-0.1,0]	-0.1 [-0.2,0]	-0.1 [-0.2,0]	-0.1 [-0.1,0]	-0.1 [-0.2,0]	-0.1 [-0.2,-0.1]	-0.1 [-0.2,0]	-0.1 [-0.2,0]	-0.1 [-0.1,0]	-0.1 [-0.2,0]	0 [-0.1,0]	-0.1 [-0.1,0]	-0.1 [-0.1,0]	-0.1 [-0.1,0]	-0.1 [-0.2,0]	-0.1 [-0.2,0]	1.1 [0.5,1.7]	0 [-0.1,0.3]
Medical history																		
ICD-10 codes available for previous visits, n (%)	791 (45.4)	800 (52)	638 (49.7)	575 (50.9)	535 (47.6)	586 (54.5)	517 (54)	334 (47)	291 (45.8)	280 (45.7)	338 (55.9)	287 (51)	89 (24.7)	120 (37.2)	110 (45.3)	43 (26.9)	1074 (89.3)	51 (66.2)
Modified Charlson comorbidity index, pts	0 [0,2]	0 [0,2]	0 [0,2]	1 [0,2]	0 [0,2]	1 [0,2]	0 [0,2]	1 [0,2]	0 [0,1]	0 [0,2]	1 [0,2]	0 [0,2]	0 [0,2]	0 [0,1.5]	0 [0,1]	0 [0,2]	3 [2,5]	1 [0,2]
Missing data, n (%)																		
Altered consciousness	1736 (99.7)	1535 (99.8)	1274 (99.3)	655 (58)	269 (24)	103 (9.6)	110 (11.5)	475 (66.8)	43 (6.8)	36 (5.9)	110 (18.2)	90 (16)	359 (99.4)	322 (99.7)	37 (15.2)	157 (98.1)	1195 (99.3)	75 (97.4)
Systolic blood pressure	1178 (67.6)	601 (39.1)	574 (44.7)	490 (43.4)	287 (25.6)	284 (26.4)	282 (29.5)	420 (59.1)	184 (28.9)	181 (29.5)	263 (43.5)	150 (26.6)	242 (67)	90 (27.9)	45 (18.5)	60 (37.5)	593 (49.3)	52 (67.5)
Mean blood pressure	1189 (68.3)	1348 (87.6)	1006 (78.4)	449 (39.8)	192 (17.1)	109 (10.1)	90 (9.4)	302 (42.5)	26 (4.1)	38 (6.2)	87 (14.4)	39 (6.9)	257 (71.2)	268 (83)	18 (7.4)	113 (70.6)	895 (74.4)	51 (66.2)
Arterial oxygen saturation	1710 (98.2)	418 (27.2)	775 (60.4)	117 (10.4)	15 (1.3)	16 (1.5)	14 (1.5)	467 (65.7)	13 (2)	7 (1.1)	30 (5)	16 (2.8)	331 (91.7)	38 (11.8)	7 (2.9)	13 (8.1)	494 (41.1)	70 (90.9)
Body mass index	741 (42.5)	672 (43.7)	500 (39)	402 (35.6)	318 (28.3)	410 (38.1)	277 (28.9)	252 (35.4)	310 (48.7)	270 (44)	189 (31.2)	287 (51)	140 (38.8)	178 (55.1)	95 (39.1)	50 (31.2)	309 (25.7)	35 (45.5)
Blood urea nitrogen	24 (1.4)	53 (3.4)	98 (7.6)	18 (1.6)	9 (0.8)	16 (1.5)	8 (0.8)	51 (7.2)	10 (1.6)	14 (2.3)	6 (1)	16 (2.8)	7 (1.9)	7 (2.2)	2 (0.8)	153 (95.6)	319 (26.5)	20 (26)
Sodium	23 (1.3)	53 (3.4)	85 (6.6)	18 (1.6)	5 (0.4)	65 (6)	8 (0.8)	15 (2.1)	9 (1.4)	6 (1)	6 (1)	6 (1.1)	7 (1.9)	5 (1.5)	1 (0.4)	4 (2.5)	267 (22.2)	16 (20.8)
Haemoglobin	22 (1.3)	21 (1.4)	49 (3.8)	1121 (99.3)	7 (0.6)	13 (1.2)	6 (0.6)	11 (1.5)	12 (1.9)	4 (0.7)	7 (1.2)	5 (0.9)	8 (2.2)	5 (1.5)	1 (0.4)	5 (3.1)	450 (37.4)	12 (15.6)
Lymphocytes count	22 (1.3)	33 (2.1)	81 (6.3)	1121 (99.3)	9 (0.8)	13 (1.2)	38 (4)	41 (5.8)	13 (2)	35 (5.7)	9 (1.5)	5 (0.9)	8 (2.2)	66 (20.4)	2 (0.8)	11 (6.9)	497 (41.3)	16 (20.8)
C-reactive protein	104 (6)	54 (3.5)	151 (11.8)	96 (8.5)	26 (2.3)	70 (6.5)	30 (3.1)	41 (5.8)	24 (3.8)	277 (45.2)	30 (5)	36 (6.4)	15 (4.2)	10 (3.1)	5 (2.1)	85 (53.1)	291 (24.2)	20 (26)
D-dimers	725 (41.6)	442 (28.7)	550 (42.9)	528 (46.8)	279 (24.8)	443 (41.2)	508 (53.1)	294 (41.4)	167 (26.3)	287 (46.8)	348 (57.5)	245 (43.5)	63 (17.5)	129 (39.9)	49 (20.2)	60 (37.5)	1034 (86)	54 (70.1)
Outcomes																		
In-hospital death, n (%)	286 (16.4)	284 (18.5)	224 (17.5)	233 (20.6)	201 (17.9)	202 (18.8)	153 (16)	110 (15.5)	92 (14.5)	120 (19.6)	118 (19.5)	97 (17.2)	25 (6.9)	62 (19.2)	48 (19.8)	20 (12.5)	302 (25.1)	6 (7.8)
ICU admission, n (%)	408 (23.4)	434 (28.2)	399 (31.1)	268 (23.7)	262 (23.3)	226 (21)	235 (24.6)	220 (30.9)	132 (20.8)	113 (18.4)	166 (27.4)	133 (23.6)	98 (27.1)	75 (23.2)	44 (18.1)	44 (27.5)	9 (0.7)	23 (29.9)

† : hospitals with a predominant activity in geriatric medicine or in physical medicine and rehabilitation. ‡ : other hospitals. Continuous variables are reported as mean (standard deviation (SD)) for normally distributed variables and median [interquartile range] for non-normally distributed variables.

Table S6. Baseline characteristics, rate of missing data and outcomes according to initial care site.

Score name	AUROC [95% CI]			
	In-hospital mortality within 30 days		In-hospital mortality or ICU admission within 30 days	
	Principal analysis: Multiple imputed datasets	Sensitivity analysis: Complete dataset†	Principal analysis: Multiple imputed datasets	Sensitivity analysis: Complete dataset†
4C Mortality Score	0.793 [0.783-0.803]	0.784 [0.763-0.804]	0.659 [0.649-0.670]	0.650 [0.631-0.669]
ABCS	0.790 [0.780-0.801]	0.765 [0.742-0.788]	0.682 [0.672-0.692]	0.642 [0.620-0.664]
COVID-GRAM*	0.771 [0.760-0.783]	0.777 [0.756-0.799]	0.688 [0.677-0.699]	0.696 [0.676-0.716]
RISE UP	0.770 [0.759-0.782]	0.750 [0.712-0.788]	0.660 [0.650-0.671]	0.629 [0.593-0.664]
CORONATION-TR*	0.769 [0.757-0.780]	0.740 [0.717-0.764]	0.724 [0.714-0.733]	0.687 [0.666-0.707]
ANDC	0.759 [0.748-0.769]	0.751 [0.736-0.765]	0.642 [0.632-0.652]	0.627 [0.614-0.640]
COVID-19 SEIMC*	0.752 [0.743-0.762]	0.764 [0.751-0.777]	0.587 [0.578-0.597]	0.611 [0.598-0.624]
COVID-AID*	0.747 [0.737-0.757]	0.766 [0.752-0.780]	0.566 [0.557-0.576]	0.600 [0.586-0.615]
COPS*	0.745 [0.734-0.755]	0.757 [0.741-0.773]	0.611 [0.599-0.622]	0.637 [0.622-0.653]
PLANS	0.745 [0.734-0.757]	0.745 [0.734-0.756]	0.635 [0.625-0.646]	0.630 [0.620-0.640]
Mei et al. (Full)*	0.737 [0.726-0.749]	0.731 [0.708-0.755]	0.684 [0.674-0.694]	0.694 [0.675-0.714]
A-DROP*	0.737 [0.725-0.749]	0.768 [0.750-0.786]	0.601 [0.589-0.614]	0.648 [0.630-0.665]
Hu et al.	0.733 [0.722-0.744]	0.716 [0.700-0.732]	0.656 [0.646-0.666]	0.635 [0.622-0.648]
Hachim et al.	0.732 [0.721-0.743]	0.730 [0.713-0.746]	0.622 [0.612-0.633]	0.608 [0.593-0.623]
SIMI	0.731 [0.720-0.742]	0.715 [0.698-0.732]	0.675 [0.666-0.685]	0.649 [0.636-0.663]
CURB-65	0.731 [0.718-0.743]	0.744 [0.728-0.759]	0.608 [0.596-0.620]	0.626 [0.611-0.642]
Wang et al. (Clinical)	0.726 [0.717-0.736]	0.726 [0.717-0.736]	0.550 [0.540-0.560]	0.550 [0.540-0.560]
Bennouar et al.	0.725 [0.714-0.736]	0.704 [0.683-0.725]	0.694 [0.685-0.704]	0.673 [0.656-0.691]
Mei et al. (Simple)	0.724 [0.712-0.735]	0.729 [0.716-0.742]	0.639 [0.628-0.650]	0.665 [0.652-0.677]
COVID-SimpleLab	0.721 [0.710-0.732]	0.716 [0.701-0.731]	0.674 [0.665-0.684]	0.671 [0.657-0.685]
COVID-NoLab	0.703 [0.692-0.715]	0.699 [0.686-0.712]	0.637 [0.627-0.647]	0.651 [0.639-0.663]
STSS	0.697 [0.683-0.712]	0.712 [0.693-0.731]	0.607 [0.593-0.621]	0.649 [0.632-0.667]
PREDI-CO	0.696 [0.684-0.708]	0.706 [0.681-0.730]	0.703 [0.694-0.712]	0.707 [0.689-0.726]
CHA(2)DS(2)-VASc	0.684 [0.673-0.694]	0.684 [0.673-0.694]	0.551 [0.542-0.561]	0.551 [0.542-0.561]
Wang et al. (Laboratory)	0.646 [0.633-0.659]	0.621 [0.598-0.644]	0.669 [0.659-0.679]	0.656 [0.639-0.673]
ABC-GOALSc*	0.646 [0.633-0.659]	0.670 [0.647-0.692]	0.656 [0.646-0.667]	0.667 [0.650-0.685]
NEWS2*	0.634 [0.618-0.651]	0.626 [0.603-0.649]	0.655 [0.641-0.668]	0.646 [0.627-0.664]
LOW-HARM Score*	0.614 [0.598-0.629]	0.628 [0.594-0.662]	0.549 [0.537-0.561]	0.567 [0.540-0.594]
PRESEP	0.595 [0.580-0.610]	0.588 [0.565-0.611]	0.626 [0.613-0.638]	0.616 [0.598-0.635]
qSOFA	0.594 [0.577-0.611]	0.598 [0.578-0.619]	0.577 [0.562-0.591]	0.588 [0.572-0.605]
KPI Score	0.586 [0.575-0.597]	0.586 [0.569-0.604]	0.614 [0.605-0.623]	0.616 [0.602-0.630]
SIRS	0.549 [0.535-0.562]	0.542 [0.526-0.558]	0.590 [0.580-0.601]	0.586 [0.574-0.599]

†i.e., considering only patients with all variables available to compute a given score (see Table S2 for sample sizes for each score). *alterations were used to compute these scores. AUROC: area under the receiver operating characteristic curve; CI: confidence interval.

Table S7. Discriminative performance of scores included in the study, ordered by performance to predict in-hospital mortality.

Score name	Low-risk cut-off value			High-risk cut-off value		
	Cut-off value	Sensitivity [95% CI]	Specificity [95% CI]	Cut-off value	Sensitivity [95% CI]	Specificity [95% CI]
4C Mortality Score	3	0.998 [0.996-1.000]	0.084 [0.077-0.092]	15	0.215 [0.196-0.234]	0.968 [0.964-0.972]
ABCS	137*	0.992 [0.988-0.996]	0.112 [0.105-0.119]	212*	0.882 [0.867-0.897]	0.512 [0.496-0.527]
RISE UP	0.05	0.998 [0.995-1.000]	0.068 [0.062-0.075]	0.5	0.508 [0.482-0.534]	0.840 [0.831-0.849]
ANDC	59	0.980 [0.974-0.986]	0.188 [0.180-0.197]	101	0.634 [0.611-0.657]	0.734 [0.719-0.749]
COVID-19 SEIMC	2	0.995 [0.991-0.998]	0.109 [0.102-0.115]	9	0.780 [0.763-0.797]	0.610 [0.601-0.619]

*Correspond to “2% mortality risk” and “9% mortality risk” in previously published study, respectively. CI: confidence interval.

Table S8. Sensitivities and specificities to predict in-hospital mortality using cut-off values from previous studies (see **Table S1**) for scores with an AUROC >0.75 in the analysis using multiple imputed data.

Score name	AUROC [95% CI]					
	In-hospital death within 30 days			In-hospital death or ICU admission within 30 days		
	First wave	Subsequent waves	p-value	First wave	Subsequent waves	p-value
4C Mortality Score	0.793 [0.779-0.807]	0.793 [0.779-0.806]	0.833	0.658 [0.643-0.673]	0.660 [0.647-0.674]	0.887
ABC-GOALSc*	0.627 [0.608-0.647]	0.660 [0.643-0.678]	**0.043**	0.648 [0.633-0.664]	0.661 [0.648-0.675]	0.355
ABCS	0.789 [0.774-0.804]	0.792 [0.778-0.806]	0.979	0.691 [0.677-0.706]	0.674 [0.661-0.688]	**0.040**
A-DROP*	0.744 [0.729-0.760]	0.730 [0.714-0.746]	0.332	0.605 [0.588-0.622]	0.598 [0.583-0.613]	0.677
ANDC	0.757 [0.741-0.772]	0.759 [0.745-0.773]	0.486	0.647 [0.632-0.662]	0.637 [0.624-0.650]	0.703
Bennouar et al.	0.721 [0.705-0.737]	0.726 [0.711-0.742]	0.837	0.694 [0.680-0.709]	0.694 [0.681-0.706]	0.828
CHA(2)DS(2)-VASc	0.674 [0.659-0.689]	0.694 [0.680-0.708]	0.077	0.543 [0.529-0.558]	0.558 [0.545-0.570]	0.184
COPS*	0.742 [0.727-0.757]	0.745 [0.730-0.749]	0.580	0.611 [0.595-0.627]	0.609 [0.595-0.623]	0.975
CORONATION-TR*	0.760 [0.743-0.777]	0.774 [0.759-0.789]	0.375	0.724 [0.710-0.739]	0.723 [0.710-0.735]	0.433
COVID-19 SEIMC*	0.750 [0.736-0.764]	0.755 [0.742-0.768]	0.555	0.589 [0.574-0.603]	0.586 [0.573-0.598]	0.437
COVID-AID*	0.741 [0.727-0.756]	0.754 [0.741-0.767]	0.063	0.562 [0.547-0.577]	0.569 [0.556-0.582]	0.352
COVID-GRAM*	0.759 [0.743-0.775]	0.779 [0.765-0.794]	0.601	0.681 [0.664-0.697]	0.692 [0.679-0.706]	0.683
COVID-NoLab	0.710 [0.694-0.726]	0.698 [0.683-0.713]	0.209	0.630 [0.615-0.646]	0.642 [0.628-0.655]	0.330
COVID-SimpleLab	0.720 [0.704-0.737]	0.720 [0.705-0.735]	0.673	0.673 [0.658-0.688]	0.674 [0.662-0.687]	0.680
CURB-65	0.733 [0.717-0.750]	0.727 [0.710-0.743]	0.805	0.610 [0.593-0.627]	0.606 [0.591-0.621]	0.923
Hachim et al.	0.731 [0.714-0.747]	0.733 [0.719-0.747]	0.844	0.631 [0.616-0.647]	0.615 [0.602-0.628]	0.150
Hu et al.	0.730 [0.713-0.746]	0.735 [0.720-0.750]	0.089	0.660 [0.646-0.675]	0.651 [0.638-0.664]	0.571
KPI Score	0.579 [0.564-0.595]	0.590 [0.575-0.605]	0.979	0.605 [0.592-0.618]	0.619 [0.607-0.632]	0.967
LOW-HARM Score*	0.610 [0.587-0.632]	0.619 [0.598-0.640]	0.141	0.549 [0.532-0.567]	0.549 [0.534-0.565]	0.341
Mei et al. (Full)*	0.730 [0.713-0.748]	0.743 [0.728-0.758]	0.226	0.684 [0.669-0.699]	0.683 [0.670-0.697]	0.446
Mei et al. (Simple)	0.714 [0.698-0.731]	0.730 [0.715-0.746]	0.373	0.634 [0.618-0.651]	0.641 [0.627-0.655]	0.441
NEWS2*	0.648 [0.627-0.668]	0.618 [0.596-0.639]	**0.002**	0.654 [0.636-0.672]	0.654 [0.638-0.670]	0.225
PLANS	0.737 [0.721-0.753]	0.754 [0.739-0.769]	0.112	0.638 [0.623-0.653]	0.633 [0.620-0.647]	0.622
PREDI-CO	0.696 [0.679-0.713]	0.693 [0.677-0.709]	0.344	0.709 [0.695-0.723]	0.697 [0.685-0.710]	0.150
PRESEP	0.604 [0.585-0.623]	0.583 [0.562-0.604]	0.057	0.629 [0.611-0.646]	0.622 [0.607-0.637]	0.364
qSOFA	0.598 [0.578-0.619]	0.584 [0.562-0.606]	0.228	0.576 [0.557-0.594]	0.575 [0.558-0.592]	0.800
RISE UP	0.765 [0.750-0.781]	0.773 [0.758-0.788]	0.583	0.661 [0.645-0.676]	0.659 [0.646-0.673]	0.936
SIMI	0.739 [0.722-0.755]	0.722 [0.707-0.737]	**0.047**	0.681 [0.667-0.695]	0.670 [0.658-0.683]	0.089
SIRS	0.547 [0.529-0.566]	0.545 [0.526-0.563]	0.611	0.588 [0.573-0.604]	0.590 [0.576-0.604]	0.893
STSS	0.706 [0.689-0.723]	0.688 [0.668-0.707]	0.120	0.607 [0.588-0.625]	0.607 [0.590-0.623]	0.755
Wang et al. (Clinical)	0.718 [0.704-0.732]	0.734 [0.722-0.747]	**0.022**	0.545 [0.530-0.559]	0.553 [0.541-0.566]	0.174
Wang et al. (Laboratory)	0.654 [0.636-0.672]	0.636 [0.619-0.653]	0.334	0.671 [0.656-0.685]	0.667 [0.654-0.680]	0.640

*alterations were used to compute these scores. P-value for interaction between score and wave of admission using multivariate logistic regression (formula: outcome~score+wave+score:wave).

Table S9. Discriminative performance of scores examined in the study according to wave of admission.

Score name	AUROC [95% CI]					
	In-hospital death within 30 days			In-hospital death or ICU admission within 30 days		
	Age ≤ 65	Age > 65	p-value	Age ≤ 65	Age > 65	p-value
4C Mortality Score	0.762 [0.736-0.788]	0.724 [0.711-0.738]	0.807	0.704 [0.688-0.719]	0.673 [0.658-0.687]	0.416
ABC-GOALSc*	0.696 [0.664-0.728]	0.636 [0.621-0.651]	**0.002**	0.673 [0.657-0.690]	0.644 [0.630-0.657]	**0.002**
ABCS	0.778 [0.750-0.805]	0.729 [0.715-0.742]	0.066	0.699 [0.684-0.714]	0.681 [0.668-0.694]	0.214
A-DROP*	0.645 [0.611-0.679]	0.660 [0.645-0.675]	0.213	0.603 [0.585-0.621]	0.595 [0.579-0.611]	**<0.001**
ANDC	0.707 [0.679-0.736]	0.686 [0.672-0.700]	0.365	0.676 [0.661-0.692]	0.636 [0.623-0.650]	0.092
Bennouar et al.	0.718 [0.690-0.746]	0.672 [0.659-0.686]	0.080	0.704 [0.689-0.719]	0.686 [0.674-0.698]	0.596
CHA(2)DS(2)-VASc	0.626 [0.595-0.657]	0.552 [0.538-0.565]	**<0.001**	0.576 [0.560-0.591]	0.500 [0.487-0.512]	**<0.001**
COPS*	0.719 [0.690-0.747]	0.653 [0.638-0.668]	**0.018**	0.644 [0.628-0.660]	0.589 [0.573-0.605]	**0.004**
CORONATION-TR*	0.768 [0.741-0.794]	0.717 [0.703-0.731]	0.113	0.733 [0.718-0.748]	0.722 [0.709-0.734]	**0.040**
COVID-19 SEIMC*	0.721 [0.693-0.749]	0.650 [0.637-0.663]	**<0.001**	0.687 [0.672-0.703]	0.535 [0.522-0.547]	**<0.001**
COVID-AID*	0.712 [0.684-0.739]	0.643 [0.630-0.656]	0.654	0.617 [0.601-0.633]	0.530 [0.517-0.543]	**0.004**
COVID-GRAM*	0.778 [0.750-0.805]	0.708 [0.694-0.723]	0.358	0.710 [0.694-0.726]	0.679 [0.665-0.694]	**0.046**
COVID-NoLab	0.685 [0.656-0.715]	0.606 [0.592-0.620]	0.276	0.635 [0.619-0.651]	0.623 [0.611-0.635]	**<0.001**
COVID-SimpleLab	0.694 [0.663-0.724]	0.652 [0.638-0.667]	0.420	0.680 [0.665-0.696]	0.674 [0.662-0.686]	0.172
CURB-65	0.669 [0.635-0.702]	0.641 [0.625-0.657]	0.087	0.607 [0.588-0.626]	0.602 [0.586-0.619]	0.136
Hachim et al.	0.740 [0.710-0.770]	0.654 [0.640-0.668]	**<0.001**	0.631 [0.615-0.647]	0.605 [0.592-0.618]	**0.005**
Hu et al.	0.675 [0.644-0.706]	0.674 [0.660-0.688]	0.123	0.678 [0.663-0.693]	0.647 [0.634-0.660]	**<0.001**
KPI Score	0.600 [0.575-0.626]	0.584 [0.571-0.596]	0.406	0.619 [0.605-0.633]	0.609 [0.597-0.621]	0.977
LOW-HARM Score*	0.577 [0.534-0.620]	0.579 [0.563-0.595]	**<0.001**	0.575 [0.557-0.594]	0.559 [0.545-0.574]	**0.001**
Mei et al. (Full)*	0.703 [0.671-0.735]	0.690 [0.677-0.704]	0.113	0.700 [0.684-0.716]	0.675 [0.663-0.688]	**<0.001**
Mei et al. (Simple)	0.710 [0.679-0.740]	0.658 [0.643-0.672]	0.218	0.664 [0.647-0.681]	0.611 [0.597-0.625]	**0.002**
NEWS2*	0.615 [0.579-0.650]	0.657 [0.641-0.674]	**0.035**	0.651 [0.631-0.670]	0.661 [0.644-0.677]	0.387
PLANS	0.680 [0.650-0.710]	0.672 [0.658-0.687]	0.531	0.666 [0.651-0.682]	0.628 [0.615-0.642]	**<0.001**
PREDI-CO	0.653 [0.625-0.681]	0.677 [0.663-0.690]	0.074	0.707 [0.692-0.721]	0.696 [0.683-0.708]	0.310
PRESEP	0.586 [0.553-0.620]	0.628 [0.612-0.644]	**0.030**	0.623 [0.604-0.642]	0.635 [0.620-0.651]	0.366
qSOFA	0.578 [0.543-0.612]	0.599 [0.582-0.617]	0.253	0.567 [0.547-0.587]	0.582 [0.563-0.600]	0.263
RISE UP	0.744 [0.715-0.774]	0.698 [0.682-0.713]	**<0.001**	0.698 [0.683-0.714]	0.653 [0.639-0.668]	**<0.001**
SIMI	0.651 [0.622-0.680]	0.673 [0.659-0.686]	0.318	0.685 [0.670-0.699]	0.670 [0.658-0.683]	0.441
SIRS	0.561 [0.529-0.593]	0.586 [0.571-0.601]	0.115	0.591 [0.574-0.608]	0.600 [0.587-0.614]	0.238
STSS	0.596 [0.560-0.633]	0.617 [0.599-0.634]	0.135	0.594 [0.574-0.614]	0.606 [0.587-0.625]	0.306
Wang et al. (Clinical)	0.705 [0.678-0.733]	0.601 [0.587-0.614]	**<0.001**	0.612 [0.596-0.627]	0.510 [0.498-0.523]	**<0.001**
Wang et al. (Laboratory)	0.610 [0.577-0.643]	0.635 [0.621-0.650]	**0.005**	0.673 [0.657-0.688]	0.661 [0.648-0.674]	0.893

*alterations were used to compute these scores. P-value for interaction between score and age group using multivariate logistic regression (formula: outcome~score+age group+score:age group).

Table S10. Discriminative performance of scores examined in the study according to age.

Score name	Area under the precision-recall curve	
	In-hospital death within 30 days	In-hospital death or ICU admission within 30 days
4C Mortality Score	0.459	0.521
ABCS	0.449	0.531
CORONATION-TR	0.432	0.583
RISE UP	0.429	0.512
COVID-GRAM	0.407	0.534
Hu et al.	0.393	0.515
ANDC	0.392	0.498
PLANS	0.391	0.479
Mei et al. (full)	0.382	0.539
COVID-AID	0.380	0.416
COVID-19 SEIMC	0.379	0.418
A-DROP	0.376	0.453
COPS	0.370	0.443
SIMI	0.364	0.504
Mei et al. (simple)	0.361	0.486
COVID-SimpleLab	0.358	0.541
CURB-65	0.355	0.453
STSS	0.351	0.477
Bennouar et al.	0.344	0.533
Hachim et al.	0.332	0.454
PREDI-CO	0.326	0.548
COVID-NoLab	0.324	0.499
Wang et al. (Clinical)	0.315	0.376
Wang et al. (Laboratory)	0.294	0.531
CHA(2)DS(2)-VASc	0.284	0.385
NEWS2	0.276	0.507
ABC-GOALSc	0.276	0.493
PRESEP	0.245	0.479
qSOFA	0.237	0.412
KPI Score	0.215	0.425
SIRS	0.199	0.425
LOW-HARM Score	0.161	0.350

Table S11. Area under the precision-recall curve of scores included in the study, ordered by performance to predict in-hospital mortality.

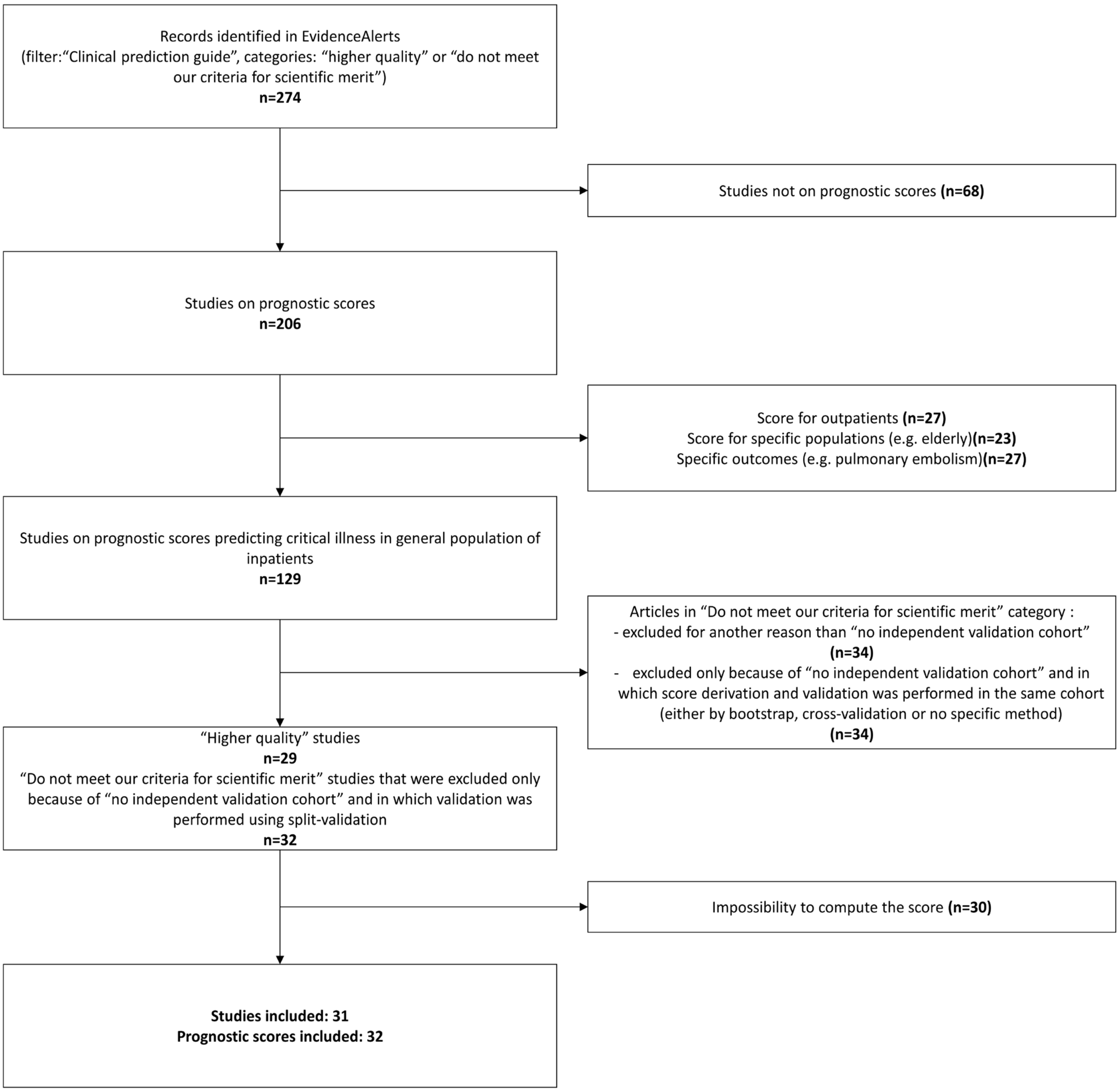
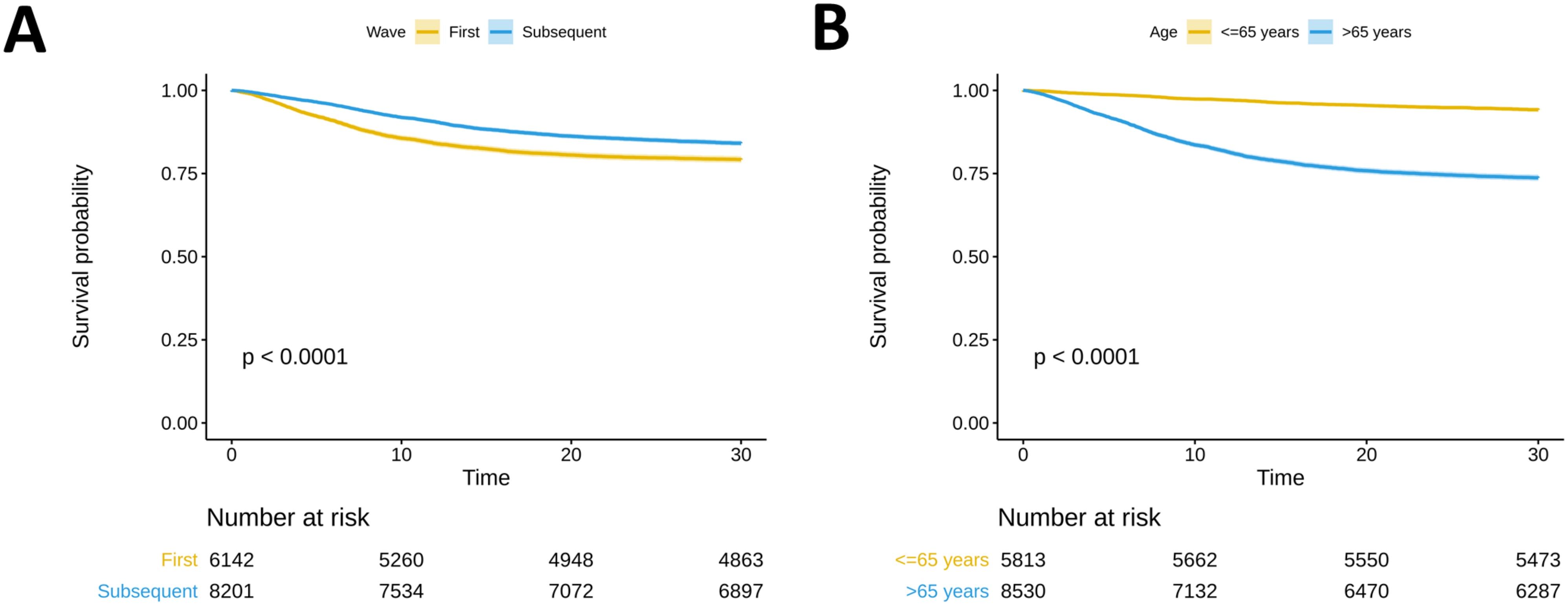


Figure S1. Flow chart for scores' selection. See **Appendix 3** for details on scores included and excluded.



All patients hospitalized for Covid-19 were considered for this analysis. P-values are from Log-Rank tests.

Figure S2. Kaplan-Meier curves for in-hospital mortality according to wave of admission or age.

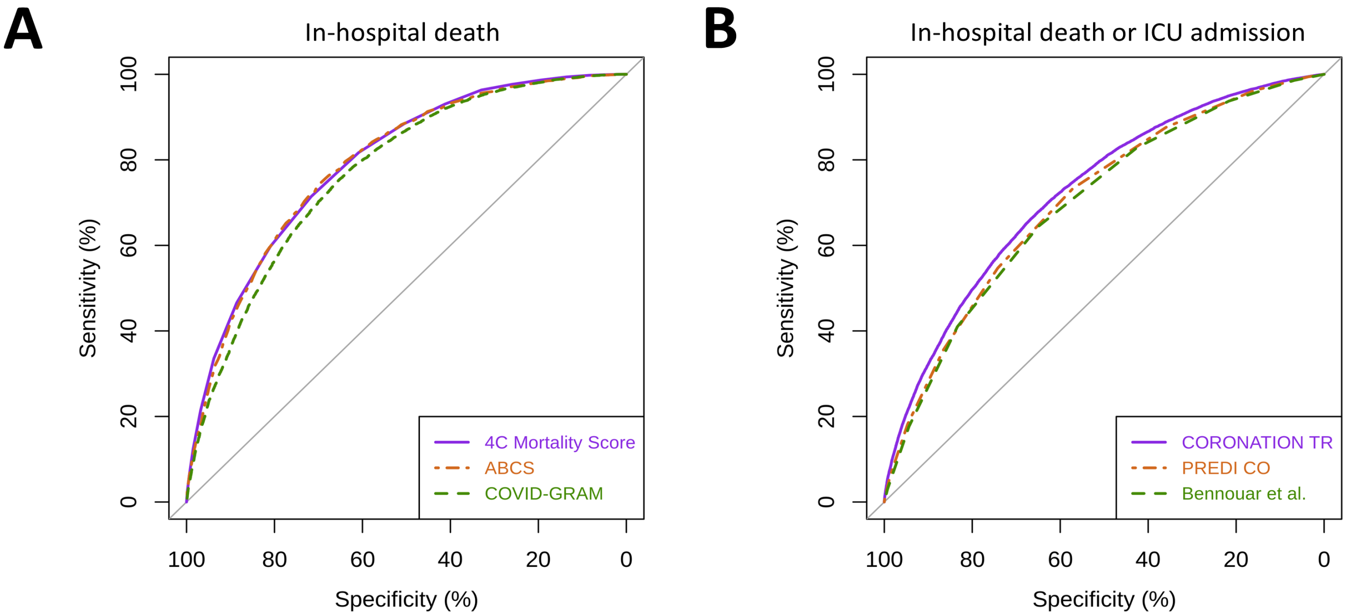
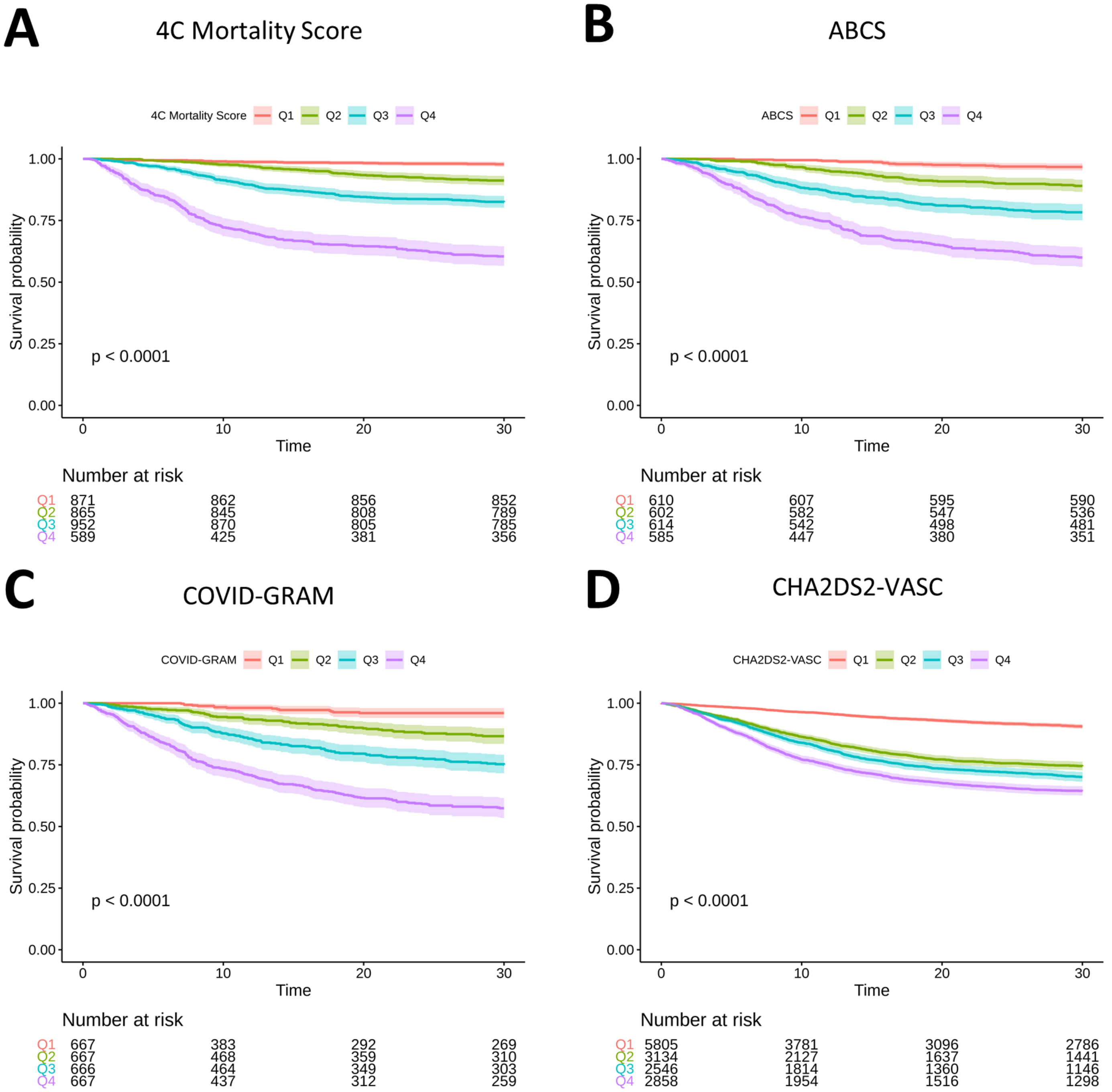


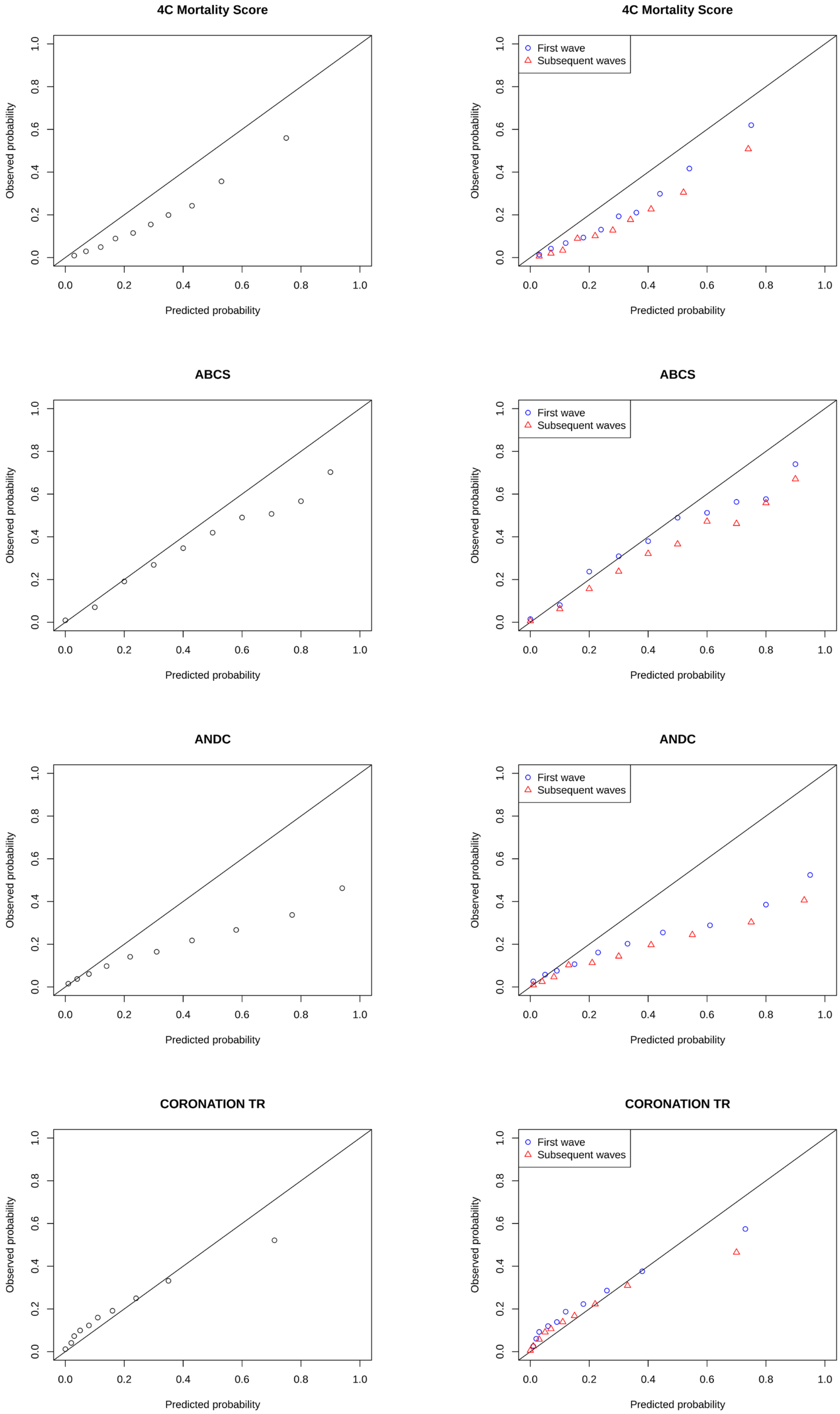
Figure S3. Receiver operating characteristic curves for prediction of in-hospital death within 30 days from admission (**A**) and in-hospital death or ICU admission within 30 days of admission (**B**) among patients hospitalized for Covid-19.

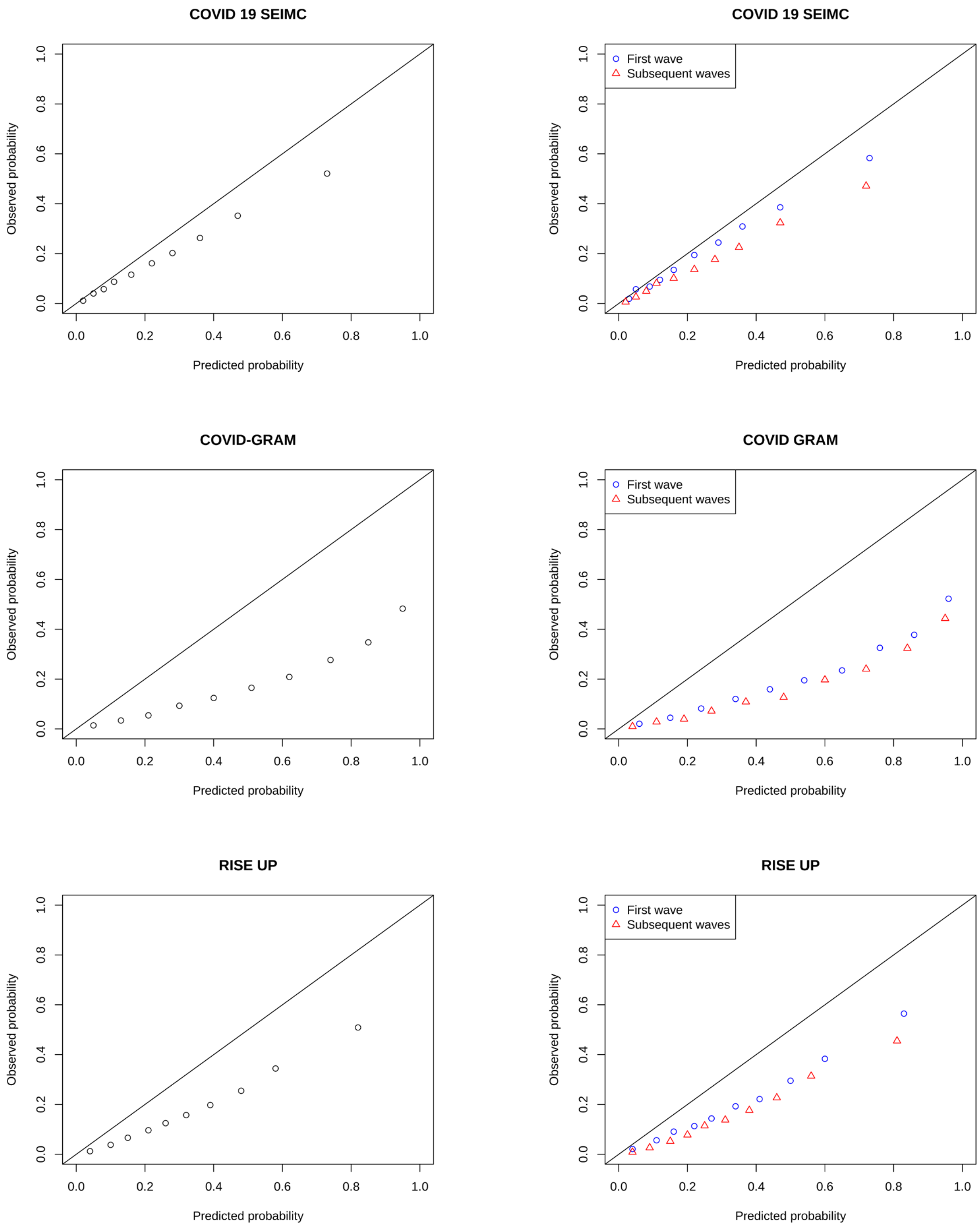


The three scores that performed best to predict in-hospital death are shown (4C Mortality Score, ABCS, COVID-GRAM), and CHA(2)DS(2)-VASc is shown for comparison purposes. For each score, complete data were used (i.e., patients with all data available to compute the score), and patients were grouped according to quartiles (Q1: lowest quartile, to Q4: highest quartile). P-values are from Log-Rank tests.

Figure S4. Kaplan-Meier curves for in-hospital mortality according to the score's value.

Regardless of wave of admission | **According to wave of admission**

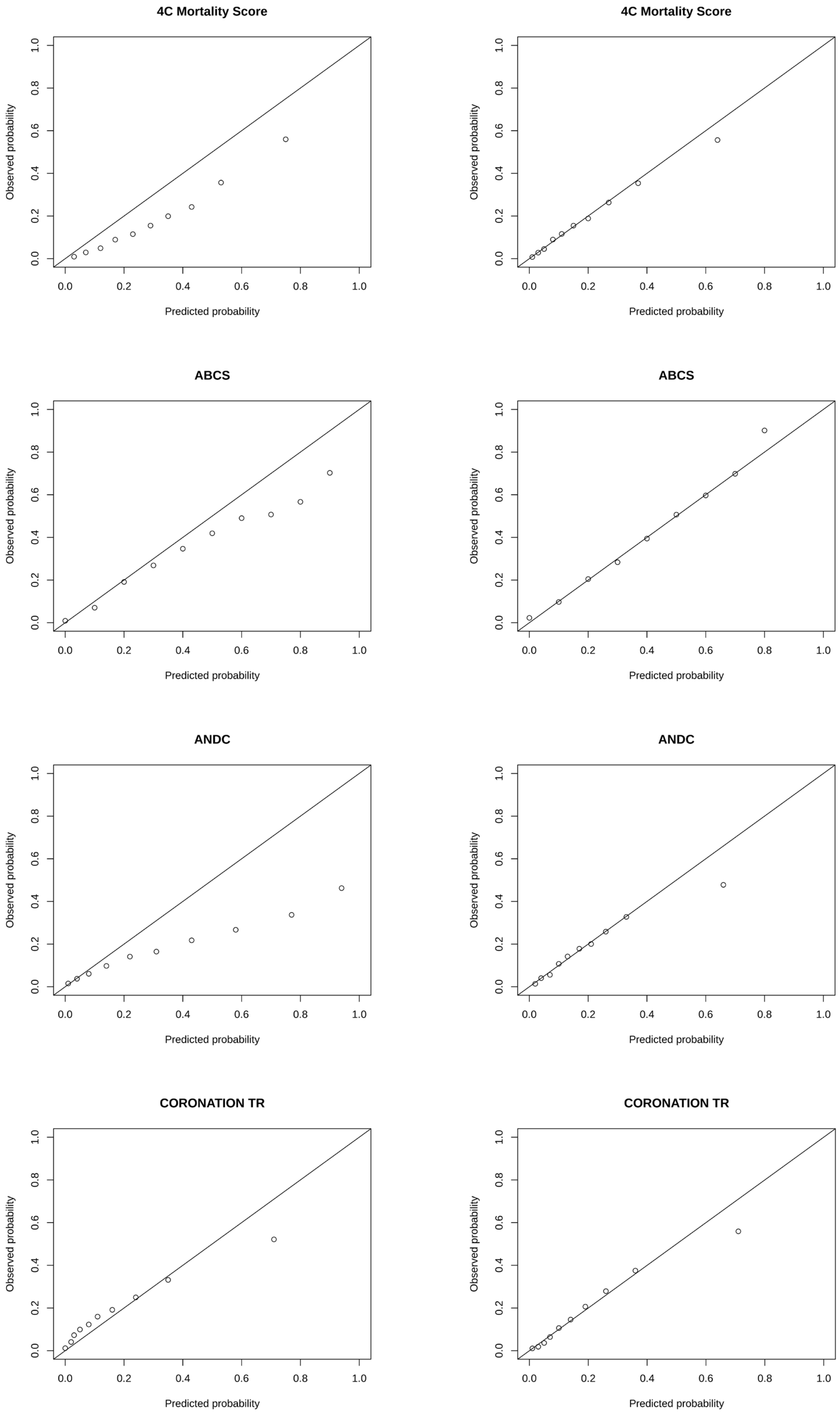


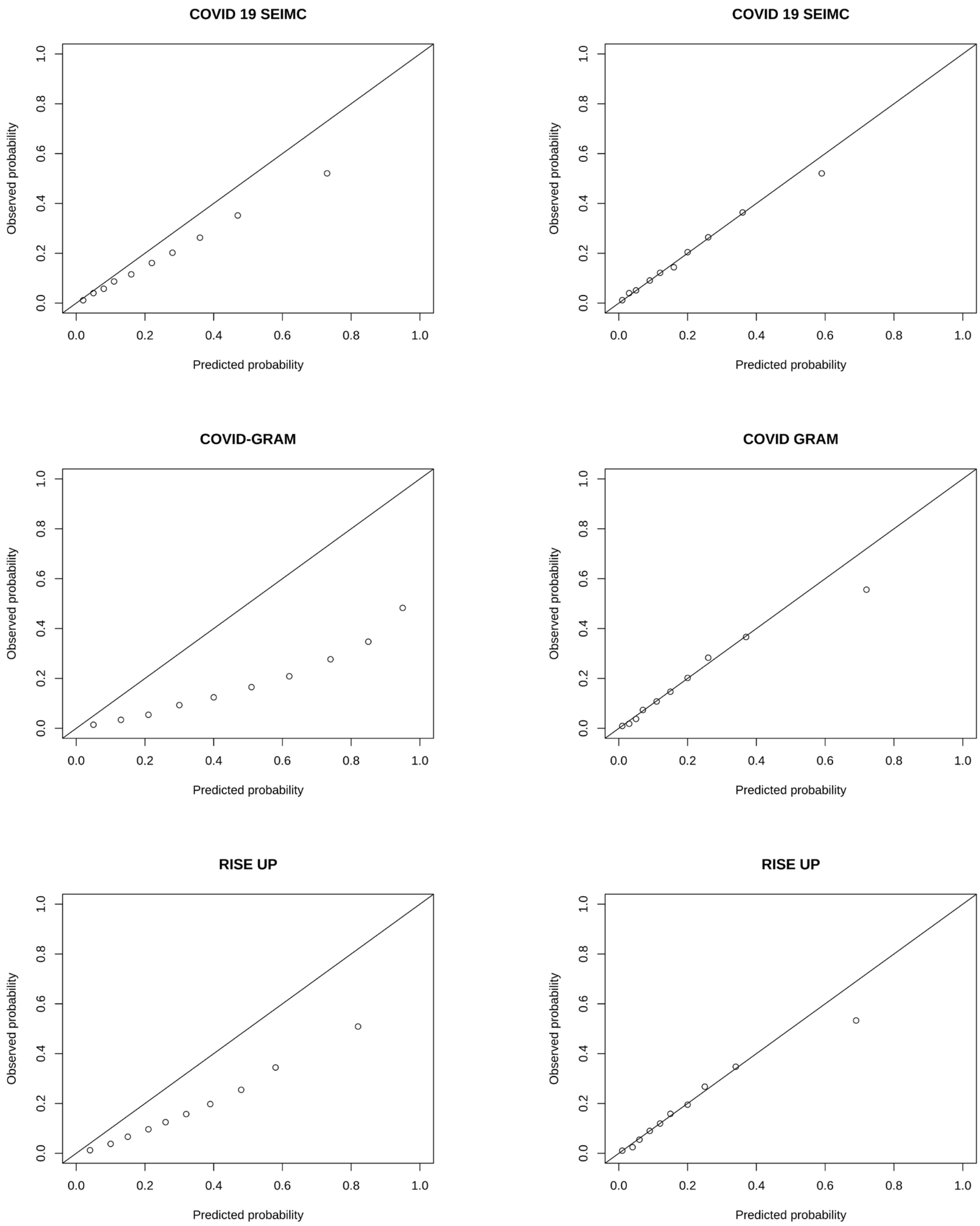


Patients are grouped according to deciles of predicted probability, except for the ABCS score where patients are grouped in classes of fixed width (0.1). Data used is from pooled multiple imputed datasets.

Figure S5. Calibration curves for prediction of 30-day in-hospital mortality for the seven scores with an AUROC > 0.75, considering patients regardless of (left panel) or according to (right panel) wave of admission.

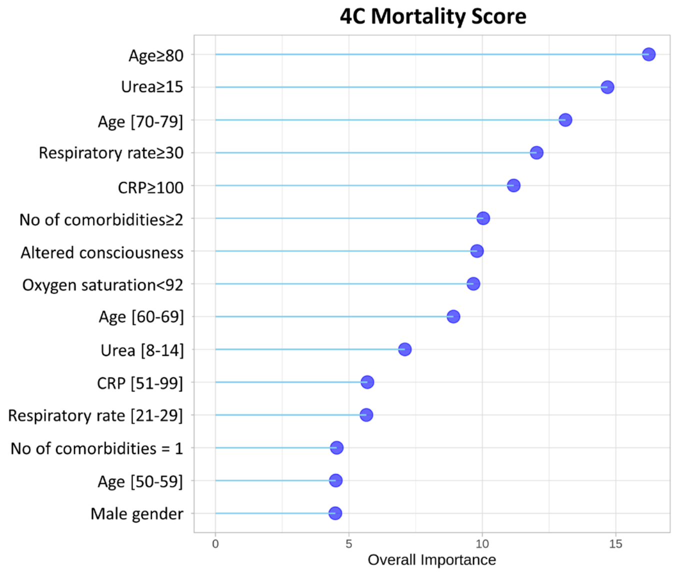
Before revision **After revision**

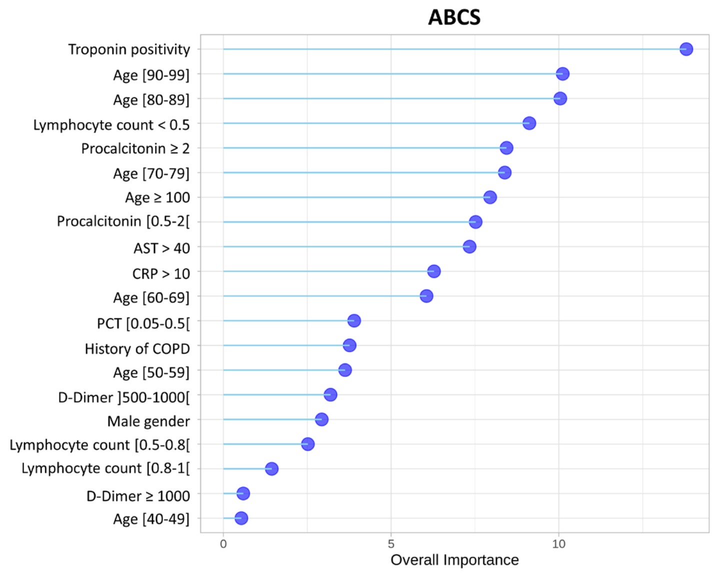


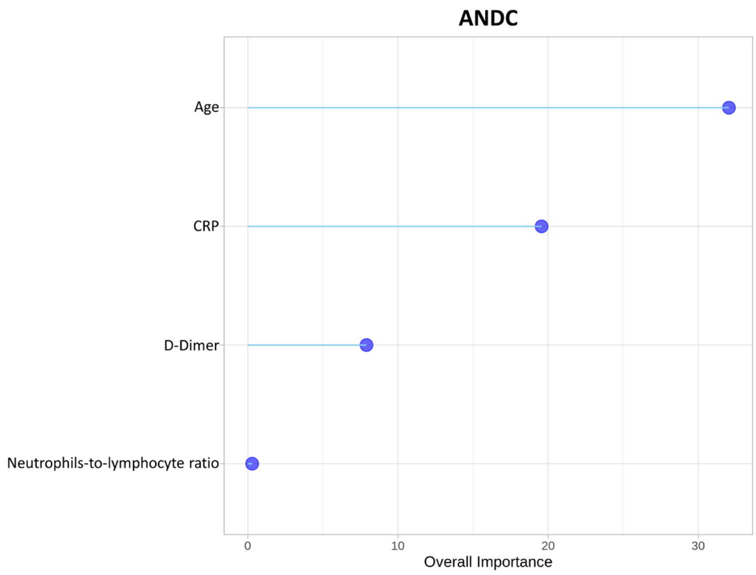


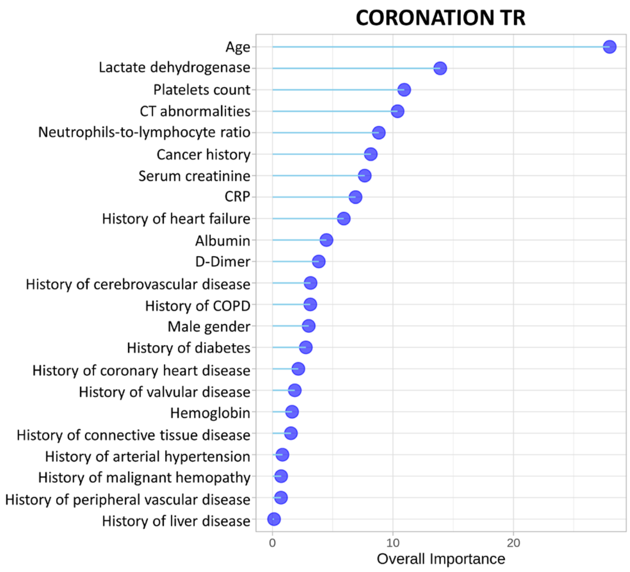
Patients are grouped according to deciles of predicted probability, except for ABCS Score where patients are grouped in classes of fixed width. Data used is from pooled multiple imputed datasets.

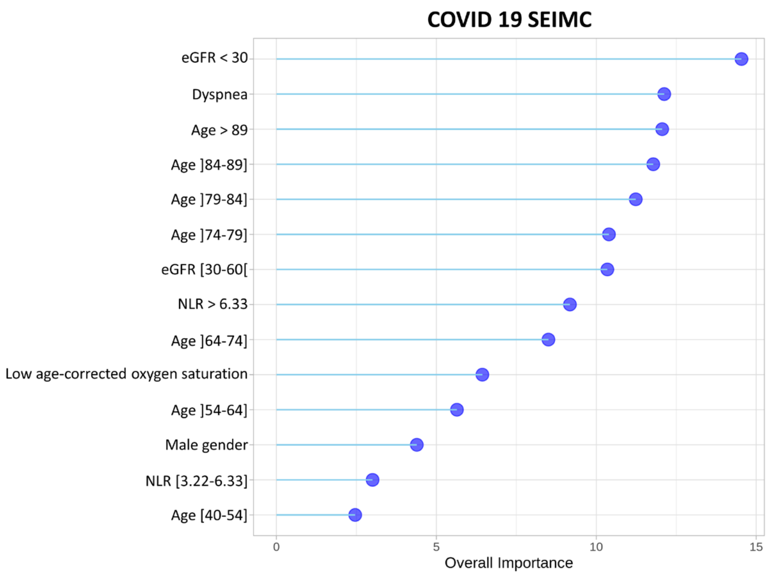
Figure S6. Calibration curves for prediction of 30-day in-hospital mortality for the seven scores with an AUROC > 0.75, before (left panel) and after (right panel) revision.

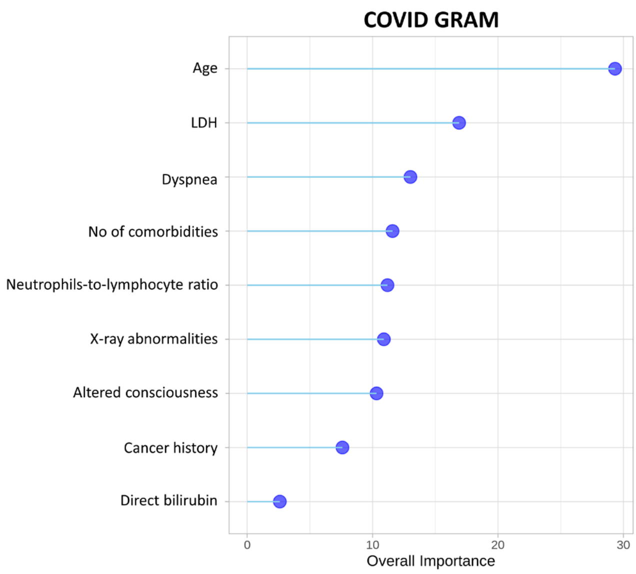

4C Mortality Score
Age≥80
Urea≥15
Age [70-79]
Respiratory rate≥30
CRP≥100
No of comorbidities≥2
Altered consciousness
Oxygen saturation<92
Age [60-69]
Urea [8-14]
CRP [51-99]
Respiratory rate [21-29]
No of comorbidities = 1
Age [50-59]
Male gender
0
5
10
15
Overall Importance

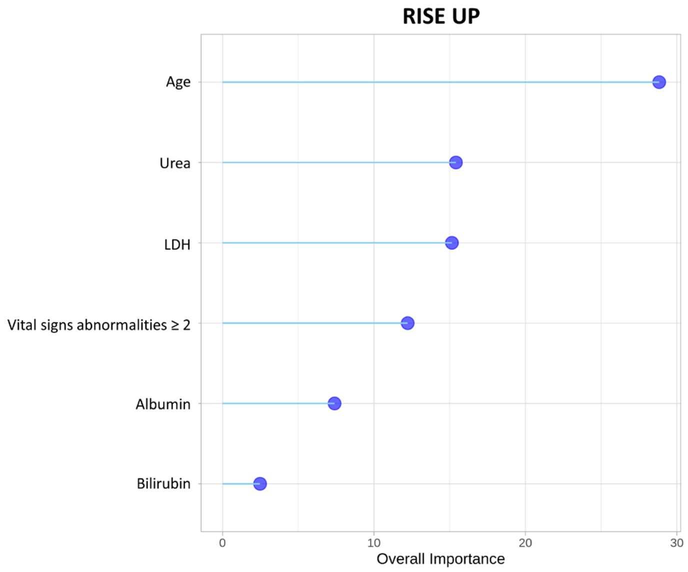

ABCS
Troponin positivity
Age [90-99]
Age [80-89]
Lymphocyte count < 0.5
Procalcitonin ≥ 2
Age [70-79]
Age ≥ 100
Procalcitonin [0.5-2[
AST > 40
CRP > 10
Age [60-69]
PCT [0.05-0.5[
History of COPD
Age [50-59]
D-Dimer]500-1000[
Male gender
Lymphocyte count [0.5-0.8[
Lymphocyte count [0.8-1[
D-Dimer ≥ 1000
Age [40-49]
0
5
10
Overall Importance


ANDC
Age
CRP
D-Dimer
Neutrophils-to-lymphocyte ratio
0
10
20
30
Overall Importance


CORONATION TR
Age
Lactate dehydrogenase
Platelets count
CT abnormalities
Neutrophils-to-lymphocyte ratio
Cancer history
Serum creatinine
CRP
History of heart failure
Albumin
D-Dimer
History of cerebrovascular disease
History of COPD
Male gender
History of diabetes
History of coronary heart disease
History of valvular disease
Hemoglobin
History of connective tissue disease
History of arterial hypertension
History of malignant hemopathy
History of peripheral vascular disease
History of liver disease
0
10
20
Overall Importance


COVID 19 SEIMC
eGFR < 30
Dyspnea
Age > 89
Age]84-89]
Age]79-84]
Age]74-79]
eGFR [30-60[
NLR > 6.33
Age]64-74]
Low age-corrected oxygen saturation
Age]54-64]
Male gender
NLR [3.22-6.33]
Age [40-54]
0
5
10
15
Overall Importance


COVID GRAM
Age
LDH
Dyspnea
No of comorbidities
Neutrophils-to-lymphocyte ratio
X-ray abnormalities
Altered consciousness
Cancer history
Direct bilirubin
0
10
20
30
Overall Importance



For the ABCS Score, classes of age [0-20[, [20-30[and [30-40[were regrouped for the analysis to be interpretable, as otherwise the reference class (i.e., [0-20[) would have had few patients (n=27). Data used is from the first imputed dataset.

Figure S7. Variable importance analysis for prediction of 30-day in-hospital mortality for the seven scores with an AUROC > 0.75.

Appendix 3. Selection, reasons for exclusion and information on scores included in the study.

Articles whose main purpose was not to derive or test prognostic scores for Covid-19 : (n = 68)

10.1007/s00330-020-07087-y. ; 10.3390/jpm11010036. ; 10.2196/23897. ; 10.7759/cureus.12565. ; 10.3389/fmed.2020.577609.; 10.1016/j.media.2020.101844.; 10.1007/s11739-020-02534-6. ; 10.1007/s00330-020-06829-2. ; 10.2196/24478. ; 10.1111/tmi.13476.; 10.1177/1753466620963019.; 10.1093/qjmed/hcaa305.; 10.3390/jcm9103350. ; 10.7326/M20-3905.; 10.1016/j.dsx.2020.03.017. ; 10.1371/journal.pone.0239474.; 10.3390/diagnostics11010041. ; 10.18632/aging.104132. ; 10.1183/13993003.03498-2020.; 10.1007/s00261-020-02823-w.; 10.1007/s11357-020-00294-x.; 10.1016/j.chest.2020.05.580. ; 10.1097/MD.0000000000022980. ; 10.1016/j.jcv.2020.104502. ; 10.1016/j.amjmed.2020.10.044. ; 10.1515/cclm-2020-0593.; 10.1007/s42979-020-00394-7.; 10.1016/j.bjid.2020.07.003.; 10.1007/s00521-020-05437-x. ; 10.1111/acem.14182. ; 10.1016/j.rmed.2020.106206. ; 10.31661/jbpe.v0i0.2008-1153. ; 10.1007/s42979-020-00394-7.; 10.1016/j.ijid.2020.09.022.; 10.1159/000512209.; 10.3390/diagnostics10090619. ; 10.1111/ijcp.13926.; 10.1007/s42399-020-00603-7.; 10.1016/j.jamda.2020.08.030. ; 10.1093/cid/ciaa322. ; 10.1038/s41598-020-76141-y. ; 10.1371/journal.pone.0243414.; 10.1016/j.dsx.2020.10.022.; 10.1016/j.bjid.2020.06.009. ; 10.1007/s00259-020-05075-4; 10.1016/j.ejrad.2020.109041.; 10.1136/bmj.m1328.; 10.1016/j.cca.2020.11.019.; 10.21037/atm.2020.03.132. ; 10.3233/XST-200735. ; 10.1016/j.ajog.2020.10.032. ; 10.1016/j.media.2020.101824. ; 10.1038/s41746-020-00372-6.; 10.4269/ajtmh.20-0730 ; 10.1371/journal.pone.0237202 ; 10.3390/jcm10040570 ; 10.1038/s41598-021-82885-y ; 10.1136/bmjopen-2020-047110 ; 10.1093/cid/ciab177 ; 10.1371/journal.pone.0248438 ; 10.1371/journal.pone.0247773 ; 10.2196/23582 ; 10.1186/s12879-021-05930-1 ; 10.18632/aging.202735 ; 10.11622/smedj.2021019 ; 10.21037/atm-20-3073 ; 10.1038/s41598-021-86735-9 ; 10.1007/s40121-021-00437-3.

f

Articles on scores to be used partially or completely for outpatients : (n = 27)

10.1038/s41598-020-75767-2. ; 10.1016/j.archger.2020.104240.; 10.1136/bmj.m3731.; 10.1093/ofid/ofaa463.; 10.1111/ijcp.13705.; 10.1093/ije/dyaa209. ; 10.1016/j.annemergmed.2020.07.022.; 10.1080/07853890.2020.1828616. ; 10.24875/RIC.20000295. ; 10.2196/21801. ; 10.1371/journal.pone.0237419.; 10.1371/journal.pone.0241825.; 10.3390/jcm9113726. ; 10.3389/fpubh.2020.587937.; 10.1136/jitc-2020-001314. ; 10.1016/S2213-8587(20)30405-8. ; 10.1371/journal.pmed.1003374. 10.1371/journal.pone.0237202. ; 10.1371/journal.pone.0236554.; 10.1016/j.ajem.2020.10.068. ; 10.1093/infdis/jiaa663. ; 10.1371/journal.pone.0240346. ; 10.1016/S2589-7500(20)30217-X. ; 10.1136/thoraxjnl-2020-216425 ; 10.1016/j.pmedr.2020.101298 ; 10.1186/s12967-021-02720-w ; 10.1002/jmv.26890 ; 10.1111/jgs.17089.

Articles on scores to be used partially or completely in a specific population (e.g. ICU patients or elderly) : (n = 23)

10.1093/ageing/afaa240. ; 10.1002/jmv.26572.; 10.7717/peerj.10083.; 10.2147/CIA.S273720. ; 10.1097/CCM.0000000000004549. ; 10.1016/j.ajem.2020.07.019.; 10.2196/23128. ; 10.3389/fonc.2020.01560. ; 10.1016/j.amsu.2020.09.044.; 10.1016/j.eclinm.2020.100426. ; 10.7717/peerj.10018.; 10.1080/03007995.2020.1825365.; 10.1371/journal.pone.0247275 ; 10.1016/j.archger.2021.104383 ; 10.3390/membranes11030170 ; 10.5603/ARM.a2020.0176 ; 10.21037/atm-20-7447 ; 10.2196/23026 ; 10.1186/s13054-021-03487-8 ; 10.1097/MD.0000000000024901 ; 10.1136/jitc-2020-002277 ; 10.7759/cureus.14051 ; 10.1093/ckj/sfab037.

Articles on scores to predict outcomes other than ICU admission, death, mechanical ventilation, or outcomes considered equivalent to those (e.g. septic shock was considered, pulmonary embolism or need for oxygen therapy was not considered) : (n = 27)

10.1016/j.ajem.2020.09.051.; 10.3389/fmed.2020.556886.; 10.1136/annrheumdis-2020-218323. ; 10.1371/journal.pone.0239172.; 10.1093/cid/ciaa443. ; 10.2147/IDR.S263157.; 10.2196/22131. ; 10.1002/iid3.353. ; 10.1093/cid/ciaa414. ; 10.1007/s11606-020-06353-5. ; 10.1007/s15010-020-01446-z. ; 10.1111/crj.13296.; 10.7883/yoken.JJID.2020.718.; 10.1038/s41746-020-00343-x.; 10.1186/s12911-020-01338-0. ; 10.7717/peerj.9945. ; 10.2214/AJR.20.24044.; 10.1016/j.ebiom.2020.102880. ; 10.1186/s12880-020-00513-z. ; 10.1093/qjmed/hcaa224. ; 10.2147/IDR.S261725.; 10.7150/ijms.47193. ; 10.7150/ijms.50007 ; PMC7821745 ; 10.1016/j.jaclp.2020.12.005 ; 10.1371/journal.pone.0248230. ; 10.1097/MD.0000000000024441

Articles in the "do not meet our criteria for scientific merit" group excluded for another reason than "no independent validation cohort": (n = 34)

10.1007/s11547-020-01200-3. ; 10.1016/j.chest.2020.04.010. ; 10.1016/j.resuscitation.2020.08.124. ; 10.1093/cid/ciaa963.; 10.1080/23744235.2020.1784457. ; 10.3346/jkms.2020.35.e234. ; 10.2196/25442. ; 10.1007/s00521-020-05592-1.; 10.3389/fpubh.2020.00475.; 10.1038/s41551-020-00633-5. ; 10.1016/j.ijid.2020.06.038. ; 10.1016/j.chest.2020.12.009. ; 10.1513/AnnalsATS.202006-698OC.; 10.5603/ARM.a2020.0176.; 10.1136/jim-2020-001525. ; 10.1183/13993003.01104-2020.; 10.1093/cid/ciaa793. ; 10.3389/fmed.2020.590460.; 10.1371/journal.pone.0236618.; 10.3348/kjr.2020.0485. ; 10.1371/journal.pone.0233328. ; 10.1097/CCM.0000000000004411. ; 10.2196/24246 ; 10.1016/S2589-7500(20)30274-0 ; 10.1016/j.media.2021.101975 ; 10.1093/jamia/ocab018 ; 10.1503/cmaj.202795; 10.1007/s11606-021-06626-7 ; 10.4414/smw.2021.20471 ; 10.1038/s41467-020-20816-7 ; 10.1016/S2666-7568(21)00006-4 ; 10.1371/journal.pone.0247676 ; 10.1080/07853890.2021.1891453 ; 10.26355/eurrev_202102_25118.

Articles in the "do not meet our criteria for scientific merit" group excluded only because of "no independent validation cohort", and in which score derivation and validation was performed in the same cohort (either by bootstrap, cross-validation or no specific method) : (n = 34)

10.1186/s12911-020-01316-6. ; 10.1136/bmjopen-2020-041983. ; 10.1016/j.acra.2020.09.004.; 10.1002/jmv.26713. ; PMID: 32913530 ; 10.1016/j.ijantimicag.2020.106110. ; 10.3390/pathogens9110880. ; 10.1016/j.bja.2020.11.034. ; 10.1183/23120541.00359-2020. ; 10.2196/24973.; 10.3390/pathogens10010058. ; 10.1017/dmp.2021.8.; 10.1016/j.jaci.2020.07.009. ; 10.7759/cureus.11786. ; 10.1088/1361-6560/abbf9e. ; 10.1111/dth.14828 ; 10.2196/24572 ; 10.3389/fmed.2020.597791; 10.1038/s41746-021-00383-x ; 10.1186/s12911-020-01359-9 ; 10.1016/j.echo.2021.02.003 ; 10.4269/ajtmh.20-1039 ; 10.1080/07853890.2021.1884744 ; 10.1038/s41598-021-83054-x ; 10.1038/s41598-021-83784-y ; 10.1136/jclinpath-2020-207157 ; 10.1038/s41598-021-83967-7 ; 10.3389/fmed.2021.608107 ; 10.1155/2021/8840835 ; 10.21037/jtd-20-2580 ; 10.2196/23948 ; 10.2196/27060 ; 10.2196/26211. ; 10.1007/s11239-021-02405-7

Articles that could not be computed in our cohort, either in the "high quality studies" group or in the "do not meet our criteria for scientific merit" group excluded only because of "no independent validation cohort" and using split validation: (n = 30)

10.26355/eurrev_202003_20709. (classifier prediction model with no information on how to compute; variables with significant importance missing or not applicable in our cohort: region, confirmed date, group, infection reason, country)
10.1016/j.jcrc.2020.10.033. (random forest with need for repeated data in a 24 hours period)
10.1259/bjr.20200634. (CT-based radiomics nomogram)
10.1055/s-0040-1716544. (score derived on patients hospitalized in GPUH hospitals)
10.1080/07853890.2020.1868564. (variables with significant importance missing or not applicable in our cohort: score mainly based on IL-10)
10.7717/peerj.10337. (deep learning prediction model with no information on how to compute)
10.1186/s12879-020-05561-y. (sample with complete data in our cohort was considered too small, mainly due to the concomitant use of LDH, ferritin, procalcitonin and D-Dimer in the score ; furthermore, sample size for split validation was considered too small: 66 patients)
10.1136/bmjspcare-2020-002602. (variables with significant importance missing or not applicable in our cohort: many variables missing among a total of 51 variables used in this score)
10.3390/ijerph17228386. (the main purpose of this study was to create various machine-learning models that cannot be computed in our cohort; for the logistic regression analysis, variables with significant importance missing or not applicable in our cohort: residential institution, oncological patient deterioration)
10.1177/0300060520955037. (sample with complete data in our cohort was considered too small, mainly due to the concomitant use of D-dimer and ferritin in the score ; furthermore, sample size for split validation was considered too small: 44 patients)
10.7717/peerj.9885. (the main purpose of this study was to create various machine-learning models that cannot be computed in our cohort; for the logistic regression analysis, variables with significant importance missing or not applicable in our cohort: BNP, platelets volume)
10.1371/journal.pone.0242953. (sample with complete data in our cohort was considered too small, mainly due to the concomitant use of LDH, troponin I, ferritin and procalcitonin in the score)
10.3389/fmed.2020.00518.(variables with significant importance missing or not applicable in our cohort: bacterial coinfection, multilobular infiltration)
10.2196/21788. (variables with significant importance missing or not applicable in our cohort: RBC distribution width, chlorine)
10.1186/s13049-020-00795-w. (variable with significant importance missing or not applicable in our cohort: smoking status)
10.1016/j.medj.2020.12.013 (variable with significant importance missing or not applicable in our cohort: platlet count decrease, neutrophils count increase, WBC count increase)
10.1038/s41598-021-81844-x (machine learning model with no information on how to compute and use of repeated data)
10.1186/s40779-021-00315-6 (variable with significant importance missing or not applicable in our cohort: IL-6)
10.1371/journal.pone.0245840 . (variable with significant importance missing or not applicable in our cohort: performance status)
10.1038/s41598-021-81732-4 (variable with significant importance missing or not applicable in our cohort: CD8+ T-cells count)
10.1038/s41598-021-82492-x (multiple machine-learnjng models to predict ARDS, using variables with significant importance missing or not applicable in our cohort)

10.1080/03007995.2021.1891036 (variable with significant importance missing or not applicable in our cohort: SaFiO2)
10.7326/M20-6754 (use of time-varying variables)
10.1038/s41598-021-84603-0 (variable with significant importance missing or not applicable in our cohort: smoking status, ethnicity)
10.1016/j.compbiomed.2021.104304 (variable with significant importance missing or not applicable in our cohort: imaging data)
10.3389/fmed.2021.629296 (variable with significant importance missing or not applicable in our cohort: alpha-hydroxybutyrate dehydrogenase, IL-6)
10.33393/jcb.2021.2194 (the main purpose of this study was to create various machine-learning models that cannot be computed in our cohort)
10.3348/kjr.2020.1104 (variable with significant importance missing or not applicable in our cohort: imaging data)
10.1016/S2589-7500(21)00039-X (variable with significant importance missing or not applicable in our cohort: imaging data)
10.1371/journal.pone.0249285 (the main purpose of this study was to create a machine-learning model that cannot be computed in our cohort)

Articles selected for further evaluation, and scores considered if multiple scores were examined : (n = 26)
10.1016/j.cmi.2020.08.003. (PREDI-CO)
10.3389/fmed.2020.585003. (Hachim et al.)
10.1371/journal.pone.0239536. (COVID-AID)
10.1016/j.resplu.2020.100042. (SIRS)
10.1093/ije/dyaa171. (Hu et al.)
10.1080/1354750X.2020.1841296. (KPI Score)
10.1136/bmj.m3339. (4C Mortality Score, A-DROP, CURB-65, qSOFA)*
10.1186/s12879-020-05688-y. (PLANS)
10.1001/jamainternmed.2020.2033. (COVID-GRAM)
10.1136/bmjopen-2020-044028. (Mei et al.: full and clinical)
10.21149/11684. (ABC-GOALSc)
10.1186/s13049-020-00764-3. (NEWS2)
10.1016/j.amjcard.2020.09.029. (CHA2DS2-VASC)
10.1002/emp2.12259. (LOW-HARM)
10.1093/cid/ciaa538. (Wang et al.: clinical and laboratory)
10.1186/s12967-020-02505-7. (ANDC)
10.22088/cjim.11.0.536 (PRESEP)
10.1136/bmjopen-2020-045141 (RISE UP)
10.1007/s11739-020-02617-4 (SIMI)
10.1136/thoraxjnl-2020-216001 (COVID-19 SEIMC)
10.1590/1516-3180.2020.0649.r1.10122020 (STSS)
10.21037/atm-20-6205 (ABCS)
10.1016/j.iccn.2021.103012 (Bennouar et al.)
10.1002/jmv.26844 (CORONATION-TR)
10.2196/26257 (COPS)
10.3122/jabfm.2021.S1.200464 (COVID-NoLab and COVID-SimpleLab)

* PSI and E-CURB were also examined in this article but were not considered as they could not be computed in our cohort (for PSI, variables not collected : nursing home, chest X-ray, hematocrit, glucose, pH; for E-CURB : sample with complete data in our cohort was considered too small, mainly due to the concomitant use of albumin and LDH) ; NEWS was not considered as NEWS2 was already considered

Articles on scores already included : (n = 5)**
10.7861/clinmed.2020-0688 (NEWS2)
10.3389/fmed.2020.624255. (NEWS2)
10.1136/bmjopen-2020-043721. (NEWS2)
10.1111/ijcp.14121 (NEWS***)
10.1007/s11239-021-02427-1. (CHA2DS2-VASC)

** the first published article on a given score was considered to get data on this score's performances
*** NEWS was not considered as NEWS2 was already considered