Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Validación multicéntrica de Sistemas de Alerta Temprana para la detección precoz de deterioro clínico en pacientes hospitalizados por COVID-19

Ivan Alfredo Huespe Indalecio Carboni Bisso Eduardo San Roman Eduardo Prado Nicolás Gemelli Jorge Sinner Marcos Las Heras Marcelo Raul Risk

PIL: S0210-5691(21)00273-4
DOI: https://doi.org/doi:10.1016/j.medin.2021.11.002
Reference: MEDIN 1718
To appear in: Medicina intensiva
Received Date: 2 September 2021
Accepted Date: 8 November 2021

Please cite this article as: Huespe IA, Bisso IC, Roman ES, Prado E, Gemelli N, Sinner J, Heras ML, Risk MR, Validación multicéntrica de Sistemas de Alerta Temprana para la detección precoz de deterioro clínico en pacientes hospitalizados por COVID-19, Medicina intensiva (2021), doi: https://doi.org/10.1016/j.medin.2021.11.002

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.
Multicenter validation of Early Warning Scores for detection of clinical deterioration in COVID-19 hospitalized patients

Validación multicéntrica de Sistemas de Alerta Temprana para la detección precoz de deterioro clínico en pacientes hospitalizados por COVID-19

Ivan Alfredo Huespe1,2,3, Indalecio Carboni Bisso1, Eduardo San Roman1, Eduardo Prado1,4, Nicolás Gemelli1, Jorge Sinner1, Marcos Las Heras1, Marcelo Raul Risk2

1 Intensive Care Unit, Hospital Italiano de Buenos Aires - Argentina.
2 Instituto de Medicina Traslacional e Ingeniería Biomédica, HIBA, IUHI, CONICET, Buenos Aires, Argentina
3 Internal Medicine Research Unit, Hospital Italiano de Buenos Aires, CABA, Argentina
4 Instituto de ciencias aplicadas Sergio Provenzano (ICAP), Facultad de Medicina, Universidad de Buenos Aires

Corresponding author: Iván Huespe
Intensive Care Unit, Hospital Italiano de Buenos Aires, Argentina.
Email: ivan.huespe@hospitalitaliano.org.ar
Address: Tte. Gral. J. D. Perón 4190 (C1199ABB) - Buenos Aires - Argentina
Phone: (+54 011) 4959-0200

Running title: Early Warning Scores for COVID-19

Abstract word count: 250

Manuscript word count: 1990

Financing: this study did not receive any funding.
Conflicts of interest: none

Ethical approval: The study was approved by the Ethics Committee of the Hospital Italiano de Buenos Aires (protocol number 5562).
CREDIT AUTHOR STATEMENT

1) Concept and design: Ivan Alfredo Huespe, Indalecio Carboni Bisso, Eduardo San Roman, Marcelo Raul Risk.

2) Acquisition, analysis, or interpretation of data: Ivan Alfredo Huespe, Indalecio Carboni Bisso, Eduardo Prado.

3) Drafting of the manuscript: Ivan Alfredo Huespe, Indalecio Carboni Bisso, Nicolás Gemelli, Eduardo Prado.

4) Critical revision of the manuscript for important intellectual content: Nicolás Gemelli, Eduardo San Roman, Jorge Sinner, Marcos Las Heras, Marcelo Raul Risk, Eduardo Prado.

5) Statistical analysis: Ivan Alfredo Huespe, Marcelo Raul Risk.

6) Supervision: Eduardo San Roman, Jorge Sinner, Marcos Las Heras, Marcelo Raul Risk.

ABSTRACT

Objective: investigate the predictive value of NEWS2, NEWS-C, and COVID-19 Severity Index for predicting intensive care unit (ICU) transfer in the next 24 hours.

Design: retrospective multicenter study

Setting: two third-level hospitals in Argentina

Patients: all adult patients with confirmed COVID-19, admitted on general wards, excluding patients with non-intubated orders.

Interventions: Patients were divided between those who were admitted to ICU and non-admitted. We calculated the three scores for each day of hospitalization.
Variables: we evaluate the calibration and discrimination of the three scores for the outcome ICU admission within 24, 48 h, and at hospital admission.

Results: we evaluate 13,768 days of hospitalizations on general medical wards of 1318 patients. Among these, 126 (9.5%) were transferred to ICU. The AUROC of NEWS2 was 0.73 (95%CI 0.68-0.78) 24 hours before ICU admission, and 0.52 (95%CI 0.47-0.57) at hospital admission. The AUROC of NEWS-C was 0.73 (95%CI 0.68-0.78) and 0.52 (95%CI 0.47-0.57) respectively, and the AUROC of COVID-19 Severity Index was 0.80 (95%CI 0.77-0.84) and 0.61 (95%CI 0.58-0.66) respectively. COVID-19 Severity Index presented better calibration than NEWS2 and NEWS-C.

Conclusion: COVID-19 Severity index has better calibration and discrimination than NEWS2 and NEWS-C to predict ICU transfer during hospitalization.

RESUMEN

Objetivo: investigar el valor predictivo de los scores NEWS2, NEWS-C y COVID-19 Severity Index para predecir la transferencia de urgencia a la unidad de cuidados intensivos (UCI) en las próximas 24 horas.

Diseño: estudio multicéntrico retrospectivo

Ámbito: dos hospitales de tercer nivel en Argentina

Pacientes: pacientes adultos con COVID-19, ingresados en salas generales, excluyendo pacientes con órdenes de no intubar.

Intervenciones: Se dividió a los pacientes entre los que ingresaron en la UCI y los que no ingresaron. Calculamos las tres puntuaciones para cada día de hospitalización.
Variables: evaluamos la calibración y discriminación de las tres puntuaciones predecir el traslado de urgencia a UCI en las 24, 48 hs previas al pase a UCI y al ingreso hospitalario.

Resultados: evaluamos 13,768 días de hospitalización en internación general de 1318 pacientes, de los cuales 126 (9,5%) fueron trasladados a UCI. El AUROC del NEWS2 fue de 0,73 (IC 95% 0,68-0,78) 24 antes del ingreso en UCI y de 0,52 (IC 95% 0,47-0,57) al ingreso hospitalario. El AUROC de NEWS-C fue de 0,73 (95% CI 0,68-0,78) y 0,52 (95% CI 0,47-0,57) respectivamente, y el AUROC del COVID-19 Severity Index fue 0,80 (95% CI 0,77-0,84) y 0,61 (95% CI 0,58-0,66) respectivamente. El COVID-19 Severity Index presentó una mejor calibración que NEWS2 y NEWS-C.

Conclusión: El COVID-19 Severity Index presentó una mejor calibración y discriminación que NEWS2 y NEWS-C para predecir la transferencia de la UCI durante la hospitalización.

Keywords: COVID-19; Early Warning Score, SARS-CoV-2; critical care.
INTRODUCTION

Early Warning Scores (EWS) detect high-risk patients during hospitalization and are vital for decision making, planning, and proper use of available resources\(^1\). The National Early Warning Score 2 (NEWS2), the most commonly used EWS, was initially recommended by the Royal College of Physicians (RCP)\(^2\). However, during the early phase of the COVID-19 pandemic, Liao et al.\(^3\) proposed the Modified NEWS (NEWS-C) adding age as a variable to reflect the emerging evidence that this variable was an independent risk factor for survival in COVID-19 patients\(^3\). In a cohort of 116 patients, the NEWS-C at hospital admission proved to be the most accurate scoring system to identify infected patients at risk of early deterioration of the respiratory function and need for intensive respiratory support\(^4\).

The increasing number of COVID-19 patients, its complexity, and the multiple variables involved in its course exposed the need for a more specific tool that could be used exclusively in this single disease to better discern among patients at risk of presenting severe infection. Therefore, Huespe et al. developed an EWS called COVID-19 Severity Index to predict progression towards critical illness. This score evaluates multiple characteristics of patients as age, comorbidities, chest X-ray vital signs, and laboratory. The design of this score included a narrative review and a Delphi process\(^5\).

Even though the COVID-19 Severity Index has good discrimination, it has a short-scale validation and was performed during the first wave of the pandemic. Therefore, we propose a large multicenter cohort validation of this EWS and a comparison with the NEWS2\(^2\) and NEWS-C\(^3\) scores.

METHODS

We carried out a retrospective multi-center study at the Hospital Italiano de Buenos Aires (HIBA) and the Hospital Italiano Agustin Rocca of San Justo, from March 2020 to October...
2020. Data was collected using the Electronic Health Record (EHR) in both centers. The HIBA is a high-complexity third-level university hospital located at the Ciudad Autónoma de Buenos Aires, Argentina, with 750 beds and 38 critical care beds for adult patients. The Hospital Italiano de San Justo is a high-complexity hospital with 125 beds and 14 critical care beds for adult patients.

We included all adult patients who were hospitalized on general medical wards with Sars-Cov2 confirmed infection by polymerase chain reaction (PCR) of a nasopharyngeal swab. Exclusion criteria included patients with limitation of care orders, such as non-ICU transfer or non-intubation orders, and patients without information on demographic data, comorbidities, and vital signs. We admitted patients without laboratory samples and Chest X-ray assuming that the lack of data was related to the lack of clinical suspicion. Therefore, we considered those missing values as “normal”.

Data on demographics, comorbidities, and clinical condition was obtained for each patient. The NEWS2 score includes the following parameters: respiratory rate, hypercapnic respiratory failure, supplemental oxygen, body temperature, systolic blood pressure, heart rate, and level of consciousness; the combination of the score obtained for each variable provides a final score that ranges between 0 and 20². NEWS-C score includes the same parameters, adding age as an extra dichotomous variable. The combination of these values provides a final score between 0 and 23³. Finally, the COVID-19 Severity Index adds to NEWS2 the following variables: male gender, bilateral infiltrates in chest x-ray, presence of dyspnea, comorbidities (heart failure, chronic obstructive pulmonary disease, and diabetes with end-organ damage) and laboratory test results (dimer-d, leucocytes, and platelets). The combination of these variables provides a score between 0 and 33⁵. Details related to the EWS’s scoring system are presented in table S1 of the Supplementary material. For the COVID-19 Severity Index, we used the last chest x-ray
and laboratory results from the last two days. The chest x-ray was evaluated by the medical team.

The main outcome was ICU admission for each patient’s hospitalization, classifying each of them as “ICU admitted” or “non-ICU admitted” according to the EHR. The three scores were measured between 8 a.m. and 10 a.m. each day during hospitalization on general medical wards.

Data was summarised as mean with standard deviation (SD) or median with interquartile range (IQR). Differences between groups were determined using Mann–Whitney test or t-test as appropriate, and differences between categorical variables were evaluated with the chi-squared test or Fisher test as appropriate.

Discrimination of NEWS2 score, NEWS-C, and COVID-19 Severity Index was evaluated by receiver-operating characteristic (ROC) curve analysis, describing areas under curves (AUROC) with a 95% confidence interval (CI) and comparing them with Hanley–McNeil test, applying Bonferroni’s correction for multiple comparisons. We compared data 24 hours prior to ICU admission (for the group of patients who required ICU transfer), with all the scores measured in the group of non-ICU patients.

To value the calibration of the scores, we presented the comparison of predicted and observed probabilities. We calculated the predicted probability of ICU transfer for the three scores, then we divided the cohort into octiles and calculated the average probability of ICU transfer for each octile. Finally, we calculated the average probability of ICU transfer predicted by the scores and the observed probability (proportion of patients transferred to ICU) within each octile stratum. We evaluated the sensitivity, specificity, positive and negative likelihood ratios for all the values of each score. The optimal threshold was chosen based on the Youden index. Statistical analysis was performed using STATA v.16 software. This was a retrospective
observational study, thus informed consent was waived. The institutional ethics committee approved this study under protocol number 5562 in March 2020.

RESULTS

During the study period, 1570 patients were hospitalized on general medical wards with SARS-CoV-2 confirmed infection in the Hospital Italiano de Buenos Aires and Hospital Italiano de San Justo. 240 patients were excluded due to not resuscitate order and 12 due to missing data. Therefore, we finally included 1318 patients (Figure 1). Among these, 126 (9.5%) were transferred to ICU and 43 (34%) patients died in the ICU. These patients had a total of 13,768 days of hospitalizations. Patients admitted to the ICU were significantly older, most of them male and had higher body mass index. Regarding comorbidities, no significant difference was found in either group. The NEWS2 score was similar both in ICU and non-ICU admitted patients, but NEWS-C and COVID-19 Severity Index were significantly higher in the first group. Patients' characteristics are presented in Table 1.

Twenty-four hours before ICU admission we observed a reduction in the oxygen saturation (Sp02) with a higher proportion of patients requiring supplemental oxygen, compared with non-ICU patients on the second day of hospitalization. Characteristics of patients 24 hours prior to ICU transfer are presented in table 2.

Twenty-four hours before ICU admission the AUROC for NEWS2 score to predict ICU transfer was 0.73 (95%CI 0.68 - 0.78), the AUROC for NEWS-C was 0.74 (95%CI 0.70 - 0.79) and the AUROC for COVID-19 Severity Index was 0.80 (95%CI 0.77 - 0.84) (Figure 2). Values of AUROCs for predicting ICU admission 48 hrs before ICU transfer and at hospital admission had lower discrimination capacity (shown in Table 3). Calibration for each of the three models is shown in Figure 3. The evolution of the scores before ICU transfer are presented in the supplementary material (Figure S1)
The Hanley–McNeil test with Bonferroni’s correction showed that the AUROC for NEWS2 score had no significant difference with the NEWS-C score’s AUROC (p=0.08), but it was significantly better for COVID-19 Severity Index compared to NEWS2 score (p<0.001) and NEWS-C score (P=0.007). Forty-eight hours before ICU transfer, the COVID-19 severity index also had the best AUROC (Table 3).

According to the Youden index, the optimal threshold for NEWS2 and NEWS-C was 3 (sensitivity of 48% and specificity of 83.39% for NEWS2 and sensitivity of 55% and specificity of 75% for NEWS-C) and 5 for the COVID-19 Severity Index (sensitivity of 69% and specificity of 67%). The sensitivity, specificity, and likelihood ratios of each score are presented in Table S2 of the Supplementary material. Additionally, for each of the scores we present the probability of ICU transfer (Figure S2) depending on the scores and the Kernel density population distribution in ICU and non ICU admitted patients (Figure S3).

**DISCUSSION**

In this study, we observed that the COVID-19 Severity index has a better calibration and discrimination than NEWS2 and NEWS-C to predict ICU transfer in the following 24 hours. When we evaluated these scores at hospital admission, we found a poor discrimination capacity to identify those patients at risk of clinical deterioration.

Previous studies that have evaluated EWS in COVID-19 patients at hospital admission, observed moderate discrimination in a short term but poor-to-moderate discrimination capacity for predicting medium-term outcome. Carr et al. observed that NEWS-C score had moderate discrimination for short term outcome (3-day ICU transfer/death), and poor-to-moderate discrimination for the medium-term outcome (14-day ICU transfer/death). Martín-Rodríguez et al. observed that the NEWS2 at hospital admission was the scale with the best capacity to predict early mortality (less than 48 hrs after hospital admission). Su et al. found
an AUROC of 0.89 (IC95% 0.82–0.96) for NEWS-C score predicting deterioration of the respiratory function and need of intensive respiratory support. Socio et al. and Hai Hu et al. reported an AUROC of 0.87 and 0.81 respectively for NEWS2 score to predict death at hospital admission. However, the patients in these studies had a NEWS 2 score of 7, 6, and 7 respectively, and were transferred to ICU in the next 24 or 48 hours. Therefore, NEWS2, NEWS-C, and COVID-19 Severity Index appear to be good models for predicting clinical deterioration only for the next 24 to 48 hours.

In the evaluation of the included variables in the COVID-19 Severity Index we observed that age and gender improved discrimination, however comorbidities did not improve the prediction of the score. These observations are interesting and have been reported by other authors. This may indicate that the effect of pre-existing health conditions could be manifest in other variables of the score such as laboratory results or physiological markers. By analyzing vital signs and laboratory tests results 24 hours prior to ICU admission, we observed a significant reduction in Sp02 and need of supplemental oxygen. Other vital signs did not present relevant alterations and most of them did not show failure of other organs’ function other than the respiratory. Regarding laboratory results, we observed that lymphocyte count showed a significant reduction prior to ICU admission, as observed in other studies in which lymphopenia on admission was associated with poor outcome. Other scores applied to COVID-19 patients add blood parameters to improve discrimination.

Our study provides a validation of COVID-19 Severity Index, and a comparison between NEWS2 and NEWS-C scores in a large multicenter cohort. This is one of the largest cohorts of COVID-19 patients with the evaluation of EWS at hospital admission and 24 hours before ICU admission, with a description of sensitivity, specificity and likelihood ratios of all the possible scores for the three EWS.
This study has some limitations. First, although both hospitals have different target populations and different health care teams, both are private and belong to the same institution. Therefore, it is assumed that care characteristics could be similar. This reduces the external validation of our study. Second, not all patients had chest X-ray and laboratory tests. We admitted patients without these studies assuming that the lack of data was related to the lack of clinical suspicion. We considered those missing values as “normal”. Third, we excluded patients with non-intubate orders, thus all critical patients in this cohort were transferred to ICU, and we did not have unexpected deaths in general medical wards hospitalized patients. Therefore we had 0% mortality in non-ICU patients. The non-intubate order decision may change in different communities, thus reducing the external validity of our work.

CONCLUSIONS
The NEWS2 has been universally used, but there is limited data about its use in COVID-19 patients. The present work is one of the largest studies of EWS in COVID-19 patients, making a comparison between different scores and showing that NEWS2, NEWS-C and the COVID-19 Severity Index at hospital admission have poor-to-moderate discrimination capacity to determine ICU admission during hospitalization. We proved that the COVID-19 Severity Index further improves discrimination of NEWS2 and NEWS-C 24 and 48 hours before ICU admission. This novel tool could be useful for hospital administration, and patient risk stratification, but further validation is needed.
REFERENCES

1. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PWB. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. Resuscitation. 2014;85:587-94.

2. National Early Warning Score (NEWS) 2, NEWS2 and deterioration in COVID-19, https://www.rcplondon.ac.uk/news/news2-and-deterioration-covid-19; 2021 [consultada el 5 de noviembre de 2021].

3. Liao X, Wang B, Kang Y. Novel coronavirus infection during the 2019–2020 epidemic: preparing intensive care units—the experience in Sichuan Province, China. Intensive Care Medicine. 2020;46:357-60.

4. Su Y, Ju M-J, Xie R-C, Yu S-J, Zheng J-L, Ma G-G, et al. Prognostic Accuracy of Early Warning Scores for Clinical Deterioration in Patients With COVID-19. Front Med. 2020;7:624255. https://doi.org/10.3389/fmed.2020.624255

5. Huespe I, Carboni Bisso I, Di Stefano S, Terrasa S, Gemelli NA, Las Heras M. COVID-19 Severity Index: A predictive score for hospitalized patients. Medicina Intensiva. 2020.

6. Nishi H, Oishi N, Ishii A, Ono I, Ogura T, Sunohara T, et al. Predicting Clinical Outcomes of Large Vessel Occlusion Before Mechanical Thrombectomy Using Machine Learning. Stroke. 2019;50:2379-88.

7. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3:32-5.

8. Carr E, Bendayan R, Bean D, Stammers M, Wang W, Zhang H, et al. Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: a multi-hospital study. BMC Med. 2021;19:1-16.

9. Kostakis I, Smith GB, Prytherch D, Meredith P, Price C, Chauhan A, et al. The performance of the National Early Warning Score and National Early Warning Score 2 in hospitalised patients infected by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Resuscitation. 2021;159:150-7.

10. Sze S, Pan D, Williams CML, Wong N, Sahota A, Bell D, et al. Letter to the Editor: Variability but not admission or trends in NEWS2 score predicts clinical outcome in elderly hospitalised patients with COVID-19. J Infect. 2021;82:159-98.

11. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Intern Med. 2020;180:1081-9.

12. Martín-Rodríguez F, Martín-Conty JL, Sanz-García A, Rodríguez VC, Rabbione GO, Cebrián Ruiz I, et al. Early Warning Scores in Patients with Suspected COVID-19 Infection in Emergency Departments. J Pers Med. 2021;11.

13. Socio GVD, De Socio GV, Gidari A, Sicari F, Palumbo M, Franciscì D. National Early Warning Score 2 (NEWS2) better predicts critical Coronavirus Disease 2019 (COVID-19) illness than COVID-GRAM, a multi-centre study. Infection. 2021.

14. Hu H, Yao N, Qiu Y. Predictive Value of 5 Early Warning Scores for Critical COVID-19 Patients. Disaster Med Public Health Prep. 2020;1-8.
15. Costa VO, Nicolini EM, da Costa BMA, Teixeira FM, Ferreira JP, Moura MA, et al. Evaluation of the Risk of Clinical Deterioration among Inpatients with COVID-19. Adv Virol. 2021;2021:6689669.

16. Huang I, Pranata R. Lymphopenia in severe coronavirus disease-2019 (COVID-19): systematic review and meta-analysis. J Intensive Care Med. 2020;8:36.
## TABLES AND FIGURES

### Table 1. Baseline characteristics of patients at hospital admission.

| Patients characteristics | All patients (n=1318) | ICU (n=126) | non-ICU (n=1192) | p values |
|--------------------------|-----------------------|-------------|-------------------|----------|
| **Patients characteristics** |                       |             |                   |          |
| Age - median [IQR]       | 59 [34]               | 68 [17]     | 57 [35]           | <0.001   |
| Female sex - n (%)       | 646 (49.0)            | 46 (36.5)   | 600 (50.3)        | 0.004    |
| Body mass index - median [IQR] | 27.5 [6.9]          | 27.9 [5.7]  | 27.4 [6.9]        | <0.001   |
| **Comorbidities**        |                       |             |                   |          |
| Obesity - n (%)          | 333 (25.3)            | 46 (36.5)   | 287 (24.1)        | <0.003   |
| Diabetes - n (%)         | 155 (11.8)            | 19 (15.1)   | 136 (11.4)        | 0.284    |
| COPD - n (%)             | 63 (4.8)              | 10 (7.9)    | 53 (4.4)          | 0.127    |
| Congestive heart failure - n (%) | 62 (4.7)            | 5 (4.0)     | 57 (4.8)          | 0.661    |
| **Vital signs on admission** |                   |             |                   |          |
| Respiratory rate - mean (SD) | 18.6 (6.50)      | 18.6 (6.82) | 18.5 (2.77)       | 0.762    |
| Heart rate - mean (SD)   | 83.6 (13.2)           | 83.5 (13.2) | 84.8 (13.6)       | 0.386    |
| SpO2 - mean (SD)         | 96.6 (2.24)           | 95.9 (2.24) | 96.7 (2.22)       | 0.003    |
| Systolic blood pressure - mean (SD) | 124 (16.0)    | 122 (13.3)  | 124 (16.2)        | 0.21     |
| Temperature - mean (SD)  | 36.6 (1.73)           | 36.7 (0.87) | 36.6 (1.80)       | 0.289    |
| Supplementary oxygen therapy - n (%) | 230 (17.4)   | 63 (50)     | 167 (14)          | <0.001   |
| **Laboratory results and Chest X-ray** |               |             |                   |          |
| Lymphocytes (mm3)\(^a\) - mean (SD) | 1090 (659)       | 1136 (671)  | 837 (535)         | 0.035    |
| D-dimer (mg/ml)\(^b\) - mean (SD) | 861 (905)        | 810 (807)   | 1259 (1424)       | 0.193    |
| Platelets (mm3)\(^c\) - mean (SD) | 191789 (78941)   | 191917 (76086) | 191183 (93206) | 0.972    |
| Bilateral Chest X-ray infiltration\(^d\) - n (%) | 94 (47.5)       | 39 (79.6)   | 55 (36.9)         | <0.001   |
| **EWS on admission - median [IQR]** |                   |             |                   |          |
| NEWS2                    | 0 [1]                | 1 [2]       | 0 [1]             | 0.12     |
| NEWS-C                   | 1 [2]                | 1 [1]       | 1 [2]             | <0.001   |
| COVID-19 Severity Index  | 3 [2]                | 4 [2]       | 3 [2]             | <0.001   |
| **Outcomes**             |                       |             |                   |          |
| Days until ICU transfer - median [IQR] | 3 [3]             |             | 3 [3]             |          |
| Length of hospital stay - median [IQR] | 7 [8]             | 6.50 [6.2]  | 24.5 [24]         | <0.001   |
| In-hospital mortality - n (%) | 43 (3.3)         | 43 (34.1)   | 0 (0)             |          |
130 patients had lymphocytes count at hospital admission. 165 patients had D-dimer dosage at hospital admission. 132 patients had platelets count at hospital admission. 198 patients had chest X-ray at hospital admission. IQR: interquartile range. SD: standard deviation.
Table 2. Characteristics of patients 24 hours before ICU admission. In non-ICU patient’s data was compared on the second day of hospitalization.

| Vital signs on admission | All patients (n=1318) | ICU (n=126) | non-ICU (n=1192) | p values |
|--------------------------|-----------------------|-------------|-------------------|----------|
| Respiratory rate - mean (SD) | 18.6 (6.39) | 18.7 (2.78) | 18.6 (6.82) | 0.828 |
| Heart rate - mean (SD) | 83.3 (13.2) | 82.4 (13.2) | 83.5 (13.2) | 0.416 |
| SpO2 - mean (SD) | 96.2 (2.55) | 95.0 (2.3) | 96.3 (2.54) | <0.001 |
| Systolic blood pressure - mean (SD) | 120 (12.2) | 123 (13) | 120 (12.1) | 0.027 |
| Temperature - mean (SD) | 36.4 (0.72) | 36.7 (0.89) | 36.4 (0.69) | <0.001 |
| Supplementary oxygen therapy - n (%) | 257 (19.5) | 78 (61.9) | 179 (15) | <0.001 |

| Laboratory results | All patients (n=1318) | ICU (n=126) | non-ICU (n=1192) | p values |
|-------------------|-----------------------|-------------|-------------------|----------|
| Lymphocytes (mm3)\(^a\) - mean (SD) | 1163 (844) | 804 (389) | 1228 (888) | <0.001 |
| D- dimer (mg/ml)\(^b\) - mean (SD) | 1042 (1202) | 1292 (1170) | 968 (1206) | 0.172 |
| Platelets (mm3)\(^c\) - mean (SD) | 193812 (81383) | 190627 (90356) | 194399 (79759) | 0.765 |

\(^a\)60 patients had lymphocytes before ICU admission and 327 non-ICU patients. \(^b\)33 patients had D- dimer before ICU admission and 112 non-ICU patients. \(^c\)59 patients had platelets count before ICU admission and 320 non ICU patients. SD: standard deviation.
Table 3. AUROCs and 95% Confidence intervals of the NEWS2, NEWS-C, and COVID-19 Severity Index 24 and 48 hours prior ICU transfer.

| AUROC                      | 24 hs pre ICU transfer | 48 hs pre ICU transfer | Hospital Admission |
|----------------------------|------------------------|------------------------|--------------------|
| NEWS2                      | 0.73 (95%CI 0.68 - 0.78) | 0.71 (95%CI 0.65 - 0.77) | 0.52 (95%CI 0.47 - 0.57) |
| NEWS-C                     | 0.74 (95%CI 0.70 - 0.79) | 0.72 (95%CI 0.68 - 0.79) | 0.56 (95%CI 0.52 - 0.61) |
| COVID-19 Severity Index    | 0.80 (95%CI 0.77 - 0.84) | 0.79 (95%CI 0.75 - 0.84) | 0.61 (95%CI 0.58 - 0.66) |

AUROCs: Area Under the ROCs curves.
**Figure 1.** Patient selection flow chart

- 1,570 patients hospitalized on general medical wards with COVID-19
  - 240 (15%) excluded due to do not expect to live order
  - 12 (1%) excluded due to missing data

1,318 (84%) COVID-19 hospitalizations included

- 126 (9.5%) transfer to ICU during hospitalization
- 1,192 (90.5%) general hospitalizations

**Figure 2.** AUROC of the NEWS2, NEWS-C. and COVID-19 Severity Index for ICU transfer in the next 24 hs.

**Figure 3.** Observed and predicted probabilities for each decile of predicted values for the three scores.
Figure S1. Longitudinal trend in NEWS2 score in ICU admitted patients prior to ICU transfer. Solid line shows fitted trend (loess regression), grey shading depicts 95% confidence intervals.

Figure S2. Probability of ICU transfer in the next 24 hours depending on the Score in the NEWS2, NEWS-C. and COVID-19 Severity Index. Solid line shows fitted trend (loess regression), grey shading depicts 95% confidence intervals.

Figure S3. Kernel density distributions of NEWS2, NEWS-C and COVID-19 Severity Index scores in patients non transferred to ICU (blue) and ICU transferred (red) patients (24 hours before ICU transfer). The vertical axis line represents the threshold proposed by the scores (4, 6 and 7 for NEWS2 and NEWS-C., and 2, 6 and 8 for COVID-19 Severity Index).