Background: We aimed to identify risk factors causing critical disease in hospitalized children with COVID-19 and to build a predictive model to anticipate the probability of need for critical care.

Methods: We conducted a multicenter, prospective study of children with SARS-CoV-2 infection in 52 Spanish hospitals. The primary outcome was the need for critical care. We used a multivariable Bayesian model to estimate the probability of needing critical care.

Results: The study enrolled 350 children from March 12, 2020, to July 1, 2020: 292 (83.4%) and 214 (73.7%) were considered to have relevant COVID-19, of whom 24.2% required critical care. Four major clinical syndromes of decreasing severity were identified: multi-inflammatory syndrome (MIS-C) (17.3%), bronchopulmonary (51.4%), gastrointestinal (11.6%), and mild syndrome (19.6%). Main risk factors were high C-reactive protein and creatinine concentration, lymphopenia, low platelets, anemia, tachycardia, age, neutrophilia, leukocytosis, and low oxygen saturation. These risk factors increased the risk of critical disease depending on the syndrome: the more severe the syndrome, the more the risk factors conferred. Based on our findings, we developed an online risk prediction tool (https://rserver.h12o.es/pediatria/EPICOAPP/, username: user, password: 9000).

Conclusions: Risk factors for severe COVID-19 include inflammation, cytopenia, age, comorbidities, and organ dysfunction. The more severe the syndrome, the more the risk factors increase the risk of critical illness. Risk of severe disease can be predicted with a Bayesian model.

Key Words: COVID-19, SARS-CoV-2, children, syndrome, Bayesian

Recent data suggest that children are less susceptible to SARS-CoV-2 infection than adults, and their symptoms are usually milder. However, it remains unclear how to identify early the patients that will have a severe disease. The heterogeneity of clinical presentation suggests that risk factors may vary depending on syndromic presentation.

There is a need to find potential predictors of severity in pediatric COVID-19 cases to stratify which patients may benefit from treatments. No models exist to predict severe disease for children with COVID-19.

This study aimed to identify risk factors associated with severe COVID-19 and to build a predictive model to anticipate the probability of need for critical care.

METHODS

Design

The Epidemiological Study of coronavirus in Children (EPICO-AEP) is a multicenter cohort study conducted in Spain to assess the characteristics of children with COVID-19. In total, 52 hospitals collected data from the beginning of the epidemic in Spain—February 25—until this analysis. The study was approved

The Pediatric Infectious Disease Journal • Volume 40, Number 8, August 2021

www.pidj.com | e287

Copyright © 2021 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.
by the Ethics Committee of the Hospital 12 de Octubre, Madrid (code 20/101), and other participating hospitals. Participants were enrolled after signed or verbal consent from parents/guardians and by the consent of patients older than 12 years.

Eligible participants were children 0–18 years of age attended in any of the hospitals of the network from March 12, 2020, to July 1, 2020, with a SARS-CoV-2 infection confirmed by real-time polymerase chain reaction (RT-PCR) or children fulfilling WHO criteria for MIS-C.11

**Laboratory Methods**

Respiratory samples were obtained from nasopharyngeal swabs and tracheal or bronchial aspirates when available. Serum samples were analyzed in local clinical microbiology laboratories using commercial kits.

**Definitions**

For analysis purposes, diagnoses were categorized into 4 syndromes: “MIS-C,” “bronchopulmonary syndrome” (including pneumonia, bronchiolitis, bronchitis, and asthma flare), “gastrointestinal syndrome” (including gastroenteritis and abdominal pain), and “mild syndrome” (including FWS, URTI, flu-like syndrome, and asymptomatic patients).

The primary outcome was need for critical care, defined as the combined outcome of admission into a pediatric intensive care unit, and need for respiratory support with high-flow oxygen, continuous positive airway pressure or mechanical ventilation. To differentiate patients admitted for COVID-19 from those admitted for other reason but with mild/asymptomatic community-acquired or nosocomial SARS-CoV-2 infection, the term “relevant COVID-19 disease” (r-COVID-19) was created. This was defined as admission due to bronchopulmonary syndrome, MIS-C, gastrointestinal syndrome, or mild syndrome with an associated diagnosis that might be considered a complication of COVID-19 and conditioned hospitalization—for example, febrile seizures.

**Data Management and Statistical Analyses**

Researchers from each participating hospital collected pseudo-anonymized data using a standardized clinical research form on the electronic data capture system REDCap.13

Statistical analyses were performed using the R language. Continuous variables including heart rate, respiratory rate, and blood pressure were categorized according to normal values for age.14 To dichotomize the continuous variables without a standardized categorization, such as platelets and oxygen saturation (SatO2), optimal cutoff points were assessed using generalized additive models implemented in the cutpointr R package.15 Each optimal cutoff point was specified in the descriptive tables and analysis. For univariable analysis, the posterior probability of a positive correlation was calculated for summary tables using Bayesian univariable logistic regression.

To build a classification model that predicts the probability for critical care, a Naive Bayes algorithm was made and implemented in a web app. All the variables with more than 80% probability of conferring risk in the univariable model and with <15% of missing values were included. Symptoms and signs that define syndromes (for instance, conjunctivitis, shock) were excluded as possible predictors because they were already included in the syndromes’ definitions (eg, MIS-C). All the included participants were randomly divided into a training dataset that was used to generate the models (70% of the original dataset), and a validation dataset used to assess the performance of the models (remaining 30%). Partitions were balanced by the outcome class, which yielded a training set (n = 151) and a validation set (n = 63). All missing values were imputed independently for both datasets.

The algorithm was trained using 5-fold cross-validation as a resampling control method to prevent overfitting. Due to the disparity in the frequencies of the observed classes, ROSE downsampling hybrid method was used. The variable importance was calculated to determine the predictors that significantly affect the algorithm classification output. A confusion matrix was built for each of the models to assess their accuracy, sensitivity, and specificity. The area under the curve of the receiver operating characteristic curves was determined for each model.

To estimate the increased probability of needing critical care according to the syndrome for each risk factor, Bayesian multivariable models were employed. Different models were employed including each risk factor, previously identified as important in the prediction Naive Bayes model, and the syndrome. Probabilities for each condition were plotted according to each condition using the ggplot2 package.16

For Bayesian analyses, Student’s t distribution with mean zero and 7 degrees of freedom was used as the weakly informative prior. All models were run with 4 Markov chains for 1000 warm-up and 100,000 sampling iterations. All models were programmed using the stan_glm function of R rstanarm package.17

**RESULTS**

**Features of the Cohort**

A total of 350 children were enrolled (Figure 1). The median age was 5.5 years (interquartile range [IQR], 0.55–12.1), 129/350 (36.9%) were ≤2 years old and 191/350 (54.6%) were male.

**Features of Patients With r-COVID-19**

A total of 214 (73.7% of 292 hospitalized) patients were considered to have r-COVID-19.

Of the 214 participants with r-COVID-19, 93/214 (45.1%) had comorbidities. Regarding major clinical syndromes, 110/214 (51.4%) had bronchopulmonary syndrome (100/214 [46.7%] pneumonia), 42/214 (19.6%) had mild syndrome, 37/214 (17.3%) MIS-C and 25/214 (11.6%) gastrointestinal syndrome.

Of the 214 patients with r-COVID-19, 52 (24.2%) required critical care during a median of 5 (IQR 3.0–8.0) days.

Clinical characteristics of patients with r-COVID-19 who needed critical care are summarized in Table, Supplemental Digital Content 1, http://links.lww.com/INF/E403. The symptoms and other features most likely to be associated with requiring critical care in the uncomfortable model are displayed in Figure 2 and Table, Supplemental Digital Content 1, http://links.lww.com/INF/E403.

**Critical Care Predictive Model**

To build a comprehensive predictive model, a Naive Bayes algorithm was trained and validated to predict critical care necessity. The predictors with higher relative importance to predict the necessity of critical care were high CRP, lymphopenia, platelets below 220,000/mm³, anemia, tachycardia, age, neutrophilia, leukocytosis, high creatinine, low oxygen saturation, fever, days of fever, high weight percentile, MIS-C, comorbidities, gastrointestinal syndrome, and bronchopulmonary syndrome. In the external validation, overall accuracy for the naïve Bayes classifier was 84% (95% confidence interval [CI]: 72.7–92.1). The model could predict the need for critical care with 80% (95% CI: 53–95.6) sensitivity and 85.4% (95% CI: 72.2–93) specificity; the positive predictive value was estimated as 63.2% (95% CI: 45.2–78.1) and the negative predictive value was 93.2% (95% CI: 83.15–97.4). The area under the curve was 76.6% (95% CI: 70.3–80.9).
In the validation set (n = 63), we compared the probability of critical care attributed by the model with the population of patients who actually needed and did not need critical care (Figure 3). There was a significant difference in the distribution of both groups, revealing the high classification ability. This model was implemented in the app (https://rserver.h12o.es/EPICOAPP/, Username: user, password: 0000).

Differences in the Effect of Risk Factors According to Syndrome

Patients diagnosed with MIS-C (density plot, Figure 4, greenish color) had the highest probability of needing critical care, followed by bronchopulmonary syndrome (bluish) and gastrointestinal syndrome (yellowish) as compared with mild syndrome. The different risk factors had different effects across the 4 syndromes.

The principal risk factors for critical care in patients with MIS-C were platelets <220,000/mm³ (31%), presence of comorbidities (26%) and lymphopenia (25%). In patients with bronchopulmonary syndrome, the 3 most important risk factors were the same; however, these factors conferred less risk than in MIS-C: 24%, 20%, and 19%, respectively. Likewise, the 3 factors conferred some risk for gastrointestinal patients (10% each) but less than in MIS-C and bronchopulmonary syndrome.

Specifically, low platelets conferred 7% more risk of critical care in the MIS-C group than in the bronchopulmonary syndrome.
FIGURE 2. Clinical features of the patients at presentation and probability of being associated with needing critical care. The posterior probability of $\beta > 0$, that is, the probability of each feature to have a positive correlation with critical care, is displayed (from 0 to 1).

FIGURE 3. Naïve Bayes predictor selection and model performance. The model was built with a set of 70% of patients and validated with a set of 30% of patients. (A) represents the relative variable importance from all the predictors included in the model. B: The patients included in the validation set (n = 63) are displayed. The Y axis represents the distribution of the predicted probability for critical care given by the model. In the X axis, patients are separated according to their actual necessity of critical care. The probability of critical care attributed by the model was significantly different ($P = 0.000002$) in the population of patients who actually needed and in the population that did not need critical care, showing the high classification ability. $P$ value was calculated using the Mann-Whitney $U$ test. This model was implemented in an app (https://rserver.h12o.es/EPICOAPP/, Username: user, password: 0000).
group and 21% more risk than in the gastrointestinal syndrome group. Likewise, the presence of comorbidities conferred 6% more risk of critical care in the MIS-C group than in the bronchopulmonary syndrome group and 16% more risk than in gastrointestinal syndrome group. Lymphopenia conferred 6% more risk of critical care in the MIS-C group than in the bronchopulmonary syndrome group and 15% more risk than in the gastrointestinal syndrome group.

By contrast, fever, oxygen saturation or a weight percentile >90 did not confer substantially different risk among the different syndromes.

**DISCUSSION**

In this study, we identified similar risk factors for critical disease as other studies. We added several new factors and the syndrome category as a specific risk factor. Remarkably, we show that most different risk factors increase the risk for critical care differently depending on the syndrome of the patient: the more severe the syndrome, the more risk the factor confers.

Some risk factors are patient-dependent, such as age and comorbidities. Other identified risk factors suggest immune dysregulation and severe inflammation in critical patients. In the predictive model, we could not use some promising biomarkers such as D-dimer, interleukin-6, or proBNP because they were not consistently measured in patients with mild disease, but those biomarkers were indeed significantly higher in patients needing critical care. High inflammatory markers as CRP or blood cell disorders, such as leukocytosis, neutrophilia, anemia, lymphopenia, and thrombopenia, were found in severe cases. In our analysis, we found that the best clinical cutoff point for platelets was 220,000/mm$^3$ instead of 150,000/mm$^3$, which is classically used for thrombopenia. The cytopenia found in severe cases suggests either damage to bone marrow or peripheral cells or migration of activated cells to tissues.

We created a novel predictive model to anticipate the probability of critical care. Early recognition of the need for critical care is relevant for starting early treatments. Through a rapid, inexpensive, and comprehensive web app, the attending physician can introduce the patient’s data at admission and the risk of severe disease can be obtained. The evaluation of the algorithm showed significant accuracy and sensitivity. To our knowledge, this is the first model with an online app to help and recognize the need for critical care in children with COVID-19.

**LIMITATIONS**

This study included children who attended in different hospitals, and we focused on those with r-COVID-19. There is a risk of selection, case identification and reporting bias. Access to SARS-CoV-2 testing was not consistent during the enrollment.
The diversity and broadness of the study are, at the same time, strengths, as they provide insight into the disease in a major clinical part of Spain through a prospective collection of data. We used a case record form with several fields shared with other international registries (ISARIC), enabling sharing, but we tailored it specifically for pediatric data collection.14,20

Although viral-bacterial coinfection was more frequent in hospitalized children, a full workup for coinfections was not done uniformly, and thus the role of coinfections is not completely clear. The study syndicated few neonates because most neonates with COVID-19 in Spain were included in a different neonatal registry.

The ethnic origin was not recorded, so we cannot compare our study with other studies suggesting worse outcomes in minorities. Interestingly, some of the factors, such as comorbidity that increased the risk significantly depending on the syndrome, had low relative importance in the model. As this artificial intelligence model is a black box, we cannot assess why this occurred.

This model was built with hospitalized children with r-COVID-19 and should not be applied to outpatients. Finally, the risk predicted by models reflects those of patients receiving care only. The prediction models should be updated regularly because the dynamics of the disease and management strategies may change.21

CONCLUSIONS

Risk factors for severe COVID-19 include inflammation, cytopenia, age, comorbidities, and organ dysfunction. The more severe the syndrome, the more the risk factor increases the risk of critical illness. Risk of severe disease can be predicted with a Bayesian model.

ACKNOWLEDGMENTS

We thank all the patients and families for their participation in the study and the laboratory staff and clinical staff members who cared for them. Thanks to Kenneth McCreath (Universidad Europea de Madrid) for the style and English revision.

EPICO-AEP Working Group: María de Ceano, Ana Méndez-Echevarría, Taulía Sainz, Clara Udaondo, Fernando Baquero (Hospital La Paz), Mar Santos, Marisa Navarro, Elena Rincón, Begoña Santiago, (Hospital Universitario Gregorio Marañón), Pablo Rojo, Daniel Báñez, Luis Prieto, Elisa Fernández-Cooke, David Torres-Fernández, Ángela Manzaneque, Jaime Carrasco, Elena Cobos-Carrasco (Hospital 12 de Octubre), Miguel Lasnapa (Hospital Sant Joan de Déu), Lourdes Calleja (Hospital Niño Jesús), María Espiau, Jacques G. Rivière (Hospital Universitari Vall d’Hebron), Mercedes Herranz (Complejo Hospitalario de Navarra), Fernando Cañabanas (Hospital Universitari Quirón salud Madrid), Rut del Valle, María Fernández, Teresa Raga, María de la Serna, Ane Plazaola, Juan Miguel Mesa (Hospital Infantia Sofia), María Dolores Martín (BR Salud), Enrique Oteo, José Luis Vázquez (Hospital Ramón y Cajal), Lola Falcón, Olaf Neth, Peter Olbrich, Walter Goicoechea (Hospital Universitario Virgen del Rocío), Laura Martin (Hospital Universitario Regional de Málaga), Lucia Figueroa (Hospital de Villalba), Maria Llorente (Hospital Universitario del Sureste), Maria Penin, Claudia Garcia, Maria Garcia, Teresa Alvaredo (Hospital Príncipe de Asturias), Mª Immaculada Olmedo, Agustín López (Hospital Puerta de Hierro), Elvira Cobo (Hospital Fundación Alcorcón), Mariam Toviz (Hospital del Tajo), Pilar Galán (Hospital Fundación Fuenlabrada), Beatriz Soto, Sara Guillén (Hospital de Getafe), Adriana Navas (Hospital Infantia Leonor) M. Luz García (Hospital de Leganés), Sara Pérez (Hospital de Torrejón), Amanda Bermejo, Pablo Mendoza (Hospital de Móstoles), Gema Sabrido (Hospital Rey Juan Carlos), María José Hernández (Hospital Central de la Defensa), Ana Belén Jiménez (Fundación Jiménez Díaz), Arantxa Berzosa, José Tomás Ramos (Hospital Clínico San Carlos), Ana López (Hospital Universitari Son Espases), Beatriz Ruiz (Hospital Universitari Reina Sofia), Santiago Alfayate, Ana Menasalvas, Elvira Cervantes (Hospital Clínico Universitario Virgen de la Arrixaca), Maria Méndez (Instituto de Investigación en Ciencias de la Salud Germans Trias y Pujol), Ángela Hurtado (Instituto Hospitalexpa de Pediatría), Cristina García, Inés Amich (Hospital de Pedro), Manuel Oltra, Álvaro Villarroya (Hospital Universitari i Politècnic La Fe), Angustias Ocaña (Hospital La Moraleja), Isabel Romero, Maria Fernanda Gazmán (Hospitales Madrid), M.J. Pascual (Hospital Nisa), Maria Sánchez-Códez (Hospital Universitario Puerta del Mar), Elena Montesinos (Consorti Hospital General Universitari de València), Julia Jensen, Maria Rodríguez (Hospital Infantana Cristina), Gloria Caro (Hospital Infantana Elena), Neus Rius, Alba Gómez (Hospital Universitari Sant Joan de Reus), Rafael Bretón (Hospital Clínico Universitario de València), Margarita Rodriguez, Julio Romero (Hospital Universitario Virgen de las Nieves), Ana Campos (Hospital Universitario Santas La Zarzuela), Mercedes Garcia (Hospital de Mérida), Rosa María Velasco (Complejo Hospitalario de Toledo), Zulema Lobato (Althaia, Xarxa Assistencial Universitaria de Manresa), Fernando Centeno, Elena Pérez (Hospital Universitario Rio Horta), Paula Vidal (Hospital Clínico Universitario Lozano Blesa), Corsino Rey, Ana Vivanco, Maruchi Alonso (Hospital Universitario Central de Asturias), Pedro Alcalá, Javier González de Dios (Hospital General Universitario de Alicante), Eduard Solé, Laura Minguell (Hospital Universitari Arnau de Vilanova), Iziar Astigarraga (Hospital Universitario de Cruescès), Mª Ángeles Vázquez, Miguel Sánchez (Hospital Universitario Torrecárdenas), Elena Diaz (Hospital Virgen de la Luz), Eduardo Consuegra (Hospital Universitario de Salamanca), Maria Cabanillas (Complejo Asistencial Universitario de Palencia), Luis Peña (Hospital Universitario Materno Infantil de las Palmas), Elisa Garrote, Maite Goicoechea (Hospital Universitario de Basurto), Irene Centelles (Hospital General Universitari de Castelló), Santiago Lapeña, Sara Gutiérrez, Soraya Gutiérrez (Complejo Asistencial Universitario de León), Amparo Cavalie (PIUS Hospital de Vallés), José María Olmos (Hospital Mare de Déu dels Lliris), Alejandro Cobos, Sara Diaz (Hospital Universitario de Canarias), Beatriz Jiménez (Hospital Universitario Marqués de Valdecilla), Raúl González (Hospital Sant Joan d’Alacant), Miguel Lafuent, Matilde Bustillo (Hospital Infantil de Zaragoza), Natividad Pons, Julia Morata (Hospital Lluís Alcanys), and Elsa Segura (Hospital Universitario Son Llatzer de Palma de Mallorca), Maria Bernardo (Universidad Europea de Madrid).

REFERENCES


