Predictors of Venous Thromboembolism in COVID-19 Patients: Results of the COVID-19 Brazilian Registry

Warley C. Silveira (warleysilveira@gmail.com)
Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil
https://orcid.org/0000-0003-2540-9083

Lucas Emanuel Ferreira Ramos
Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil
https://orcid.org/0000-0001-7844-0581

Rafael Tavares Silva
Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil
https://orcid.org/0000-0002-9270-5328

Bruno Barbosa Miranda de Paiva
Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil
https://orcid.org/0000-0001-8378-3650

Polianna Delfino Pereira
Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil
https://orcid.org/0000-0003-2406-6576

Alexandre Vargas Schwarzbold
Hospital Universitário de Santa Maria. Av. Roraima, 1000, prédio 22, Santa Maria, Brazil
https://orcid.org/0000-0002-5535-6288

Andresa Fontoura Garbini
Hospital Nossa Senhora da Conceição and Hospital Cristo Redentor. Av. Francisco Trein, 326, Porto Alegre, Brazil https://orcid.org/0000-0002-3467-2150

Bruna Schettino Morato Barreira
Pontifícia Universidade Católica de Minas Gerais. Rua do Rosário 1081, Betim, Brazil
https://orcid.org/0000-0003-4114-2577

Bruno Mateus de Castro
Hospital de Clínicas de Porto Alegre. Rua Ramiro Barcelos, 2350, Porto Alegre, Brazil
https://orcid.org/0000-0002-7013-2127

Carolina Marques Ramos
Hospital Julia Kubitschek. Av. Professor Alfredo Balena 190, Belo Horizonte, Brazil
https://orcid.org/0000-0002-8258-0891

Caroline Danubia Gomes

Hospital Mãe de Deus. Rua José de Alencar, 286, Porto Alegre, Brazil https://orcid.org/0000-0002-5254-4320

Christiane Corrêa Rodrigues Cimini

Mucuri Medical School – FAMMUC, Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM. Rua Cruzeiro, 01, Teófilo Otoni, Brazil https://orcid.org/0000-0002-1973-1343

Elayne Crestani Pereira

Hospital SOS Cárdio. Rodovia, SC-401, 121, Florianópolis, Brazil https://orcid.org/0000-0002-1178-669X

Eliane Würdig Roesch

Hospital de Clínicas de Porto Alegre. Rua Ramiro Barcelos, 2350, Porto Alegre, Brazil https://orcid.org/0000-0002-9507-6344

Emanuele Marianne Souza Kroger

Hospital Julia Kubitschek. Av. Professor Alfredo Balena 190, Belo Horizonte, Brazil https://orcid.org/0000-0001-5108-4337

Felipe Ferraz Martins Graça Aranha

Universidade do Sul de Santa Catarina. Av. Pedra Branca, 25, Palhoça, Brazil https://orcid.org/0000-0003-0429-7641

Fernando Anschau

Hospital Nossa Senhora da Conceição and Hospital Cristo Redentor. Av. Francisco Trein, 326, Porto Alegre, Brazil https://orcid.org/0000-0002-2657-5406

Fernando Antonio Botoni

Hospital Julia Kubitschek. Av. Professor Alfredo Balena 190, Belo Horizonte, Brazil https://orcid.org/0000-0001-6268-8507

Fernando Graça Aranha

Hospital SOS Cárdio. Rodovia, SC-401, 121, Florianópolis, Brazil https://orcid.org/0000-0001-9173-8892

Gabriela Petry Crestani

Hospital Mãe de Deus. Rua José de Alencar, 286, Porto Alegre, Brazil https://orcid.org/0000-0001-9668-1434

Giovanna Grunewald Vietta

Hospital SOS Cárdio. Rodovia, SC-401, 121, Florianópolis, Brazil https://orcid.org/0000-0002-0756-3098

Gisele Alsina Nader Bastos

Hospital Moinhos de Vento. Rua Ramiro Barcelos, 910, Porto Alegre, Brazil https://orcid.org/0000-0001-9668-1434

Jamille Hemétrio Salles Martins Costa

Hospital Márcio Cunha. Av. Kiyoshi Tsunawaki, 48, Ipatinga, Brazil https://orcid.org/0000-0002-7577-6060

Jéssica Rayane Corrêa Silva da Fonseca

Hospital Semper. Alameda Ezequiel Dias, 389. Belo Horizonte, Brazil https://orcid.org/0000-0002-6453-3741

Karen Brasil Ruschel

Institute for Health Technology Assessment (IATS/ CNPq). Rua Ramiro Barcelos, 2359, Prédio 21 | Sala 507, Porto Alegre, Brazil https://orcid.org/0000-0002-6362-1889

Leonardo Seixas de Oliveira

Hospital Santa Rosalia. Rua do Cruzeiro, 01, Teófilo Otoni, Brazil https://orcid.org/0000-0003-1575-6559

Lílian Santos Pinheiro

Hospital Santa Rosalia. Rua do Cruzeiro, 01, Teófilo Otoni, Brazil https://orcid.org/0000-0003-2957-7674

Liliane Souto Pacheco

Hospital Universitário de Santa Maria. Av. Roraima, 1000, prédio 22, Santa Maria, Brazil https://orcid.org/0000-0002-0229-0692

Luciana Borges Segala

Hospital Universitário de Santa Maria. Av. Roraima, 1000, prédio 22, Santa Maria, Brazil https://orcid.org/0000-0002-2628-5396

Luciana Siuves Ferreira Couto

Universidade Federal de Ouro Preto. Rua Diogo de Vasconcelos, 122, Ouro Preto, Brazil https://orcid.org/0000-0002-7948-222X

Luciane Kopittke

Hospital Nossa Senhora da Conceição and Hospital Cristo Redentor. Av. Francisco Trein, 326, Porto Alegre, Brazil https://orcid.org/0000-0002-6606-7756

Maiara Anschau Floriani

Hospital Moinhos de Vento. Rua Ramiro Barcelos, 910, Porto Alegre, Brazil https://orcid.org/0000-0002-2981-9445

Majlla Magalhães Silva

Hospital Moinhos de Vento. Rua Ramiro Barcelos, 910, Porto Alegre, Brazil https://orcid.org/0000-0001-8669-0214

Marcelo Carneiro

Hospital Santa Cruz. Rua Fernando Abott, 174, Santa Cruz do Sul, Brazil https://orcid.org/0000-0003-3603-1987

Maria Angélica Pires Ferreira

Hospital de Clínicas de Porto Alegre. Rua Ramiro Barcelos, 2350, Porto Alegre, Brazil https://orcid.org/0000-0003-0961-524X

Matheus Carvalho Alves Nogueira

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0002-0241-9046

Maria Auxiliadora Parreiras Martins

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0002-5211-411X

Marina Neves Zerbini de Faria
Hospital Julia Kubitschek. Av. Professor Alfredo Balena 190, Belo Horizonte, Brazil https://orcid.org/0000-0003-3903-2475

Milton Henriques Guimarães Júnior
Hospital Márcio Cunha. Av. Kiyoshi Tsunawaki, 48, Ipatinga, Brazil https://orcid.org/0000-0002-2127-8015

Natália da Cunha Severino Sampaio
Hospital Eduardo de Menezes. Rua Dr. Cristiano Rezende, 2213, Belo Horizonte, Brazil https://orcid.org/0000-0002-5799-8006

Neimy Ramos de Oliveira
Hospital Eduardo de Menezes. Rua Dr. Cristiano Rezende, 2213, Belo Horizonte, Brazil https://orcid.org/0000-0001-5408-9459

Nicole de Moraes Pertile
Hospital Moinhos de Vento. Rua Ramiro Barcelos, 910, Porto Alegre, Brazil https://orcid.org/0000-0001-8442-4350

Pedro Guido Soares Andrade
Hospital Semper. Alameda Ezequiel Dias, 389. Belo Horizonte, Brazil https://orcid.org/0000-0002-4511-5149

Pedro Ledic Assaf
0Hospital Metropolitano Doutor Célio de Castro. Rua Dona Luiza, 311, Belo Horizonte, Brazil https://orcid.org/0000-0002-8419-5719

Reginaldo Aparecido Valacio
Hospital Metropolitano Odilon Behrens. Rua Formiga, 50, Belo Horizonte, Brazil https://orcid.org/0000-0003-2666-7092

Rochele Mosmann Menezes
Hospital Santa Cruz. Rua Fernando Abott, 174, Santa Cruz do Sul, Brazil https://orcid.org/0000-0002-1548-1607

Saionara Cristina Francisco
0Hospital Metropolitano Doutor Célio de Castro. Rua Dona Luiza, 311, Belo Horizonte, Brazil https://orcid.org/0000-0002-9655-6294

Silvana Mangeon Mereilles Guimarães
Hospital Semper. Alameda Ezequiel Dias, 389. Belo Horizonte, Brazil https://orcid.org/0000-0002-9655-6294

Silvia Ferreira Araújo
Hospital Semper. Alameda Ezequiel Dias, 389. Belo Horizonte, Brazil https://orcid.org/0000-0003-4782-5440

Suely Meireles Rezende

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0002-3083-7093

Susany Anastácia Pereira

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0001-5039-443X

Tatiana Kurtz

Hospital Santa Cruz. Rua Fernando Abott, 174, Santa Cruz do Sul, Brazil https://orcid.org/0000-0003-2124-711X

Tatiani Oliveira Fereguetti

Hospital Eduardo de Menezes. Rua Dr. Cristiano Rezende, 2213, Belo Horizonte, Brazil https://orcid.org/0000-0001-5845-0715

Carísi Anne Polanczyk

Institute for Health Technology Assessment (IATS/ CNPq). Rua Ramiro Barcelos, 2359, Prédio 21 | Sala 507, Porto Alegre, Brazil https://orcid.org/0000-0002-2447-2577

Magda Carvalho Pires

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0003-3312-4002

Marcos André Gonçalves

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0002-2075-3363

Milena Soriano Marcolino

Universidade Federal de Minas Gerais. Av. Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil https://orcid.org/0000-0003-4278-3771

Research article

Keywords: Covid-19, Pulmonary embolism, Deep vein thrombosis, Risk Factors, Thromboprophylaxis

Posted Date: January 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1180396/v1

Version of Record: A version of this preprint was published at Internal and Emergency Medicine on June 1st, 2022. See the published version at https://doi.org/10.1007/s11739-022-03002-z.

Abstract

Background

Previous studies that assessed risk factors for venous thromboembolism (VTE) in COVID-19 patients have shown inconsistent results. Our aim was to investigate VTE predictors by both logistic regression (LR) and machine learning (ML) approaches, due to their potential complementarity.

Methods

This substudy of a large Brazilian COVID-19 Registry included COVID-19 adult patients from 16 hospitals. Symptomatic VTE was confirmed by objective imaging. LR analysis, tree-based boosting and bagging were used to investigate the association of variables upon hospital presentation with VTE.

Results

Among 4,120 patients (55·5% men, 39·3% critical patients), VTE was confirmed in 6·7%. In multivariate LR analysis, obesity (OR 1·50, 95%CI 1·11-2·02); being an ex-smoker (OR 1·44, 95%CI 1·03-2·01); surgery ≤ 90 days (OR 2·20, 95%CI 1·14-4·23); axillary temperature (OR 1·41, 95%CI 1·22-1·63); D-dimer ≥ 4 times above the upper limit of reference value (OR 2·16, 95%CI 1·26-3·67), lactate (OR 1·10, 95%CI 1·02-1·19), C-reactive protein levels (CRP, OR 1·09, 95% CI 1·01-1·18); and neutrophil count (OR 1·04, 95%CI 1·005-1·075) were independent predictors of VTE. Atrial fibrillation, peripheral oxygen saturation/inspired oxygen fraction (SF) ratio and prophylactic use of anticoagulants were protective. Temperature at admission, SF ratio, neutrophil count, D-dimer, CRP and lactate levels were also identified as predictors by ML methods.

Conclusion

By using ML and LR analyse, we showed that D-dimer, axillary temperature, neutrophil count, CRP and lactate levels are risk factors for VTE in COVID-19 patients.

Research In Context

Evidence before this study

The body of evidence available to date has shown that COVID-19 can lead to coagulation dysfunction and endotheliitis, predisposing to a higher risk of venous and arterial thromboembolic events.

Studies have reported variable rates of VTE in patients with COVID-19, ranging from 20%-60% in critically ill patients admitted to intensive care units (ICU) and 5%-20% in those hospitalized in wards.

Despite consistent data indicating this increased incidence, there is a lack of robust data on which clinical and laboratory risk factors influence the occurrence of thromboembolic events in those patients.

Added value of this study

The present study has one of the largest sample sizes among the studies on the subject. In addition to the traditional logistic regression analysis, we also investigated VTE predictors by machine learning approaches, thus contributing to making the available body of evidence more robust. Beside that, our study provides information of thromboembolic complications in the Latin American population with COVID-19, which has not been included in previous studies. To identify a subgroup with a higher thrombotic risk is particularly important considering the Brazilian population, who was severely hit by the pandemic.

Implications of all the available evidence

The available evidence shows that COVID-19 increases the risk of VTE, which, in turn, leads to high mortality in these patients. Therefore, establishing independent predictors for increased risk of thromboembolic complications in COVID-19 patients may help to identify a subgroup of patients who could benefit from a more intensive thromboprophylaxis, as well as to warn about earlier diagnosis and treatment, potentially reducing mortality- and morbidity-associated VTE.

Introduction

Venous thromboembolism (VTE) is an underdiagnosed disease, with an estimated incidence of 10 million cases per year worldwide, and more than half a million deaths.(1) However, its incidence varies widely, depending on the prevalence of genetic and acquired risk factors, such as age, sex, comorbidities, acute illnesses and immobilization in a population.(2) As it leads to high morbidity and mortality,(3) early recognition and prompt treatment are essential.(4)

Coronavirus disease 19 (COVID-19) can trigger an intense endotheliitis and hypercoagulability state, which can lead to an increased thromboembolic risk.(5–7) Several reports have described a high incidence of VTE in patients hospitalized with COVID-19, ranging from 20%-60% in critically ill patients admitted to intensive care units (ICU) and 5%-20% in those hospitalized in wards.(8–10) The incidence of VTE remained high even when thromboprophylaxis was used.(11, 12) In those patients, pulmonary embolism (PE) represents a major diagnostic challenge, as its symptoms and signs overlap with the ones of the severe acute respiratory syndrome (SARS). The occurrence of VTE in patients with COVID-19 has been shown to increase mortality.(13–16) Therefore, there has been a major worldwide effort to identify predictors of VTE in hospitalized COVID-19 patients, as a path to promote prevention, early diagnosis and treatment.(8, 9, 17–19)

Clinical prediction scores for VTE in patients with COVID-19 are still lacking and the main available scores for clinical prediction in medical patients do not seem to perform well in patients with COVID-19.(20)

Additionally, there is still a major inconsistency among the potential predictors of VTE identified by previous studies.(21) In this context, machine learning (ML) techniques, which can identify complex (non-linear) correlations among potential predictors, may be useful tools.(22) However, to the best of our knowledge, the use of ML as an approach to assess VTE predictors in COVID-19 patients has not yet been reported. Thus, this study aims at identifying predictors of VTE in a large cohort of patients hospitalized with COVID-19 in Brazil, using traditional statistical methods as well as ML techniques approaches. We also reported the incidence of thromboembolic complications in COVID-19 and their prognostic impact.

Methods

Study design and settings

This is a retrospective multicenter cohort study, a substudy of the Brazilian COVID-19 Registry, which has been conducted in 37 Brazilian hospitals, described in detail elsewhere.(23) Due to previous evidence of the importance of D-dimer as a predictor of VTE in COVID-9 patients upon hospital admission,(6, 9, 10, 13, 14, 17, 19, 24, 25) we restricted the present analysis to the 16 hospitals in which D-dimer was routinely performed at hospital admission. The hospitals were located in three Brazilian states (Minas Gerais, Santa Catarina and Rio Grande do Sul).

Study subjects

Consecutive adult patients (≥ 18 years) with laboratory-confirmed COVID-19, according to World Health Organization interim guidance,(26) admitted to participating hospitals, between March and September 2020 were enrolled. Patients who were transferred from the participating hospital to another hospital (not part of the cohort) within 30 days and did not have VTE within that period were not included. We also excluded patients who were admitted for other reasons and developed COVID-19 symptoms during their stay (Figure 1).

Data collection and quality assessment

Demographic information, clinical characteristics, laboratory and outcome data were collected by trained hospital staff or undergraduate medical or nurse interns from medical records, by using a validated case report form (CRF) on Research Electronic Data Capture (REDCap).(27, 28) The CRF was hosted at the Telehealth Center, University Hospital, *Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil*.

A detailed data management plan (DMP) was developed and provided to all participating centers. An online DMP training was mandatory prior to data collection. Comprehensive data quality checks were undertaken, to ensure high quality, as previously described.(29) Definitions of variables were published elsewhere.(23) In case the patient was transferred from one participant hospital to another, information about the patient was merged and considered as a single entry.

All covariates in the present study were assessed upon hospital admission, except for in-hospital anticoagulation.

During hospitalization, prophylactic anticoagulation was considered as the use of low-molecular-weight heparin, such as enoxaparin 40 mg once a day, unfractionated heparin 5,000 international units, twice or three times a day, or fondaparinux at a dose of 2·5 mg a day.

Therapeutic anticoagulation, on the other hand, referred to the use of enoxaparin 1mg/kg, twice a day (or once a day, if estimated glomerular filtration rate <30 mL/min/1.73m^2), unfractionated heparin with titrated dose to 1·5-2·5 times the baseline of activated partial thromboplastin time (aPTT) when compared to control or fondaparinux at doses of 5 mg, 7·5 mg or 10 mg once a day, depending on the patient's weight.

Some centers used an intermediate dose of heparin for routine thromboprophylaxis, since this was an available approach at the beginning of the pandemic. Others have used this dose for patients considered to be at high risk for VTE, as defined by the International Society on Thrombosis and Haemostasis (ISTH) guideline.(30) The intensity of the intermediate dose varied according to centers, and was either enoxaparin 40 mg twice-daily, enoxaparin 0·5 mg/kg twice-daily or enoxaparin 1mg/kg once daily (in the absence of severe renal dysfunction). Some institutions, on the other hand, used full-dose anticoagulation for prophylaxis, that is treatment dose with the intention of prevention, in the absence of suspected or confirmed VTE.

Outcomes

Symptomatic VTE was diagnosed based on clinical manifestations confirmed by objective imaging such as compression ultrasonography (CUS) with Doppler or bedside compression ultrasonography for deep venous thrombosis (DVT) and computed tomography pulmonary angiography or ventilation-perfusion scan, for PE. If hemodynamic instability made it impossible to perform the previous tests in patients suspected of PE, the presumptive diagnosis was performed by abnormalities suggestive of acute right ventricular overload on echocardiogram or at the point-of-care multi-organ ultrasonography.(31) Catheter-associated thrombosis or visceral thrombosis were not considered as outcomes.

We also assessed mortality, need for invasive mechanical ventilation, renal replacement therapy and bleeding in patients with confirmed and suspected VTE. Bleeding was classified as: 1) severe if: fatal, critical location (intracranial, spinal, pericardial, articular, retroperitoneal or intramuscular with compartment syndrome), shock, permanent disability, and/or fall in hemoglobin level ≥2 g/dL (1·24 mmol/L) or leading to transfusion of two or more units of whole blood or red cells. 2) not severe, but clinically relevant when it did not meet the criteria for severe bleeding, but required medical intervention, temporary interruption of treatment or caused pain. And 3) Non-serious if none of the previous definitions.

D-dimer assessment

We opted to include centers where D-dimer upon hospital presentation had less than 35% missing values. Assessment of D-dimer levels was not performed with the same method among the 16 centers (Table S1). To allow for a unified analysis, we presented D-dimer levels in relative values, that is, the number of times the D-dimer was increased in relation to the upper limit of the reference value of the test used. Then, we stratified it into five groups, as shown in Table 1.

Statistical analysis

Statistical analyses were conducted using R software (version 4.0.2). Descriptive analyses were used to summarize all the variables: continuous variables were summarized using medians and interquartile ranges (IQR) and categorical variables with counts and percentages.

Logistic regression was used to investigate the associations (odds ratio [OR], 95% confidence interval [95% CI]) of variables at hospital presentation as potential risk factors for VTE (demographic characteristics, underlying medical conditions, home medications, clinical characteristics and laboratory analysis at hospital presentation). Bivariate analysis considered the use of prophylactic anticoagulants at any dosage. For the multivariate model, variables with p<0·15 in bivariate analysis were included and model selection was based on Akaike information criterion (AIC). Before multivariable analysis, missing values were handled by using multiple imputation with chained equations, under the missing at random assumption (mice R package, 10 sets of imputations). We performed these analyses considering first only those patients with confirmed VTE and then including both patients with confirmed and suspected VTE.

Machine learning approaches

We evaluated tree-based boosting (such as extreme gradient boosting machines and light gradient boosting machines) and bagging (essentially random forests) ML algorithms, combined with Shapley Additive ExPlanation (SHAP) values(32) to obtain feature importances and impact of variables on predictions over the same imputed data used with the statistical tools. All algorithms were trained in a 10-fold cross-validation procedure, using a grid search algorithm for hyperparameter tuning. Our particular choice of these tree-based algorithms is due to their higher interpretability,(33) especially when compared with neural or deep learning solutions. Additionally, the use of SHAP values allows us to learn and infer more interesting patterns, such as non-linear correlations, as well as interpreting individual model predictions.

Patient and public involvement

Due to the fact that this was an urgent public health research study in response to a Public Health Emergency of international concern, patients or the public were not involved in the design, conduct, interpretation or presentation of results of this research.

Results

Patients

Among 4,120 consecutive patients included, the median age was 61 years (IQR, 48-72); 55·5% were male; 39·3% critical patients and 60·1% hospitalized in the ward. The most common comorbidities were hypertension, diabetes mellitus and obesity (Table 1). Most patients (91%) received thromboprophylaxis, either at the usual prophylactic (low) dose (78·1%), intermediate (0·7%) or even full dose (12·1%), during hospitalization (Table 1).

Venous thromboembolism was confirmed in 274 (6·7%) patients of whom 74·8% had PE, 19·7% DVT and 5·4% had both conditions.

Risk factors associated with venous thromboembolism

Table S2 shows the results of the bivariate analysis. In multivariable logistic regression analysis (Table 2), the following variables were shown to be independent predictors of VTE: obesity (OR 1·5, 95% CI 1·11-2·02, p<0·01), being an ex-smoker (OR 1·44, 95% CI 1·03-2·01, p=0·03), surgery in the past 90 days (OR 2·2, 95% CI 1·14-4·23, p<0·01), temperature on admission (OR 1·41, 95% CI 1·22-1·63, p<0·01), D-dimer equal or above 4 times the reference value (OR 2·16, 95% CI 1·26-3·67, p<0·01), lactate (OR 1·10, 95% CI 1·02-1·19, p=0·01) and C-reactive protein values (OR 1·09, 95% CI 1·01-1·18, p=0·01), neutrophil count (OR 1·04, 95% CI, 1·01-1·08, p=0·02). Among the protective factors, there were atrial fibrillation/flutter (OR 0·30, 95% CI 0·09-0·99, p=0·04), SF ratio (OR 0·997, 95% CI 0·996-0·998), p<0·01) and prophylactic anticoagulation (OR 0·20, 95% CI 0·15-0·26, p<0·01).

Patients with confirmed VTE had higher mortality (28·4% *vs* 18·5%, p<0·001), required mechanical ventilation (58·4% *vs* 26·4%, p<0·001) and renal replacement therapy (21·5% *vs* 9·7%, p<0·001) more frequently, and bleed more (5·8% vs 1·5%, p<0·001), when compared to the group without confirmed VTE (Table 3).

Machine learning

Figure 2 shows the impact of variables on final prediction of VTE by SHAP values. D-dimer value was the most important feature in predicting VTE, followed by urea, axillary temperature and neutrophils' count. In addition to the D-dimer, axillary temperature and neutrophils count, three other variables identified by the ML methods coincided with those shown by the logistic regression and maintained the direction of the correlation: high C-reactive protein and lactate values increased the risk of VTE (red tone of the graph shifted to the right from point 0), while high SF ratio was associated with lower incidence of the outcome (red tone of the graph left shifted from point 0).

The figure also shows that for hemoglobin, values either too high or too low yield higher risk. For urea, creatinine and lymphocytes count, low values yield higher risk.

[insert Figure 2]

Machine learning vs Traditional statistics

Figure 3 shows the comparison of predictors from LR and ML analyses. There is some intersection between the most important variables identified by the regression analysis and the boosting + shap-values analysis. This intersection is, in particular, expected for the important factors, such as D-dimer, but it is also expected that we find more variables in the boosting algorithm's feature importance, seeing as factors like collinearity do not hinder its performance, meaning that collinear variables are still used, to the degree that they encode some level of new informations. In the analysis by logistic regression, this collinearity can negatively impact the results, being sometimes necessary to remove some variables. Furthermore, variables like Hemoglobin, in which either values too high and too low increase risk, can only be safely captured as important by the shap-values approach, seeing as the regression analysis cannot capture nonlinear approaches without explicit modelling. [insert Figure 2]

[insert Figure 3]

Discussion

This multicenter study, involving 16 Brazilian hospitals, investigated predictors of VTE in hospitalized patients with COVID-19 through logistic regression analysis and ML approaches. Among the variables identified by the multivariate logistic regression analysis as predictive of VTE, six were confirmed by ML analysis, which include D-dimer levels, axillary temperature, neutrophil count, C-reactive protein and lactate levels and SF ratio. D-dimer was one of the main predictors identified in both methods. The incidence rate of VTE was 6·7%, confirming the increased thrombotic risk in COVID-19 patients. Mortality, need for mechanical ventilation and renal replacement therapy were higher in patients who developed VTE in comparison with the patients who did not, highlighting the severity of this complication in the prognosis of COVID-19.

Our study is one of the largest individual studies on risk factors of VTE in COVID-19 patients. It confirms and extends the findings of a recent meta-analysis of 38 studies, which is the largest meta-analysis on the topic to date. This meta-analysis included 7,847 patients and showed that, among other factors, D-dimer, CRP and neutrophil count are independent predictors of VTE in COVID-19 patients. However, the results were limited due to high heterogeneity between individual studies. Overall, the number of studies assessing each clinical characteristic and laboratory tests was small, as well as the sample sizes, which may lead to the lack of power to detect associations we observed in the present analysis.(21)

D-dimer is an important marker of COVID-19 coagulopathy(6) and is often increased in patients with COVID-19, regardless of the presence of thrombosis. A previous study in 1,240 patients showed that the mean D-dimer in undiagnosed patients with pulmonary embolism was 1,371 µg/L.(19) Although D-dimer is also elevated in other conditions, such as sepsis, recent surgery and malignancy,(34) some studies have shown that D-dimer is an important predictor of VTE in COVID-19 patients.(9, 21, 25) A recent meta-analysis suggested that the traditional D-dimer cut-off value (<500 µg/L) used to exclude VTE in the general population seems applicable also to patients with COVID-19.(8) However, as a VTE risk predictor,

there are still uncertainties about which levels would, in fact, predict a VTE. Additionally, the interpretation of D-dimer results is challenging due to the great diversity of methods, cutoff values, measurement units and whether presented as D-dimer units (DDU) or fibrinogen equivalent units (FEU), which are approximately twice those of DDU. A literature review has shown that the majority of studies that assessed D-dimer in COVID-19 patients did not make these points about the test clear, impairing the interpretability of the results.(35) Therefore, the analysis of D-dimer in relative values, compared to the reference value, seems to be more proper. A North-American study showed a significant increase in the strength of association the higher the levels of D-dimer,(13) while a Chinese study indicated that the most significant association with VTE occurred when D-dimer increments ≥ 1·5 fold.(36) In our study, the association of D-dimer with VTE was significant when D-dimer was four or more times above the reference value. A recent North American retrospective cohort observed the same cut-off for VTE or mortality.(25) These data suggest that this cut-off value may be a predictor of VTE in hospitalized COVID-19 patients. However, the previous study used maximum D-dimer levels during hospitalization, which may have been assessed at the moment of the suspicion of the diagnosis. As the main gap is to identify risk factors to help clinical decision, using D-dimer at hospital presentation seems more appropriate and more suitable for clinical application.

Our study showed other independent risk factors as predictors of VTE in COVID-19 patients, which were not previously identified in other studies,(21) such as recent surgery, being an ex-smoker, axillary temperature and lactate levels. Although some authors have questioned the role of traditional risk factors of venous thromboembolic disease as predictors of VTE among COVID-19 patients,(19, 37) our study reassures recent surgery and obesity as independent predictors.

Surgery has been consistently recognized as a major transient risk factor for VTE, among the general population, and recent data has shown that risk may remain elevated for 7 to 12 weeks after surgery.(38) It was quite unexpected that such association was not observed among COVID-19 patients in previous studies. We hypothesized that this may be due to the lack of power or to lack of collection of information on recent surgery in the previous studies. While obesity is considered a minor risk factor for VTE among the general population,(39) in COVID-19 patients, it has been shown to be associated with severe disease and increased risk of mortality.(13, 14, 19, 36, 40) Besides the association of obesity with venous stasis and decreased mobility, patients with obesity also have coagulation abnormalities, which may lead to increased risk of thromboembolic disease.(41–46) Increased plasma levels of fibrinogen, plasminogen activator inhibitor-1, factors VII and VIII, von Willebrand factor, increased platelet activation and higher circulating procoagulant microparticles as well as endothelial dysfunction have been reported in obese patients.(41–46)

Smoking is another minor risk factor for VTE in the general population.(47) It has not been observed to be a predictor among patients with COVID-19 in the more recent individual studies.(21) In our study, previous smoking was an independent predictor of VTE, but current smoking was not. This may be due to underreporting of current smoking, as the rate was less than 4%.

Unlike previous reports, our study identified axillary temperature upon hospital admission as an independent predictor of VTE risk. One of the hypotheses for this may be the consequent contraction of volume secondary to insensitive losses, contributing to the venous stasis of the Virchow's triad.(7)

In our study, inflammatory markers such as C-reactive protein and neutrophil count were independently associated with the occurrence of VTE, in agreement with other reports.(19, 21) However, unlike other publications, we also found that lactate level was an independent predictor of VTE in COVID-19 patients. Lactate level is a marker of disease severity and corroborates previous evidence that indicates an increased thrombotic risk in patients hospitalized with severe infections, such as sepsis and septic shock. (48)

Hospitalization due to acute infections has shown to be a strong trigger for VTE, independent of immobilization. It is hypothesized that this may be a consequence of the inflammatory response, leading to local activation of tissue, factor-mediated coagulation and to local vasoconstriction.(49, 50) In hospitalised patients with COVID-19, the cytokine storm, excessive inflammation and the consequent endothelial injury, inflammatory endotheliitis, besides hypoxia and disseminated intravascular coagulation are believed to play a key role in this process.(5, 7, 51)

In our study, we found that atrial fibrillation and flutter, SF ratio (peripheral oxygen saturation over inspired oxygen fraction) and prophylactic use of anticoagulant were protective factors for VTE, in hospitalized COVID-19 patients. The highest levels of SF ratio (peripheral oxygen saturation over inspired oxygen fraction) likely reflects a diminished severity of the inflammatory response. In fact, SF ratio was an important predictor of mortality in the ABC_2-SPH score, derived from this same cohort.(52) Although less used than the ratio of arterial oxygen partial pressure over inspired oxygen fraction (PaO_2/FiO_2), SF ratio has been validated as a surrogate for the PaO2/FiO2 ratio, to assess the severity of hypoxemia, in patients with acute respiratory distress syndrome.(53) As it does not require vascular puncture, the SF ratio is less invasive, less painful and simpler than the PaO_2/FiO_2 ratio.

The findings of this study confirm those of a previous study which showed that pre-existing cardiovascular diseases are not associated with a higher PE risk, in COVID-19 patients.(19) However, in our study, the presence of atrial fibrillation or flutter was shown to be a protective factor of VTE. This is likely to be a proxy of anticoagulant use, since more than 90% of these patients in our study were using anticoagulants prior to admission and the vast majority of these patients had oral medication changed to therapeutic heparin during hospitalization. Despite this possible reduction in the rate of VTE with full anticoagulation, it does not mean that full-dose anticoagulation should be routinely administered to patients with COVID-19. A recent meta-analysis showed that the indiscriminate use of a full dose of anticoagulant significantly increased the incidence of bleeding and mortality.(54) On the other hand, a recently published randomized multiplatform trial indicated a potential benefit of routine therapeutic anticoagulation for patients hospitalized for non-critical COVID-19, in relation to days free of cardiovascular or respiratory organ support.(55) Another recently published randomized study not included in the meta-analysis(54) showed that the empirical use of anticoagulant at a therapeutic dose

reduced the occurrence of thromboembolic events in patients hospitalized in a ward with D-dimer ≥ 4 times the reference value,(25) the same cut-off we observed as a predictor of VTE in the present study. More studies are still needed to better guide when and for whom to use the full dose of anticoagulant as a prophylactic strategy. However, our study corroborates the most recent evidence that a possible cut-off value of the D-dimer four times the upper limit of reference may be a guide for a more aggressive anticoagulation approach. As expected, in our study, the use of anticoagulants at a prophylactic dose reduced the risk of VTE in COVID-19 patients, corroborating data already available.(19)

In the present study, ML approaches detected other fourteen potential predictors of VTE in patients with COVID-19 in addition to the six variables identified by logistic regression analysis. One of the main advantages in traditional methods, such as regressions, lies in how simple they are and in how just analyzing the model (i.e. looking at the coefficients, for instance) can properly explain what was learned in the model.(32) Despite that, many of such techniques fall short in the sort of patterns they can learn, mostly remaining restricted to linear associations among variables, manually crafted non-linearities and other simpler variable associations. Additionally, LR's performance usually deteriorates in presence of collinearity, which may be especially problematic when the variables are not perfectly collinear and discarding some of them may result in useful information loss. Furthermore, missing values have to be replaced with some form of artificial values, which may also generate problems. Machine learning approaches have the ability of dealing with collinearity and redundancy, which may have occurred among some variables, as well as the ability to assess non-linear correlations.

Among the chief advantages of using ML models is their learning capacities, enabling them to capture much more complex patterns, sometimes even ascending into semantic and abstract levels. Albeit requiring substantially more data points in exchange. In the particular case of decision trees, random forests and gradient boosting machines, collinearity is not a problem, which means no potentially predictive information has to be discarded, and missing values do not require any form of filling.(56, 57) However, there is also an increased risk of identifying spurious (non-significant) associations, mainly due to issues of overfitting.(58)

In multivariate logistic regression analysis, we have not observed an association with some variables which were significant in the aforementioned meta-analysis,(21) including white blood cell count, alanine aminotransferase (ALT), lactate dehydrogenase (LDH), and prolonged prothrombin time, but these variables were observed as predictors in the ML model.

The rate of VTE in our study was at the lowest limit of that usually described in the literature.(9) Several factors may have influenced this finding. Some of the previous studies performed routine imaging exams or even excluded patients who had not performed imaging exams for VTE while asymptomatic for the disease, which may have overestimated the rate of thromboembolic events.(19) Additionally, the first cases of infection in Brazil only occurred at the end of February 2020, while the first wave of the disease only happened between April and May of that same year. At that time, the thrombogenic potential of the disease was already known and the routine use of thromboprophylaxis for patients hospitalized for

COVID-19 was already widespread.(59) The rate of use of thromboprophylaxis, either at high or low dose, in our study was high, with more than 90% of the participants having used it. Therefore, we believe that this is the major reason for the lower incidence in our study. Furthermore, since the publication of the Recovery trial,(60) in June 2020, dexamethasone has been included in the treatment of patients with COVID-19, when they require using oxygen therapy or ventilatory support. It is possible that this has also influenced the decreased incidence of VTE, through reduced inflammation and, therefore, the thrombotic potential. There is already previous evidence suggesting that, in proinflammatory conditions, there is an increase in anticoagulant factors and a reduction in fibrinogen and procoagulant factors.(61) We also hypothesize there could be an underestimation of the occurrence of VTE due to limited access to objective tests, to avoid spreading out the disease. Nevertheless, even considering an incidence of 6·7%, it was higher than that described in other viral infections, supporting the thrombogenic potential of COVID-19.(62)

When compared to the group without VTE, the use of invasive mechanical ventilation, the need for renal replacement therapy and inhospital mortality were about twice as high in patients with VTE, reinforcing the prognostic importance of thrombotic events in patients with COVID-19.(13, 14, 19, 36, 40) As expected, the bleeding rate was higher in groups with VTE, due to the more frequent use of therapeutic doses of anticoagulants. However, most of these bleeding events were non-serious (Table 1). There was no difference in the severity of bleeding between the groups (Table 3).

This study has some limitations. First, the data collection was retrospective, which resulted in missing data on some laboratory tests. As a pragmatic study, they were requested at the discretion of the attending physician. In order to minimize this impact, we agreed, from the beginning of the study, that we would select, for data analysis, those centers with missing data accounting for less than 35% for D-dimer, as this is one of the potential VTE predictors at COVID-19 that most consistently appear in studies on the subject.(8, 9, 13, 36) Second, all variables analysed were collected upon hospital admission, as we would like to provide evidence to alert clinicians, so they could be able to identify, as soon as possible, patients at the highest risk of VTE, allowing for prompt diagnosis and treatment. Therefore, other relevant factors that could increase the risk of VTE, occurring during hospitalization, were not evaluated. Third, laboratory tests were not centralized. In particular, D-dimer was performed by using different methodologies, according to local hospitals. However, we evaluated D-dimer in relative values, stratified into five groups, in relation to the upper limit of normality of the reference values. Due to high variability in D-dimer assays among different institutions, we strongly believe that the way we analysed, increases the applicability of our findings.

Conclusion

We evaluated predictors of VTE in a large cohort of patients with COVID-19 by using both LR analysis and ML approaches. There was consistency between them, by which we identified that D-dimer, axillary temperature, neutrophils count, C-reactive protein and lactate as risk factors for VTE. We suggest that patients presenting these risk factors at admission should be more closely monitored for VTE

development. SF ratio, prophylactic use of anticoagulant and atrial fibrillation, probably as a proxy of anticoagulant use, are protective of VTE development in COVID-19 patients. Finally, we observed that the occurrence of VTE had an impact on higher mortality, the need for mechanical ventilation and renal replacement therapy, reinforcing the importance of early diagnosis and treatment.

Abbreviations

AIC
Akaike information criterion.
ALT
Alanine aminotransferase.
aPTT
activated partial thromboplastin time
95% CI
95% confidence interval.
COVID-19
Coronavirus disease 19.
CRF
Case report form.
CRP
C-reactive protein.
CUS
Compression ultrasonography.
DDU
D-dimer units.
DMP
Data management plan.
DVT
Deep venous thrombosis.
FEU
Fibrinogen equivalent units.
ICU
Intensive care units.
ISTH
International Society on Thrombosis and Haemostasis.
IQR
Interquartile ranges.
LDH
Lactate dehydrogenase.
LR

Logistic regression.
ML
Machine learning.
OR
Odds ratio.
PaO_2/FiO_2
Ratio of arterial oxygen partial pressure over inspired oxygen fraction.
PE
Pulmonary embolism.
REDCap
Research Electronic Data Capture
SARS
Severe acute respiratory syndrome.
SHAP
Shapley Additive ExPlanation.
SF ratio
Peripheral oxygen saturation/inspired oxygen fraction.
VTE
Venous thromboembolism.

Declarations

Ethics

The study was approved by the Brazilian National Commission for Research Ethics (CAAE 30350820.5.1001.0008). Individual informed consent was waived due to the severity of the situation and the use of unidentified data, based on medical chart review only.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported in part by Minas Gerais State Agency for Research and Development (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG*) [grant number APQ-00208-20], National Institute of Science and Technology for Health Technology Assessment (*Instituto de Avaliação de Tecnologias em Saúde – IATS*)/ National Council for Scientific and Technological Development (*Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq*) [grant number 465518/2014-1], and CAPES Foundation (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*) [grant number 88887.507149/2020-00].

The sponsors had no role in study design; data collection, management, analysis, and interpretation; writing the manuscript; and decision to submit it for publication. MSM and MP had full access to all the data in the study and had responsibility for the decision to submit for publication.

Authors' contribution

Substantial contributions to the conception or design of the work: MSM and WCS.

Substantial contributions to the acquisition, analysis, or interpretation of data for the work: MSM, WCS, MCP, LEFR, RTS, BBMP, AVS, AFG, BSMB, BMC, CMR, CDG, CCRC, ECP, EWR, EMSK, FA, FAB, FFMGA, FGA, GPC, GGV, GANB, JHSMC, JRCSF, KBR, LSO, LSP, LSP, LBS, LSFC, LK, MAF, MMS, MC, MAPF, MCAN, MAPM, MNZF, MHGJ, NCSS, NRO, NMP, PGSA, PLA, RAV, RMM, SCF, SMMG, SFA, SAP, TK, TOF and MAG.

Drafted the work: MSM, MCP, PDP, BBMP and WCS.

Revised the manuscript critically for important intellectual content: all authors.

Final approval of the version to be published: all authors.

Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: MSM, MCP and WCS.

Acknowledgments

We would like to thank the hospitals which are part of this collaboration, for supporting this project: Hospital de Clínicas de Porto Alegre; Hospital Eduardo de Menezes; Hospital Julia Kubitschek; Hospital Mãe de Deus; Hospital Márcio Cunha; Hospital Mater Dei Betim-Contagem; Hospital Mater Dei Contorno; Hospital Mater Dei Santo Agostinho; Hospital Metropolitano Dr. Célio de Castro; Hospital Moinhos de Vento; Hospital Nossa Senhora da Conceição; Hospital Santa Cruz; Hospital Santa Rosália; Hospital Semper; Hospital SOS Cárdio; Hospital Universitário de Santa Maria.

We also thank all the clinical staff at those hospitals, who cared for the patients, and all undergraduate students who helped with data collection.

References

1. Haematology TL. Thromboembolism: an underappreciated cause of death. The Lancet Haematology. 2015;2(10).
2. JA H, WM OF, TM P, CM L, MD S, DN M, et al. Relative impact of risk factors for deep vein thrombosis and pulmonary embolism: a population-based study. Arch Intern Med. 2002;162(11):1245–8.
3. V T, V P, SR K. S S. Incidence of and mortality from venous thromboembolism in a real-world population: the Q-VTE Study Cohort. The American journal of medicine. 2013;126(9).
4. SB S, JB G, JM M, NA Z, RE C. TI M. Early anticoagulation is associated with reduced mortality for acute pulmonary embolism. Chest. 2010;137(6).
5. al PLe. COVID-19 is, in the end, an endothelial disease. European Heart Journa. 2020;41:3038–44.
6. U MRAB. DD, E B, M DP, M R, et al. The procoagulant pattern of patients with COVID-19 acute respiratory distress syndrome. Journal of thrombosis and haemostasis: JTH. 2020;18(7).
7. Lowenstein CJ, Solomon SD. Severe COVID-19 Is a Microvascular Disease. Circulation. 2020;142(17):1609–11.
8. Suh YJ, Hong H, Ohana M, Bompard F, Revel MP, Valle C, et al. Pulmonary Embolism and Deep Vein Thrombosis in COVID-19: A Systematic Review and Meta-Analysis. Radiology. 2021;298(2):E70–80.
9. Kefale B, Tegegne GT, Degu A, Tadege M, Tesfa D. Prevalence and Risk Factors of Thromboembolism among Patients With Coronavirus Disease-19: A Systematic Review and Meta-Analysis:. SAGE Journal. 2020;26.
10. JB H. D G, M C, B S, S P, K C, et al. Frequency of venous thromboembolism in 6513 patients with COVID-19: a retrospective study. Blood advances. 2020;4(21):5373–7.
11. C JH, I L-L TFS. M O, X D, et al. High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study. Intensive care medicine. 2020;46(6):1089–98.
12. JF L, C ML, JM C, M M. R, M A, et al. High incidence of venous thromboembolic events in anticoagulated severe COVID-19 patients. Journal of thrombosis haemostasis: JTH. 2020;18(7):743–1746.
13. Bilaloglu S, Aphinyanaphongs Y, Jones S, Iturrate E, Hochman J, Berger JS. Thrombosis in Hospitalized Patients With COVID-19 in a New York City Health System. JAMA. 2020;324(8):799–801.
14. Mahmoud B. Malas INN, Elsayed N, Mathlouthi A, Marmor R, Clary B. Thromboembolism risk of COVID-19 is high and associated with a higher risk of mortality: A systematic review and meta-analysis | Elsevier Enhanced Reader. EClinicalMedicine. 2020;29.
15. Stokes EKZL, Anderson KN, et al. Coronavirus Disease 2019 Case Surveillance - United States, January 22-May 30, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(24):759–65.
16. Wang DHB, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. JAMA. 2020;323(11):1061–9.

17. RP R, C G, GA RC, JS D. G, J H, et al. Diagnosis and Treatment of Pulmonary Embolism During the Coronavirus Disease 2019 Pandemic: A Position Paper From the National PERT Consortium. Chest. 2020;158(6).
18. RL MM, EC ZKWS, NT CVC. C, et al. VTE in ICU Patients With COVID-19. Chest. 2020;158(5):2130–5.
19. C F, O W. A T, D M, T P, N P, et al. Pulmonary embolism in COVID-19 patients: a French multicentre cohort study. European heart journal. 2020;41(32):3058–68.
20. BV S, C J, R P, C M, ML U. T R, et al. Pulmonary embolism and COVID-19: A comparative analysis of different diagnostic models performance. The American journal of emergency medicine. 2021;50.
21. Joshua Henrina ICSP, Irvan Cahyadi S, Lawrensia HFeliciaH, Gunawan A, Cahyadi. Jennifer Franke, Leonardo Paskah Suciadi. Clinical characteristics and outcomes of venous thromboembolism in patients hospitalized for COVID-19: Systematic review and meta-analysis. Thrombosis Update. 2021;2.
22. Hou L, Hu L, Gao W, Sheng W, Hao Z, Chen Y, et al. Construction of a Risk Prediction Model for Hospital-Acquired Pulmonary Embolism in Hospitalized Patients:. https://doiorg/101177/10760296211040868. 2021;27:1-9.
23. MS M, PK Z, MVR S-S, IJB N, LM O, LS M, et al. Clinical characteristics and outcomes of patients hospitalized with COVID-19 in Brazil: Results from the Brazilian COVID-19 registry. International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases. 2021(21).
24. Cerdà P, Ribas J, Iriarte A, Mora-Luján JM, Torres R, Del Río B, et al. Blood test dynamics in hospitalized COVID-19 patients: Potential utility of D-dimer for pulmonary embolism diagnosis. PLoS One. 2020;15(12).
25. SL C, MA EG. B, S C, N K, M L, et al. Prevalence and Predictors of Venous Thromboembolism or Mortality in Hospitalized COVID-19 Patients. Thromb Haemost. 2021;121(8):1043–53.
26. Organization WH. Diagnostic testing for SARS-CoV-2: interim guidance, 11 September 2020. 2020.
27. PA H, R T, R T. J P, N G, JG C. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81.
28. PA H, R T, BL M. V E, M F, L ON, et al. The REDCap consortium: Building an international community of software platform partners. Journal of biomedical informatics. 2019;95.
29. Gregory KERL. Research strategies that result in optimal data collection from the patient medical record. Appl Nurs Res. 2012;25(2):108–16.
30. AC S, JH L, JM WA, BJ C. H, T I, et al. Scientific and Standardization Committee communication: Clinical guidance on the diagnosis, prevention, and treatment of venous thromboembolism in hospitalized patients with COVID-19. Journal of thrombosis haemostasis: JTH. 2020;18:8.
31. SV K, C GM, GJ BHB, VP G. H, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). European heart journal. 2020;41(4):543–603.

32. Lundberg S, Lee S-I. A Unified Approach to Interpreting Model Predictions. 2017.
33. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. ed. New York: Springer; 2009.
34. Christopher Kabrhel M, MPH D, Mark Courtney MD, Carlos A, Camargo MD Jr, DrPH MC, Plewa MD, Kristen E, Nordenholz MD, Christopher L, Moore MD, Peter B, Richman MD, Howard MBA, Smithline A, Daren MD, Beam M. and Jeffrey A. Kline, MD. Factors Associated With Positive D-dimer Results in Patients Evaluated for Pulmonary Embolism. Acad Emerg Med. 2010;17:589–97.
35. EJ F. J T. Reporting of D-dimer data in COVID-19: some confusion and potential for misinformation. Clinical chemistry laboratory medicine. 2020;58(8):1191–9.
36. JY L, HF W, DL PYDL. W, P P, et al. Clinical characteristics and risk factors for symptomatic venous thromboembolism in hospitalized COVID-19 patients: A multicenter retrospective study. Journal of thrombosis haemostasis: JTH. 2021;19(4):1038–48.
37. COVID-19 and Thrombotic or Thromboembolic Disease. Implications for Prevention, Antithrombotic Therapy, and Follow-Up | Elsevier Enhanced Reader. JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY. 2021;75(23).
38. Caron A, University Lille CL, EA 2694, Santé publique: épidémiologie et qualité des soins, Lille, France, Depas N, University Lille CL, EA 2694, Santé publique: épidémiologie et qualité des soins, Lille, France, Chazard University Lille E CL, EA 2694, Santé publique: épidémiologie et qualité des soins, Lille, France, et al. Risk of Pulmonary Embolism More Than 6 Weeks After Surgery Among Cancer-Free Middle-aged Patients. JAMA Surgery. 2019;154(12):1126-32.
39. Tsai AW, Cushman M, Rosamond WD, Heckbert SR, Polak JF, Folsom AR. Cardiovascular Risk Factors and Venous Thromboembolism Incidence: The Longitudinal Investigation of Thromboembolism Etiology. Arch Intern Med. 2002;162(10):1182–9.
40. S C, T Z, S W, Y Y, P W, Y S, et al. Association between risk of venous thromboembolism and mortality in patients with COVID-19. International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases. 2021;108:543–9.
41. I M, LF VG. Obesity, haemostasis and the fibrinolytic system. Obesity reviews: an official journal of the International Association for the Study of Obesity. 2002;3(2):85–101.
42. MT SBGP, MR G, C MFS. P, et al. Insulin resistance as a determinant of platelet activation in obese women. J Am Coll Cardiol. 2006;48(12):2531–8.
43. I J-V MCA. A M, PE M. Plasminogen activator inhibitor-1, inflammation, obesity, insulin resistance and vascular risk. Journal of thrombosis haemostasis: JTH. 2003;1(7):1575–9.
44. N P, V DM. R M, R G, G DP. Effect of glucose tolerance status on PAI-1 plasma levels in overweight and obese subjects. Obesity research. 2002;10(8):717–25.
45. L BG. G, D D, S V, L M, JL S, et al. Circulating procoagulant microparticles in obesity. Diabetes Metab. 2006;32(1):82–5.
46. O M, L FLLG J, N M, D D, et al. Short-term very low-calorie diet in obese females improves the haemostatic balance through the reduction of leptin levels, PAI-1 concentrations and a diminished

release of platelet and leukocyte-derived microparticles. International journal of obesity (2005). 2011;35(12):1479-86.

47. ER P, FR R. CJ D. Smoking increases the risk of venous thrombosis and acts synergistically with oral contraceptive use. Am J Hematol. 2008;83(2):97–102.
48. TC DK, CG C, RC ESM, LW P. K, et al. VTE Incidence and Risk Factors in Patients With Severe Sepsis and Septic Shock. Chest. 2015;148(5):1224–30.
49. T I, YIGV GG, JB TSKB. H. Acute infection as a trigger for incident venous thromboembolism: Results from a population-based case-crossover study. Research practice in thrombosis haemostasis. 2017;2(1):85–92.
50. MJ S, JJ H. H Z, AS S. Pulmonary coagulopathy as a new target in therapeutic studies of acute lung injury or pneumonia–a review. Critical care medicine. 2006;34(3).
51. RA C. WH F. Thrombotic Complications of COVID-19 Infection: A Review. Cardiology in review. 2021;29(1).
52. MS M, MC P, LEF R, RT S, LM O, RLR C, et al. ABC 2-SPH risk score for in-hospital mortality in COVID-19 patients: development, external validation and comparison with other available scores. International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases. 2021;110:281–308.
53. TW R, AP W, GR B, DL H. DA S, LB W. Comparison of the SpO2/FIO2 ratio and the PaO2/FIO2 ratio in patients with acute lung injury or ARDS. Chest. 2007;132(2):410–7.
54. C M. D S, T C, P R, N U. Anticoagulation and In-Hospital Mortality From Coronavirus Disease 2019: A Systematic Review and Meta-Analysis. Clinical and applied thrombosis/hemostasis: official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis. 2021;27.
55. PR L, EC G, JS B, MD N, BJ M, JC N, et al. Therapeutic Anticoagulation with Heparin in Noncritically Ill Patients with Covid-19. N Engl J Med. 2021;385(9):790–802.
56. MS M, Gerda HRA. D, J. PJ. Healthcare Expenditure Prediction with Neighbourhood Variables – A Random Forest Model. Statistics Politics Policy. 2020;11(2):111–38.
57. Friedman JH. Greedy function approximation: A gradient boosting machine. https://doiorg/101214/aos/1013203451. 2001;29(5):1189-232.
58. Sagawa S, Raghunathan A, Koh PW, Liang P. An Investigation of Why Overparameterization Exacerbates Spurious Correlations. 2020;3.
59. Extremely High Incidence of Lower Extremity. Deep Venous Thrombosis in 48 Patients With Severe COVID-19 in Wuhan. Circulation. 2021;142(2):181–3.
60. PH, WS L, JR. E, M M, JL B, L L, et al. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med. 2021;384(8):693–704.
61. vZ B, VE ENAS, HR G. B, OM D, et al. Systematic review on the effect of glucocorticoid use on procoagulant, anti-coagulant and fibrinolytic factors. Journal of thrombosis haemostasis: JTH. 2010;8(11):2483–93.

62. PE B, SM H, MD MNKSWCL. C. Pandemic H1N1 influenza infection and vascular thrombosis. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America. 2011;52(2).

Tables

Table 1. Demographic and clinical characteristics of cohort of Brazilian patients admitted to hospital with COVID-19

Characteristics	Confirmed VTE (n= 274[1])		Non VTE (n = 3,846[1])	
	Frequency (%) or median (IQR)	Non missing cases (%)	Frequency (%) or median (IQR)	Non missing cases (%)
Age (years)	63·0 (51·0, 72·0)	274 (100%)	60·0 (48·0, 72·0)	3,846 (100%)
Sex at birth		274 (100%)		3,845 (100%)
Male	150 (54·7%)		2,134 (55·5%)	
Comorbidities				
Hypertension	151 (55·1%)	274 (100%)	2,092 (54·4%)	3,846 (100%)
Coronary artery disease	16 (5·8%)	274 (100%)	192 (5·0%)	3,846 (100%)
Heart failure	15 (5·5%)	274 (100%)	242 (6·3%)	3,846 (100%)
Atrial fibrillation/flutter	3 (1·1%)	274 (100%)	137 (3·6%)	3,846 (100%)
Stroke	9 (3·3%)	274 (100%)	141 (3·7%)	3,846 (100%)
Asthma	19 (6·9%)	274 (100%)	272 (7·1%)	3,846 (100%)
COPD	24 (8·8%)	274 (100%)	233 (6·1%)	3,846 (100%)
Diabetes mellitus	87 (31·8%)	274 (100%)	1,084 (28·2%)	3,846 (100%)
Obesity[a]	68 (24·8%)	274 (100%)	698 (18·1%)	3,846 (100%)
Cirrhosis	2 (0·7%)	274 (100%)	21 (0·5%)	3,846 (100%)
Chronic kidney disease	8 (2·9%)	274 (100%)	204 (5·3%)	3,846 (100%)
Rheumatological disease	0 (0.0%)	274 (100%)	3 (0·1%)	3,846 (100%)
HIV infection	4 (1·5%)	274 (100%)	42 (1·1%)	3,846 (100%)
Cancer	14 (5·1%)	274 (100%)	170 (4·4%)	3,846 (100%)
Surgery in previous 90 days	14 (5·1%)	274 (100%)	87 (2·3%)	3,841 (100%)
Previous Transplant	1 (0·4%)	274 (100%)	17 (0·4%)	3,846 (100%)
Medications on admission				
NSAIDs	9 (3·3%)	274 (100%)	135 (3·5%)	3,846 (100%)
Potassium sparing	9 (3·3%)	274 (100%)	106 (2·8%)	3,846 (100%)

diuretic				
Thiazide diuretic	32 (11·7%)	274 (100%)	496 (12·9%)	3,846 (100%)
Hypoglycemic (non-insulin)	55 (20·1%)	274 (100%)	693 (18·0%)	3,846 (100%)
Immunosuppressant	3 (1·1%)	274 (100%)	21 (0·5%)	3,846 (100%)
ACE or BRA inhibitor	102 (37·2%)	274 (100%)	1,313 (34·1%)	3,846 (100%)
Insulin	19 (6·9%)	274 (100%)	270 (7·0%)	3,846 (100%)
Statin	53 (19·3%)	274 (100%)	714 (18·6%)	3,846 (100%)
Amiodarone	0 (0·0%)	274 (100%)	48 (1·2%)	3,846 (100%)
Oral anticoagulant	13 (4·7%)	274 (100%)	290 (7·5%)	3,846 (100%)
Beta blocker	38 (13·9%)	274 (100%)	694 (18·0%)	3,846 (100%)
Calcium channel blocker	30 (10·9%)	274 (100%)	469 (12.2%)	3,846 (100%)
Inhaled corticosteroid	6 (2·2%)	274 (100%)	127 (3·3%)	3,846 (100%)
Oral corticosteroids	7 (2·6%)	274 (100%)	78 (2·0%)	3,846 (100%)
Digitalic	0 (0·0%)	274 (100%)	20 (0·5%)	3,846 (100%)
Loop diuretic	16 (5·8%)	274 (100%)	278 (7·2%)	3,846 (100%)
Clinical Characteristics at admission				
Temperature (°C)	36·6 (36·1, 37·4)	169 (62%)	36·5 (36·0, 37·2)	2,617 (68%)
Systolic blood pressure (mmHg)		261 (95%)		3,687 (96%)
> 90 mmHg without amine	227 (87·0%)		3,445 (93·4%)	
< 90 mmHg without amine	9 (3·4%)		45 (1·2%)	
Any value, but with amine	25 (9·6%)		197 (5·3%)	
Diastolic blood pressure (mmHg)		261 (95%)		3,685 (96%)
> 60 mmHg without amine	207 (79·3%)		3,010 (81·7%)	
< 60 mmHg without amine	29 (11·1%)		478 (13·0%)	
Any value, but with amine	25 (9·6%)		197 (5·3%)	

Heart rate (bpm)	90·0 (80·0, 103·0)	264 (96%)	88·0 (78·0, 100·0)	3,693 (96%)
Respiratory rate (bpm)	21 (18, 25)	222 (81%)	20 (18, 24)	3,153 (82%)
Glasgow coma score < 15	44 (16·1%)	274 (100%)	504 (13·1%)	3,846 (100%)
Laboratory tests				
D-dimer/maximum reference value		239 (87%)		3,069 (80%)
≤ 1 x	30 (12·6%)		684 (22·3%)	
1-1·9 x	54 (22·6%)		857 (27·9%)	
2-3·9 x	36 (15·1%)		539 (17·6%)	
4-9·9 x	37 (15·5%)		267 (8·7%)	
≥ 10 x	82 (34·3%)		722 (23·5%)	
C-reactive protein (mg/L)	94·3 (54·2, 183·7)	243 (89%)	72·8 (33·4, 130·1)	3,460 (90%)
Hemoglobin (g/L)	13·1 (11·8, 14·2)	269 (98%)	13·4 (12·2, 14·5)	3,777 (98%)
Leukocytes count (cels/mm^3)	8·8 (6·0, 11·9)	269 (98%)	6·9 (5·1, 9·4)	3,777 (98%)
Neutrophils count (cels/mm^3)	6,928·0 (4,310·0, 9,205·0)	269 (98%)	4,946·1 (3,374·0, 7,452·0)	3,658 (95%)
Lymphocytes count (cels/mm^3)	1,000·0 (684·5, 1,355·0)	267 (97%)	1,058·0 (730·0, 1,478·5)	3,656 (95%)
Neutrophils-to-lymphocytes ratio	6·2 (4·0, 10.6)	267 (97%)	4·7 (2·8, 8·0)	3,654 (95%)
Platelet count (10^9/L)	214·0 (162·0, 282·2)	268 (98%)	197·0 (155·0, 256·0)	3,742 (97%)
TGP/ALT (U/L)	35·5 (23·0, 56·0)	207 (76%)	34·9 (22·0, 56·0)	2,791 (73%)
TGO/AST (U/L)	43·0 (32·0, 63·8)	205 (75%)	40·0 (28·9, 59·6)	2,806 (73%)
Arterial pO_2 (mmHg)	76·0 (63·0, 100·0)	237 (86%)	76·0 (64·0, 97·8)	3,186 (83%)
Arterial pCO_2 mmHg	35·7 (32·0, 40·0)	237 (86%)	35·0 (31·9, 39·0)	3,196 (83%)

SF ratio	350·0 (120·0, 441·7)	266 (97%)	428·6 (328·6, 452·4)	3,755 (98%)
Creatinine (mg/dL)	0·9 (0·7, 1·2)	264 (96%)	0·9 (0·8, 1·2)	3,683 (96%)
Sodium (mmol/L)	138·0 (135·0, 140·1)	260 (95%)	138·0 (135·0, 140·0)	3,507 (91%)
Lactate (mmol/L)		274 (100%)		3,842 (100%)
Lactate dehydrogenase (U/L)	421·0 (336·1, 629·0)	177 (65%)	373·0 (272·0, 511·0)	2,443 (64%)
INR	1·1 (1·0, 1·2)	210 (77%)	1·1 (1·0, 1·2)	2,410 (63%)
Lifestyle habits				
Illicit drugs	1 (0·4%)	274 (100%)	32 (0·8%)	3,846 (100%)
Alcoholism	9 (3·3%)	274 (100%)	155 (4·0%)	3,846 (100%)
Current smoking	8 (2·9%)	274 (100%)	144 (3·7%)	3,846 (100%)
Ex-smoker	53 (19·3%)	274 (100%)	591 (15·4%)	3,846 (100%)
Anticoagulant during hospitalization*				
Prophylactic use of anticoagulant [ƒ]	166 (60·6%)	274 (100%)	3,081 (80·1%)	3,846 (100%)
Full-dose anticoagulation for prophylaxis	0 (0·0%)	274 (100%)	498 (12·9%)	3,846 (100%)
Therapeutic use of anticoagulant	204 (74·5%)	274 (100%)	612 (15·9%)	3,846 (100%)
Admission to intensive care	195 (71·2%)	274 (100%	1,426 (37·1%)	3,846 (100%)

ACEi: angiotensin-converting enzyme inhibitor; ARB: Angiotensin receptor blocker; bpm: beats per minute; COPD: chronic obstructive pulmonary disease; FiO_2: fraction of inspired oxygen; HIV: human immunodeficiency virus; INR: international normalized ratio; IQR: Interquartile range; NSAIDs: nonsteroidal anti-inflammatory drugs; O_2 saturation (%): peripheral oxygen saturation; PaO_2: partial pressure of oxygen; $PaCO_2$: partial pressure of carbon dioxide; SF ratio: peripheral O_2 saturation/FiO_2; TGO/AST: aspartate aminotransferase; TGP/ALT: Alanine aminotransferase; VTE: venous thromboembolism. [a]BMI >30kg/m^2. *The rate of anticoagulant use, summing the three strategies (usual prophylactic use, full dose of anticoagulation for prophylaxis and therapeutic use), exceeds 100%, due to the fact that the same patient transitioned from prophylactic dose to full dose of anticoagulation, and vice versa, in the same hospitalization. [ƒ] Of these, 29 patients (0·7% in total) used an intermediate dose of anticoagulation.

Table 2. Multivariable analysis for prediction of symptomatic venous thromboembolism, based on variables available upon hospital presentation

Variable	Frequency (%) or median (IQR)	Confirmed VTE	
		Odds ratio (95% CI)	P value
Obesity[a]	766 (18·6%)	1·50 (1·11-2·02)	<0·01
Atrial fibrillation/flutter	140 (3·4%)	0·30 (0·09-0·99)	0·04
Previous use of beta blocker	732 (17·8%)	0·73 (0·50-1·07)	0·11
Ex-smoker	644 (15·6%)	1·44 (1·03-2·01)	0·03
Surgery in previous 90 days	101 (2·5%)	2·20 (1·14-4·23)	<0·01
Temperature (°C)[bc]	36·5 (36·0, 37·2)	1·41 (1·22-1·63)	<0·01
SF ratio[bd]	428·6 (317·9, 452·4)	0·87 (0·83-0·93)	<0·01
D-dimer/maximum reference value[b]			
1-1·9x	911 (22·1%)	1·32 (0·83-2·09)	0·239
2-3·9x	575 (13·9%)	1·19 (0·72-1·96)	0·486
4-9·9x	304 (7·3%)	2·16 (1·26-3·67)	<0·01
≥ 10x	804 (19·5%)	1·89 (1·18-3·01)	<0·01
Lactate[be]	1·4 (1·1, 1·9)	1·10	0·01

		(1·02-1·19)	
C-reactive protein (mg/L)[bd]	74·4 (34·0, 134·1)	1·09 (1·01-1·18)	0·01
Neutrophils' count[bf]	5,045·0 (3,400·0, 7,613·8)	1·04 (1·005-1·075)	0·02
Prophylactic use of anticoagulant	3,247 (78·8%)	0·20 (0·15-0·26)	<0·01
Full-dose anticoagulation for prophylaxis	498 (12·1%)	NA	0·95

IQR: Interquartile range; SF ratio: oxygen saturation/inspired oxygen fraction. [a]BMI (Body mass index)>30kg/m^2; [b]Data regarding hospital presentation; [c]Increment of 1·0°C; [d]Increment of 50 units; [e]Increment of 1 unit; [f]Increment of 1000 units.

Table 3. Outcomes in patients with and without confirmed venous thromboembolism

Outcomes	Diagnosis of VTE		P value[2]
	No[1] (n= 3846)	Yes[1] (n= 274)	
Invasive Mechanical Ventilation	1016 (26·4%)	160 (58·4%)	<0·001
Need for renal replacement therapy	373 (9·7%)	59 (21·5%)	<0·001
Death	710 (18·5%)	77 (28·4%)	<0·001
Bleeding	56 (1·5%)	16 (5·8%)	<0·001
Severity of bleeding			0·311
Severe	26 (46·4%)	4 (25·0%)	
Not severe, but clinically relevant	18 (32·1%)	8 (50·0%)	
Not severe	12 (21·4%)	4 (25·0%)	

[1]Statistics presented: n (%); [2]Statistical tests performed: Chi-square test of independence; Fisher's exact test.

Supplementary Tables

Table S1. D-dimer methods used by each center included in the study.

Study center code	Method	Equipment	Reference values	Measurement unit	Biological fluid
1001	Enzyme-linked fluorescente assay ELFA)	Biomérieux	<500	ng/mL FEU	Citrated plasma.
	Enzyme-linked fluorescent assay (ELFA)	Alere	<400	ng/mL D-DU	Total blood in ethylenediamine tetraacetic acid (EDTA)
	Immunoturbidimetry	Siemens	<550	ng/mL FEU	Citrated plasma.
1002	Enzyme-linked fluorescent assay ELFA)	Biomérieux	<500	ng/mL FEU	Citrated plasma.
	Enzyme-linked fluorescent assay (ELFA)	Alere	<400	ng/mL D-DU	Total blood in ethylenediamine tetraacetic acid (EDTA)
	Immunoturbidimetry	Siemens	<550	ng/mL FEU	Citrated plasma.
1004	Immunoturbidimetry	CS-2500 (Sysmex)	<0·5	µg/mL DDU	Citrated plasma.
1006	Immunoturbidimetry	ACL TOP	<0·5	µg/mL DDU	Citrated plasma.
1008	Imunoassay	CS-2500 CA660	<500	ng/mL FEU	Citrated plasma.
1009	Immunoturbidimetry	CS-2500 (Sysmex)	<0·5	mg/L FEU	Citrated plasma.
1012	Fluorescent immunoassay (FIA)	Ichroma II	<500	ng/mL FEU	Total blood
1013	Immunofluorescence	Ichroma D-dimer	<500	ng/mL FEU	Citrated plasma.
1015	Fluorescence polarization immunoassay (FPIA)	Fanecare	<400	ng/mL FEU	Total blood
	Immunoturbidimetry	Fanecare	<500	ng/mL FEU	Total blood
	Fluorescent immunodetection	Fanecare	<0·5	mg/mL FEU	Total blood
1018	Immunoturbidimetry	Compact	<0·5	µg/mL DDU	Citrated plasma
1019	Immunoturbidimetry	ACLTOP 550	<500	ng/mL FEU	Citrated plasma

1021	Immunoturbidimetry	ACLTOP 550	< 500	ng/mL FEU	Citrated plasma
1022	Enzyme Linked Immunosorbent Assay (ELISA) (D-dimer HS 500)	ACL TOP 500	<0·5	µg/mL FEU	Citrated plasma
1027	Immunoturbidimetry	ACL TOP350 Werfen	<500	ng/ml FEU	Citrated plasma
1037	Immunoturbidimetry	Satellite STA	< 500	ng/mL FEU	Citrated plasma

FEU: Fibrinogen Equivalente Units; D-DU: D-Dimer Units. It is generally accepted that 1 D-DU = 2 FEU (lack of standardization and different antibody configurations reduce the reliability of this conversion factor.

Table S2. Bivariable analysis for the prediction of venous thromboembolism		
Characteristics	**Confirmed VTE (n= 274)**	
	Odds Ratio (95% IC)	P value
Age (years)	1 (1-1·01)	0·194
Sex at birth	0·99 (0·77-1·28)	0·961
Hypertension	1·11 (0·86-1·43)	0·41
Coronary artery disease	1·37 (0·77-2·28)	0·255
Heart failure	0·87 (0·48-1·46)	0·612
Atrial fibrillation/flutter	0·25 (0·06-0·68)	0·02
Stroke	0·94 (0·44-1·8)	0·868
Chagas disease	NA	0·994
Rheumatic valve disease	NA	0·997
Other cardiovascular disease	1·22 (0·7-2)	0·46
No relevant disease	0·88 (0·68-1·13)	0·327
Asthma	0·95 (0·56-1·51)	0·825
COPD	1·65 (1·02-2·54)	0·031
Pulmonary fibrosis	NA	0·991
No respiratory disease	0·86 (0·61-1·24)	0·408
Diabetes mellitus	1·24 (0·95-1·63)	0·113
Obesity (BMI > 30kg/m2)	1·63 (1·20-2·17)	0·001
No Metabolic Diseases	0·63 (0·49-0·81)	<0·001
Hematologic neoplasm	NA	0·989
Solid organ Neoplasm	1·54 (0·83-2·67)	0·142
No neoplasm information	NA	0·995
Cirrhosis	1·4 (0·22-5·07)	0·661
Psychiatric illness	0·8 (0·51-1·21)	0·302
Chronic kidney disease	0·6 (0·3-1·06)	0·102
Rheumatological disease	1·03 (0·42-2·19)	0·939

HIV infection	1·15 (0·38-2·88)	0·783
Cancer	1·11 (0·63-1·83)	0·705
Number of Comorbidities*		
0	1·42 (0·96-2·14)	0·084
1	1·42 (0·96-2·13)	0·084
2	1·63 (1·05-2·54)	0·029
3	1·49 (0·85-2·55)	0·153
4	1·13 (0·5-2·32)	0·756
5	1·83 (0·42-5·64)	0·35
6	NA	0·991
7	NA	0·997
8	1·42 (0·96-2·14)	0·084
Number of Cardiovascular Comorbidities***		
0	1·14 (0·87-1·5)	0·328
1	0·97 (0·61-1·49)	0·883
2	1·42 (0·69-2·68)	0·306
3	NA	0·99
4	NA	0·994
5	1·14 (0·87-1·5)	0·328
Surgery in previous 90 days	2·34 (1·22-4·18)	0·006
Surgery in previous 90 days	2·34 (1·22-4·18)	0·006
Previous Transplant	0·41 (0·02-2·06)	0·387
Solid Organ Transplantation	0·41 (0·02-2·06)	0·387
No relevant health conditions	1·11 (0·86-1·45)	0·431
Illicit drugs	0·51 (0·03-2·44)	0·516
Alcoholism	0·82 (0·38-1·55)	0·572
Current smoking	0·75 (0·33-1·46)	0·433
Ex-smoker	1·45 (1·04-1·98)	0·024
NSAIDs	0·84 (0·39-1·61)	0·63

Potassium sparing diuretic	1·17 (0·53-2·27)	0·666
Thiazide diuretic	0·94 (0·63-1·37)	0·755
Hypoglycemic (non-insulin)	1·13 (0·82-1·54)	0·434
Immunosuppressant	1·69 (0·39-5·18)	0·414
ACEi or ARB inhibitor	1·22 (0·94-1·58)	0·134
Insulin	1·06 (0·63-1·7)	0·813
Statin	1·09 (0·79-1·49)	0·586
Amiodarone	NA	0·986
Oral anticoagulant	0·55 (0·30-0·96)	0·047
Beta blocker	0·72 (0·5-1·02)	0·075
Calcium channel blocker	0·89 (0·58-1·3)	0·55
Inhaled corticosteroid	0·66 (0·25-1·4)	0·328
Oral corticosteroids	1·2 (0·49-2·51)	0·662
Digitalic	NA	0·991
Loop diuretic	0·76 (0·43-1·26)	0·312
Other continuous use medication	0·92 (0·7-1·19)	0·517
Does not use continuous medication	0·98 (0·74-1·27)	0·856
Duration of symptoms (days)	1 (0·98-1·02)	0·964
Temperature (°C)[a]	1·25 (1·06-1·47)	0·006
Fever on admission	1·54 (1·02-2·28)	0·035
Heart rate (bpm)	1·01 (1-1·02)	0·004
Respiratory rate (bpm)	1·04 (1·01-1·06)	0·006
Mechanical ventilation on admission	3·08 (2·13 – 4·46)	<0·001
Systolic blood pressure		
< 90 mmHg without amine	3·43 (1·5-7·12)	0·002
Any value, but with amine	2·42 (1·5-3·77)	<0·001
Diastolic blood pressure		
< 60 mmHg without amine	0·85 (0·55-1·26)	0·436

Any value, but with amine	2·31 (1·43-3·6)	<0·001
Glasgow coma score < 15	1·3 (0·91-1·82)	0·138
Mental Status		
Sedated	3·0 (2·06-4·35)	< 0·001
Coma	0·99 (0·12-7·8)	0·993
Altered without coma	0·96 (0·43-2·13)	0·93
Mental confusion	0·36 (0·13-0·99)	0·048
Alert		
D-dimer/maximum reference value		
≤ 1x		
1-1·9x	1·57 (0·99-2·52)	0·06
2-3·9x	1·59 (0·96-2·66)	0·071
4-9·9x	3·18 (1·90-5·36)	<0·001
≥ 10x	2·74 (1·78-4·32)	<0·001
C-reactive protein (mg/L)[b]	1·23 (1·15-1·31)	<0·001
Leukocytes count[c] (cels/mm^3)	1·01 (1-1·03)	0·05
Neutrophils count[c] (cels/mm^3)	1·09 (1·06-1·12)	<0·001
Lymphocytes count[d] (cels/mm^3)	0·98 (0·96-1·00)	0·18
Neutrophils-to-lymphocytes ratio	1·01 (1-1·02)	0·187
Rods count[d] (cels/mm^3)	1·01 (1·004-1·034)	0·014
Platelet count (10^9/L)	12·9 (3·6-46·4)	<0·001
Hemoglobin (g/L)	0·93 (0·88-0·99)	0·029
TGP/ALT (U/L)	1 (0·99-1·00)	0·987
TGO/AST (U/L)	1 (0·99-1·00)	0·724
pH	0·34 (0·08-1·69)	0·169
Arterial pO_2	1 (0·99-1)	0·07

(mmHg)		
Arterial pCO_2 (mmHg)	1·01 (1-1·02)	0·17
O_2 saturation (%)	0·95 (0·93-0·97)	<0·001
FiO_2 at admission	5·01 (2·9-8·45)	<0·001
SF ratio[b]	0·82 (0·78-0·86)	<0·001
pO_2/FiO_2 Ratio	0·998 (0·997-0·999)	<0·001
HCO_3^-(mmoL/L)	0·99 (0·96-1·03)	0·647
Creatinine (mg/dL)	0·9 (0·75-1·03)	0·168
Urea (mg/dL)	1 (1-1)	0·714
Potassium (mmoL)	1·09 (0·88-1·33)	0·417
Sodium (mmol/L)	1·03 (1-1·05)	0·041
Lactate	1·01 (1·006-1·02)	<0·001
Lactate[a] dehydrogenase	1·14 (1·06-1·23)	<0·001
INR	1·09 (0·93-1·24)	0·181
Anticoagulant during hospitalization		
Non-fractionated heparin dose		
Prophylactic	1.08 (0.79-1.48)	0.624
Therapeutic	32.38 (19.9-54.13)	<0.001
Low molecular weight heparin		
Prophylactic	1.2 (0.85-1.73)	0.31
Therapeutic	39.67 (27.9-57.09)	<0.001
Fondaparinux dose		
Prophylactic	2.93 (0.15-17.26)	0.322
Therapeutic	NA	0.997

95% CI: confidence interval; ACEi: angiotensin-converting enzyme inhibitor; ARB: Angiotensin receptor blocker; BMI: body mass index; bpm: breaths per minute; COPD: chronic obstructive pulmonary disease; HIV: human immunodeficiency vírus; INR: international normalized ratio; NSAIDs: nonsteroidal antiinflammatory drugs; O_2 saturation (%): oxygen saturation; PaO_2: partial pressure of oxygen; $PaCO_2$: partial pressure of carbon dioxide; SF ratio: O_2 saturation/FiO_2; FiO_2: fraction of

inspired oxygen; TGO/AST: aspartate aminotransferase; TGP/ALT: Alanine aminotransferase; VTE: venous thromboembolism; [a]Increment of 1ºC; [b]Increment of 50 units; [c]Increment of 1000 units; [d]Increment of 100 units.

Figures

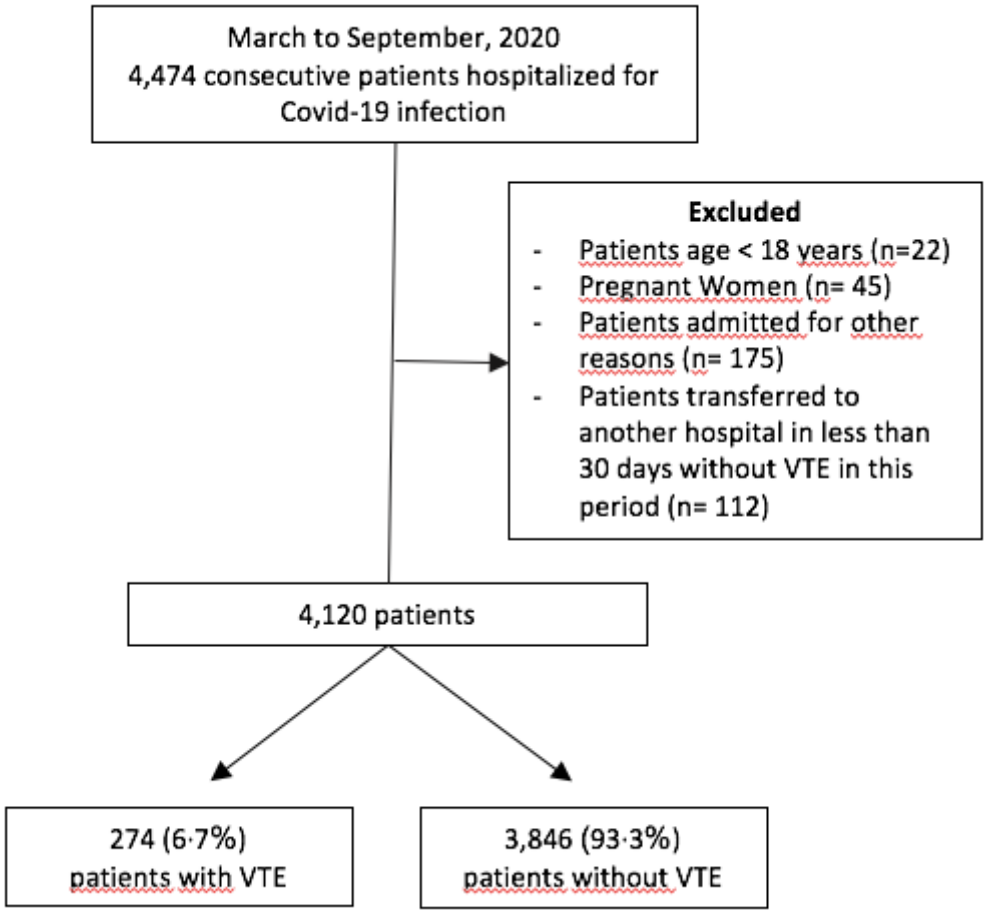

Figure 1

Flowchart of Brazilian patients included in the study.

VTE: Venous thromboembolism.

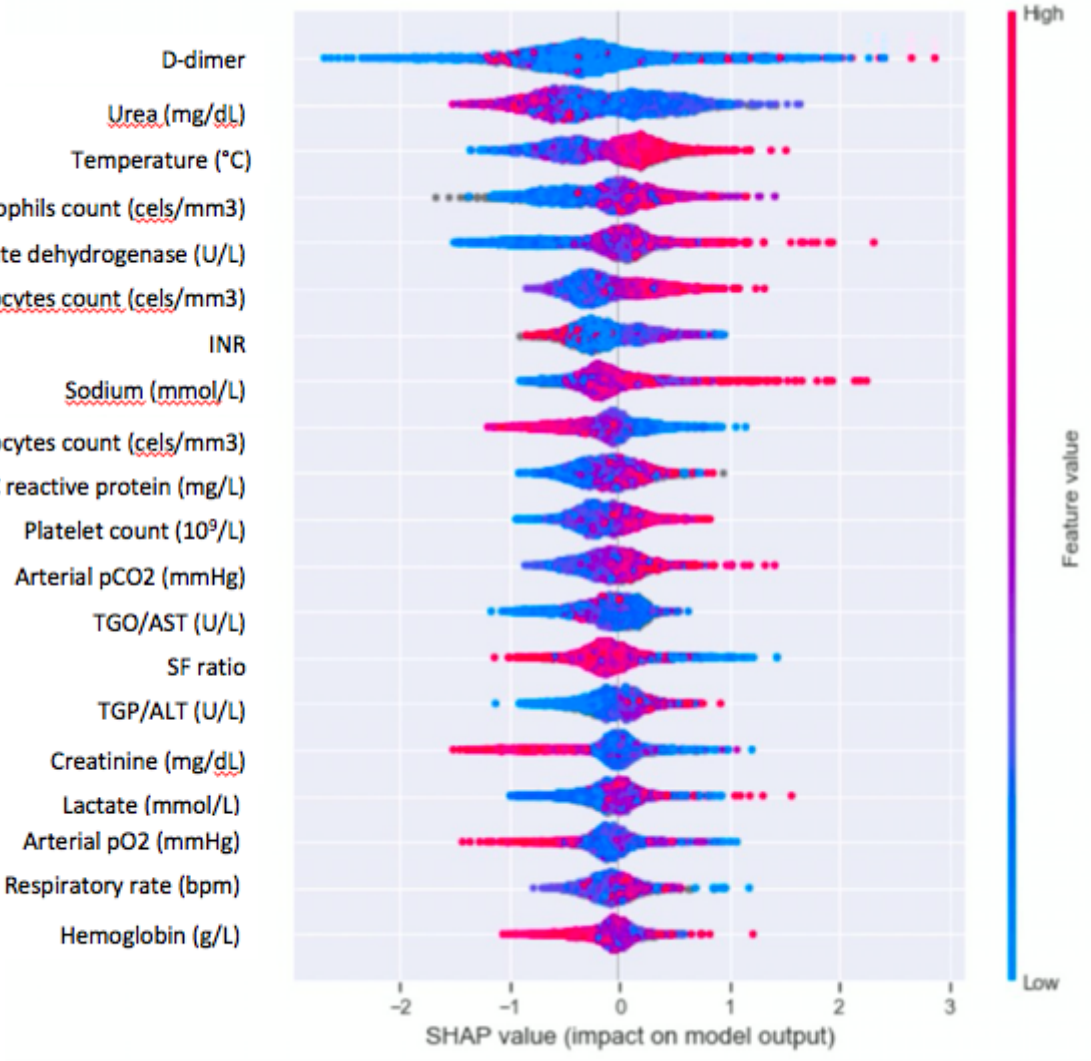

Figure 2

Impact of variables on the prediction of venous thromboembolism by machine learning.

Variables closer to the top are those with the highest correlation with the outcome. Red means probability of the outcome being predicted while blue means a smaller probability. Values to the right mean higher input values of the variable, while values to the left mean otherwise. FiO_2: fraction of inspired oxygen; INR: international normalized ratio; PCO2: arterial carbon dioxide partial pressure; PaO2: arterial oxygen partial pressure; SF ratio: peripheral oxygen saturation/FiO_2, TGO/AST: aspartate aminotransferase; TGP/ALT: alanine aminotransferase.

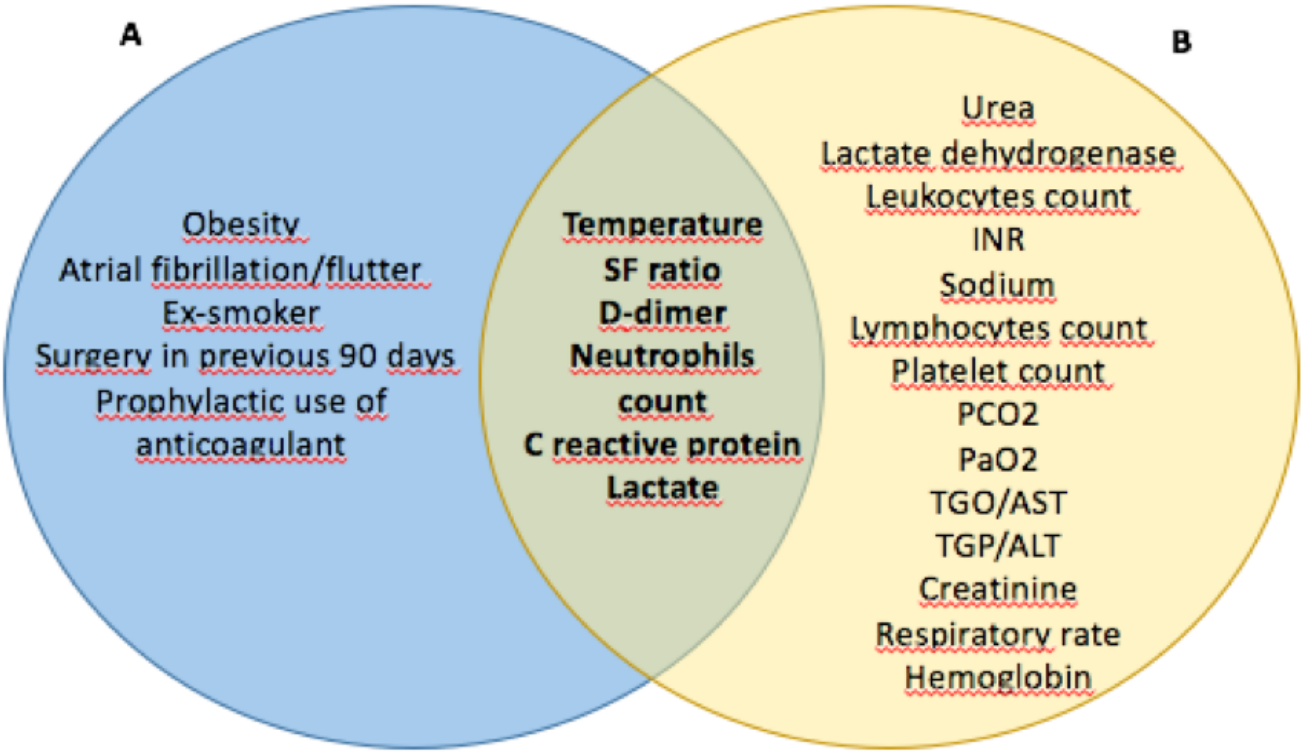

Figure 3

Comparison of VTE risk predictors identified by logistic regression analysis (A) and Machine Learning approaches (B).

SF ratio: oxygen saturation/inspired oxygen fraction; INR: international normalized ratio; PCO2: arterial carbon dioxide partial pressure; PaO2: arterial oxygen partial pressure; TGO/AST: aspartate aminotransferase; TGP/ALT: alanine aminotransferase.