Personalized Preventive Corticosteroid Medication Recommendation System for Postacute COVID-19 Treatment

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

16-09-2021 / 19-09-2021

CITATION

Myska, Vojtech; Burget, Radim; Genzor, Samuel; Mezina, Anzhelika; Mizera, Jan; Stybnar, Michal; et al. (2021): Personalized Preventive Corticosteroid Medication Recommendation System for Postacute COVID-19 Treatment. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.16616308.v1>

DOI

[10.36227/techrxiv.16616308.v1](https://doi.org/10.36227/techrxiv.16616308.v1)

Personalized Preventive Corticosteroid Medication Recommendation System for Postacute COVID-19 Treatment

Vojtech Myska¹, Radim Burget¹, Samuel Genzor¹,
Anzhelika Mezina¹, Jan Mizera¹, Michal Stybna¹, and Malay Kishore Dutta²

Abstract—Severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) is a pathogen responsible for one of the most massive pandemics in modern history. In certain patients with prolonged resorption of pulmonary involvement, organizing pneumonia develops which may lead to irreversible fibrotic damage of the lungs. According to some studies, corticosteroid treatment increases the chance for successful recovery. However, the distinction between patients who would benefit from corticotherapy, and those, who would recover spontaneously, is unclear. This paper introduces an artificial intelligence-based recommendation system for a personalised selection of patients for corticotherapy. In this study, 101 patients were enrolled. Every patient conducted an examination at the start and 3 months after the post-COVID treatment. It included physical examination, blood tests, functional lung tests, and health state based on the high-resolution computed tomography scan results. The proposed methodology recommends the application of CS and achieved balanced accuracy 86.18% whether a patient will recover or not without CS medication. The study identifies the most accurate algorithm and the most significant attributes for this prediction. This paper also introduces a simplified and easy human-interpretable model, which reaches 83.23% balanced accuracy.

With the proposed methodology, PF preventive treatment can be applied to those patients who will benefit from the treatment with high probability.

Index Terms—artificial intelligence, COVID-19, respiratory system

I. INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) is a pathogen responsible for coronavirus disease 2019 (COVID-19), a respiratory illness that is responsible for a significant rise in morbidity and mortality globally. Although the majority of patients have only a mild course of the disease without any further complications, in more severe cases, pneumonia, acute respiratory distress syndrome, or even multi-organ failure may develop [1], [2]. The virus usually enters the body via the respiratory epithelium (nasal cavity or lower airway). In its acute phase, it replicates down the respiratory tract, and in some cases, it also causes significant damage to the lungs, which leads to pneumonia. These cases may have moderate or severe clinical courses. The progress of the disease is depicted in the fig. 1. Commonly, the first symptoms after the onset are noticeable after 7 days. If the virus attack also hits the respiratory tract, it often leads to pneumonia or acute respiratory distress syndrome ARDS. ARDS can cause many different long-term sequelae, where many of them can also be severe.

This work was supported by the Ministry of the Interior of the Czech Republic, under Grant no. VI04000039.

Vojtech Myska, Radim Burget, Anzhelika Mezina are with the Brno University of Technology, Brno, Czech Republic (e-mail: {xmyska04, burgetrm, xmezin00}@vut.cz).

Samuel Genzor, Jan Mizera, and Michal Stybna, were with University Hospital of Olomouc, Olomouc, Czech Republic. (e-mail: {samuel.genzor, jan.mizera, michal.stybna}@fnol.cz).

Malay Kishore Dutta is with Centre for Advanced Studies, Dr. A. P. J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India (e-mail: mkd@cas.res.in).

As the post-acute phase is commonly considered a period after 3 or 4 weeks after the acute phase [3], [4]. In more severe cases, persistent pulmonary interstitial damage may develop. The most typical finding is organising pneumonia, which in certain cases may progress to irreversible fibrotic changes. The treatment of persistent lung interstitial damage in post-acute COVID-19 is still not standardised. According to the Czech national position document for the treatment of pulmonary involvement in COVID-19 [5], corticosteroid (CS) treatment is indicated in cases of prolonged resorption of pulmonary infiltrates as a prevention of the development of fibrosis. Similarly, Bieksiene et al., in their systematic review, states that appropriately timed CS therapy may be beneficial for selected patients [6].

This paper focuses on the treatment of lung interstitial damage in the post-acute phase of COVID-19. The treatment was indicated in concordance with the Czech national position document (Sova et al.) [5]. Only patients with persistent lung interstitial involvement (all of them had COVID-19 pneumonia in the acute phase) were indicated for CS treatment. The aim of this study is to establish a model that would aid in correctly distinguishing between post-COVID patients who would benefit from corticotherapy, and those, whose lung involvement would resolve without CS therapy, thus avoiding overtreatment.

The main contributions of this paper are:

- A unique public dataset FNOL.PulFib2021¹ of patients from post-acute treatment which contains 101 cases is released. Each patient is subjected to rigorous testing at the start of the post-COVID treatment and 3 months after, including information whether the patient benefited from CS treatment.
- The artificial intelligence-based (or also called evidence-based) methodology can predict with 86.18% balanced accuracy whether a patient will benefit from CS therapy. The most accurate algorithm is k -NN. We also identified the most valuable attributes used for the classification.
- Simplified interpretable decision tree-based model, which can be easily incorporated into the clinical practice. This model reaches 83.23% balanced accuracy.

The rest of the paper is structured as follows. Section II describes related work, which discusses possible post-acute COVID-19 complications and focuses on mortality, complications like pulmonary fibrosis and predictive machine learning approaches. It also identifies the areas that are still waiting for evidence-based studies. Section III describes the data used for the experiment and applied machine learning methods. Section IV presents results of the experiment and discusses their meaning and section V describes the limitations of this study. Section VI concludes the paper.

II. RELATED WORK

COVID-19 course can develop a systemic inflammatory response of the body, leading to lung injury or multisystem organ dysfunction. CS are known to have an anti-inflammatory effect and might prevent

¹Download FNOL.PulFib2021, <https://tinyurl.com/rcxb7c3x>

Fig. 1. Process of the COVID-19 disease including onset, acute and post-acute stages of the disease

or mitigate these deleterious effects. Unfortunately, opinions on their use in the treatment differ.

Myall et al. [7] found that approximately 39% patients with COVID-19 pneumonia remain symptomatic after the post-acute phase of the disease. These patients were further examined in their observational study using pulmonary function tests (PFT). In patients with a significant decrease in diffusion lung capacity for carbon monoxide (DLCO), high-resolution computed tomography (HRCT) was performed. CS treatment was indicated in 4,80 % (35 subjects) of all included patients. The radiology finding was consistent with the diagnosis of organising pneumonia, and histology verification was not performed. All subjects treated with CS showed an increase in DLCO (mean increase of 31,60 %) and significant improvement of subjective dyspnea. The main study limitation was the lack of a control group without CS treatment.

On the other hand, the post-hoc analysis of COVADIS study revealed contradictory results. The study aim was to establish the possible role of CS treatment in the post-acute phase of COVID-19 induced ARDS. 57 patients received CS, 290 were without treatment. The used dose of CS was 1 mg of prednisolon/kg, and no benefit of CS was found.

Especially, the last two COVID-19 phases are not standardized, so the physicians held professional discussions about the most appropriate way. In the case of moderate or severe course of the current phase of COVID-19 disease, the main goal is to keep the patient alive. There are several studies, which deal with the influence of CS on the patient's condition during the acute phase of COVID-19. According to [8] in selected critically ill patients, the application of CS may reduce 28-days mortality. The treatment seems beneficial, especially for patients with hypoxemia and C-reactive protein elevation above 200mg/l. A higher probability of therapeutic response was reported from patients with systematic CS treatment.

Another study [9] claims CS did not reduce or increase mortality within 42 days of admission to the intensive care unit. In this study, 422 patients were included.

The authors of [10] came to the completely opposite opinion. In this relatively extensive study published in 01/2021, which included 21 350 patients, authors concluded that overall mortality using CS was higher than without CS. It also concluded that an extended course

of CS beyond 10 days has no evidence of decreasing the risk of pulmonary fibrosis. On the other hand, 10 days course of CS had a positive impact on 28-days mortality. It must be emphasized that findings are short term, and PF would impact increased mortality in the horizon of years.

On the other hand, there is no yet evidence that long-term application of steroids can protect the patient against the potential onset of PF [10]. In the case of CS, one of the main issues is to determine the appropriate dosage. Work [11] deals with the analysis of the effect of the administered dose of CS on the patient's health. The study included 573 patients, the vast majority of whom consisted of men - 74.70% with median age 64 years. The majority of the patients (379 - 69.10%) were treated with standard doses (1.0-1.5mg/kg/day) and the rest received high doses (250-1000mg/day). Patients receiving higher doses of CS showed higher mortality and an increased risk of requiring mechanical ventilation. This phenomenon has been observed mainly in the elderly. It should be noted that the paper report bias in the experiment - the group of patients who were prescribed a higher dose consisted of elderly people who suffered much more often from some comorbidity and at the same time showed worse respiratory function. The authors recommend not to exceed the dose of CS in the amount of 1.0-1.5 mg/kg/day.

Despite a positive attitude towards the effectiveness of CS, it can not be prescribed to all patients, mainly given the strong side effects. Therefore, it is an effort to identify patients or individuals who will benefit from the treatment. It is the goal of several works that have focused on predicting the benefits of CS treatment. Work [12] proposes a machine learning-based approach which identifies patients for whom treatment with a CS or remdesivir will increase survival time. The method is based on the Gradient-boosted decision-tree model. They use data from 2364 patients acquired from 10 US hospitals for the experiment. Authors describe some limitations of their work, including retrospective character: it is unknown how treatment recommendations can impact prescribing practices and patient outcomes in clinical settings. According to the authors, this work is the first one, which applied the machine learning method to evaluate the effectiveness of treatment. There is also a work, which utilized more machine learning algorithms [13]. The objective of the experiment was to evaluate the response to CS therapy on

TABLE I

INFORMATION ABOUT DATASET AND DEMOGRAPHY OF PATIENT, WHO WERE INCLUDED INTO THE STUDY

DEMOGRAPHIC & HABITS						
Attributes	Number of patients with CS			Number of patients without CS		
Total number	54			47		
Male	34			25		
Female	20			22		
Smoker	4			2		
Ex-smoker	10			4		
Non-smoker	36			33		
THERAPY						
Attributes	Number of patients with CS			Number of patients without CS		
	Yes	No	N/A	Yes	No	N/A
Hospitalized	43 (80%)	10 (18%)	1 (2%)	35 (74%)	12 (26%)	0
Oxygen	38 (70%)	12 (22%)	4 (7%)	26 (55%)	13 (28%)	8 (17%)
Remdesivir	2 (4%)	48 (89%)	4 (7%)	1 (2%)	46 (98%)	0
PERSISTENT HEALTH ISSUES						
Attributes	Number of patients with CS			Number of patients without CS		
	Yes	No	N/A	Yes	No	N/A
Dyspnea	39 (72%)	15 (28%)	0	26 (55%)	21 (45%)	0
Cough	22 (41%)	32 (59%)	0	12 (26%)	35 (74%)	0
Fatigue	10 (19%)	44 (81%)	0	13 (28%)	34 (72%)	0
Olfactory loss	1 (2%)	53 (98%)	0	5 (11%)	42 (89%)	0
Gastrointestinal problems	6 (11%)	48 (89%)	0	10 (21%)	37 (79%)	0
COMORBIDITY						
Attributes	Number of patients with CS			Number of patients without CS		
	Yes	No	N/A	Yes	No	N/A
EAP	1 (2%)	53 (98%)	0	4 (9%)	43 (91%)	0
DM	15 (28%)	39 (72%)	0	9 (19%)	38 (81%)	0
ICHs	3 (6%)	51 (94%)	0	2 (4%)	45 (96%)	0
COVID-19 TESTING						
Attributes	Number of patients with CS			Number of patients without CS		
	Yes	No	N/A	Yes	No	N/A
IgM (qualitatively)	41 (76%)	12 (22%)	1 (2%)	41 (87%)	5 (11%)	1 (2%)
IgG (qualitatively)	52 (96%)	1 (2%)	1 (2%)	45 (96%)	1 (2%)	1 (2%)

a dataset of 666 patients. The used ML algorithms are Logistic Regression (LR), Support Vector Machine (SVM), Gradient Boosted Decision Tree (GBDT), k -Nearest Neighbor (k -NN), and Neural Network (NN). The results seem to be promising: AUC of 0.81 in the internal validation set and 0.85 in the external validation set. Also, an unsupervised machine learning approach, clustering, was used to evaluate response to CS therapy and the association between CS treatment and mortality [14]. According to the results, CS showed a positive effect on the survival of critically ill patients with the hyper-inflammatory phenotype.

Approximately one in three patients with a symptomatic course of COVID-19 disease suffers from at least one health sequelae even 12 weeks after infection [15]. Similar to the acute phase, CS therapy is also used in this phase. They have been used successfully to alleviate olfactory loss [16], [17]. Nevertheless, one of the most serious health sequelae is pulmonary fibrosis. It may significantly reduce the patient's quality of life and can reduce the length of survival.

To summarise the studies, there is not a clear opinion on the CS treatment. Some of the studies report positive effects [8], [7] but there are also many studies claiming the opposite [10]. Unfortunately, most of the studies are often not so easy to easily compare between each other. The course of COVID-19 disease can be divided into three phases: onset, acute and post-COVID phase (see fig. 1), and they compare results in different periods, some of them just after the acute phase, some of them in +1 month, +3 months or +6 months.

Another issue is the metric that the studies usually use for comparison. They commonly compare the results according to the

morbidity rate. It is an objective and relatively easily accessible value. However, COVID-19 has not only an impact on morbidity. Also, a significant percentage of patients have impaired quality of life, which is more challenging to measure objectively. With a high probability, the quality of life will also be reflected in the increased morbidity in future. However, this is expected to be reflected in the long-term horizon (5+ years).

There are still many open questions regarding CS treatment. Including in which cases their benefits outweigh their negative side effects (i.e. when they should be applied), in which phase they should be applied, how much of the medicament should be applied and how long the medication should last.

Another limitation is the level of detail the studies use about each patient. Most of them use just a basic set of information (age, sex, height, sometimes selected symptoms, etc.). However, more information is commonly available at the start of post-COVID treatment (e.g. blood tests, functional lung tests, symptoms and patient status). This basic information cannot be used for a qualified decision regarding future optimal treatment.

Currently, there is no study devoted to post-COVID CS treatment. Study [13] introduces an AI-powered acute-phase recommendation system. However, it is not suitable for the post-acute phase since the patient is not yet in an immediate life-threatening condition.

III. METHODOLOGY

This section describes overall methodology used in this experiment. In particular it describes the dataset and which patients were enrolled into the experiment (see section III-A), the whole experiment and

optimization techniques used (see section III-B) and metrics used for the evaluation (see section III-C).

A. Dataset

The experiment enrolled in total 101 patients. All the patients were infected by COVID-19 disease with lung involvement of different severity. The occurrence of PF to patients with no pneumonia is quite rare, so CS treatment is not considered in those cases.

Each patient was subjected to initial examination at the start of post-COVID treatment, which is approximately 3 weeks after first visible symptoms (see fig. 1). The data collected for the experiment include physical examination (age, body mass index, is a smoker, presence of other commodities, etc.), information related to acute-phase treatment, pulmonary function tests data and blood tests data (e.g. immunoglobulins IgG, IgM). The patients are divided into two groups, 1) to who received CS therapy (54) and who did not (47). The data also includes the result of the treatment after approx. 3 months, i.e. whether the decision to apply CS treatment or not was correct. Detailed information about the dataset is available online.

To reduce the bias and the risk of overfitting, some attributes from the dataset were removed. The possible bias can occur because of not a random selection of the patients, whether they received CS or not. The complete dataset was released, including detailed information about each attribute to make it easier to integrate with other datasets in future. In this paper, only those parameters, which were identified as significant, are explained. Demographical information about the dataset is shown in Table I. The data from spirometry contains for each attribute values marked "(pred)", "(abs)" and "(%pred)", where "(pred)" stands for predicted normal value based on height, age, etc., "(abs)" stand for actual absolute measured value. For the the analysis only the "(%pred)" values were used, since they reflects the best body composition of each patient.

From the point of the statistical analysis, it would be optimal to select the patients for preventive CS treatment on a random basis. Since the main priority is patients' health and to provide them with the best possible treatment, this would be unethical and not possible. This can be connected with possible bias in separating the patients into these two groups. According to results of Mann-Whitney U-test (U-value is 299397, p -value – 0), generally, the two groups of patients (which have got CS therapy and which have not got CS) are different. However, more detailed testing for each qualitative feature is shown in Table II. It can be seen that just some attributes have low p -value: Olfactory loss (0.03), Gastrointestinal problems (0.08), pulmonary artery embolization (0.06) and IgM (0.07), which indicate that samples from these attributes are different. However, considering the typical threshold of p -value (≤ 0.05), only one attribute matches this condition.

An overview of selected qualitative parameters is shown in table I. It contains information about both groups of patients: 1) who received and 2) who did not receive CS treatment. Percentages from the total number of respective groups for each parameter are also provided. The Czech national position document was used for indication of CS treatment in all patients. According to this national guideline the patients with persistent pulmonary interstitial damage induced by COVID-19 pneumonia are indicated for the CS therapy. Recommended dose is 0,5 mg of prednisolon per kg of body weight (with maximum dose of 40 mg) for 2 weeks, with subsequent 4 weeks of 20 mg of prednisolon treatment and further gradual tapering of the doses until withdrawal. Bacterial (or other) superinfection was excluded prior CS therapy. Moreover, possible signs of pre-existing interstitial lung disease were assessed - either by the typical picture on HRCT scans, or by biopsy in uncertain cases. Only subjects with clear post-COVID lung damage were involved in the analysis.

TABLE II
MANN-WHITNEY U-TEST AND P-VALUE TESTS FOR SELECTED ATTRIBUTES

THERAPY		
Attributes	U-value	P-value
Hospitalized	1197.5	0.75
Oxygen	1100.0	0.91
Remdesivir	1157.0	0.96
PERSISTENT HEALTH ISSUES		
Attributes	U-value	P-value
Dyspnea	1054.5	0.96
Cough	1076.0	0.95
Fatigue	1385.0	0.14
Olfactory loss	1380.5	0.03
Gastrointestinal probl.	1398.0	0.08
COMORBIDITY		
Attributes	U-value	P-value
EAP	1353.5	0.06
DM	1159.5	0.84
ICHS	1252.5	0.62
COVID-19 TESTING		
Attributes	U-value	P-value
IgM (qualitatively)	1415.5	0.07
IgG (qualitatively)	1261.5	0.56

During the selection of the treatment we tried to minimise possible bias as much as possible.

B. Experiment

The experiment was conducted separately on the two subgroups. The first subgroup of patients received CS therapy, and the second subgroup did not. On the subgroup of patients who did not receive the CS medication, one extra experiment was done, which objective was to create a simplified interpretable model. We examined in total six machine learning algorithms. The selected algorithms are mainly suitable for the analysis of small datasets. They are linear regression [18, pp. 208-210], k -Nearest Neighbours (k -NN) [18, pp. 208-210], decision tree [18, Ch. 7], XGBoost [18, pp. 192-193], Random Forest [18, pp. 187-189] and Support Vector Machine (SVM) [18, pp. 145-155] with radial basis function (rbf) as a kernel. The model was statistically evaluated using leave-one-out cross-validation [18, pp. 75-77]. Hyperparameters of each machine learning algorithm were tuned using the grid search, and features were pre-selected using Forward Feature Selection (FFS) method [18, pp. 100-102].

C. Evaluation metrics

To evaluate and compare the results from experiments described in the previous section, we selected several metrics. These metrics include: accuracy (see eq. 1), balanced accuracy (see eq. 5), F1 score (see eq. 2), sensitivity (see eq. 3), and specificity (see eq. 4). For a description of the following metrics, there were used the following abbreviations: TP – number of true positive cases, TN – number of true negative cases, TN – number of true negative and FN – number of false negative cases.

Accuracy measures the ratio of correctly predicted labels over total number of evaluated samples [19]:

$$Acc = \frac{TN + TP}{TP + TN + FP + FN} \quad (1)$$

F_1 score is a combination of precision and recall metrics, which capture properties of them both [19]:

$$F_1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (2)$$

TABLE III

ACCURACY OF PREDICTING WHETHER PATIENT WILL RECOVER WITH CS. CONFIDENCE LEVEL USED TO COMPUTE CI WAS 95%.

Algorithm	Accuracy	Balanced Accuracy	F1-score	Sensitivity	Specificity
Logistic regression	77.78% (CI: 11.09%)	50.00% (CI: 13.34%)	0.00%	0.00 %	100.00%
k-NN	79.63% (CI: 12.18%)	66.08% (CI: 13.29%)	47.62%	41.67%	90.48%
Decision Tree	81.48% (CI: 10.48%)	61.31% (CI: 12.99%)	37.50%	25.00%	97.62%
XGBoost	88.89% (CI: 10.33%)	75.00% (CI: 12.49%)	66.67%	50.00%	100.00%
Random Forest	77.78% (CI: 11.09%)	50.00% (CI: 13.34%)	40.00%	25.00 %	100.00%
SVM (rbf)	83.33% (CI: 9.94%)	62.50% (CI: 12.91%)	50.00%	33.33%	100.00%

Sensitivity measures, how the model can correctly predict positive samples [19]:

$$Sen = \frac{TP}{TP + FN}. \quad (3)$$

Specificity measures, how the model can correctly predict negative samples [19]:

$$Spe = \frac{TN}{TN + FP}. \quad (4)$$

Balanced Accuracy is average value between sensitivity and specificity [20]:

$$Acc_{BAL} = \frac{Spe + Sen}{2}. \quad (5)$$

IV. RESULTS AND DISCUSSION

The objective of this experiment is to show a refined data-driven methodology, which will identify those patients more accurately. In particular, it will 1) identify those patients who were pre-selected for medication but will most probably not benefit from the treatment (their condition won't be improved) or 2) initiate the medication to those patients who were not pre-selected to the treatment, and they will not recover themselves. The results for groups who were pre-selected for this study are described in the following sections.

A. Analysis of patients with Corticotherapy

The first analysed subgroup contained patients who were according to the existing practice recommended for CS therapy. Not every patient benefited from the treatment, and the objective was to recognise who will not benefit from the treatment and who only will have side

effects from the treatment. According to the current Czech COVID-19 positional document [5], only patients with a high probability they will benefit were selected for the treatment. The number of persons included in the study was 54. In total, 42 benefited and 12 did not (their health status did not improve after 3 months after CS medication). Cases of patients who had no positive response to CS was quite limited, and therefore, the dataset is significantly imbalanced.

Results of the experiment is shown in table III. The best possible balanced accuracy was $75.00\% \pm 12.49\%$. The confusion matrix for the best performing model (XGBoost, see table III) is shown in Table IV. The parameters used for training the model was 50 trees, maximal tree depth limit was set to 5, the minimum number of rows to assign to the terminal node was 10 and number of bins for its histograms was 20.

These results are worse than the initial bias in the dataset (88.89%). It predicted correctly 48 patients and 6 incorrectly. The accuracy of the experiment was 88.89%, and balanced accuracy was 75.00%. In all 6 cases, it predicted the patients would recover themselves without CS medication, but they, in fact, did not. The overall class precision of "will recover" is 87.50%, and the precision for the class "Will not recover without medication" is 100.00%.

The most important attributes used for the prediction were identified: the sex of the patient and RV(%pred). Unfortunately, from the clinical point of view, these attributes seem not to be very reliable, and decisions based on them seem to be not much trustworthy. Instead, it seems that the algorithm recognized some initial bias in the data. As a result, we conclude that all the patients who were included by the Czech COVID-19 positional document [5] should receive CS therapy. The success rate is relatively high $42 + 6/54 = 88.89\%$ accuracy.

TABLE IV

CONFUSION MATRIX, XGBOOST, MEDICATION WITH CS.

		Ground truth		Class precision
		Will recover	Will not recover	
Predicted	Will recover	42 (70.37%)	6 (12.96%)	87.50% (CI: 9.36%)
	Will not recover	0 (0.00%)	6 (9.26 %)	100.00% (CI: --%)
	Class recall	100.00% (CI: --%)	50.00% (CI: 28.29%)	

B. Analysis of patients without Corticotherapy

The objective was to predict who of the patients will recover themselves without the medication and who will have serious health complications even after 3 months period after the onset of the disease. The analysed subgroup contained 47 patients. None of them received the CS medication. Number of patients who recovered without CS medication was 17, and who did not recover 30. According to the Czech positional document [5], the selection of the patients was not random, and the medication followed internal selection practice. It is based on the priority not to hurt patients with possible unnecessary medication (with harmful side effects).

Results are shown in Table V. The most significant attributes, which were identified from the analysis, were patient's age,

TABLE V

ACCURACY OF PREDICTING THE PATIENT WILL RECOVER WITHOUT CS. CI STANDS FOR CONFIDENCE INTERVAL, CONFIDENCE LEVEL IS 95%.

Algorithm	Accuracy	Balanced Accuracy	F1-score	Sensitivity	Specificity
Logistic regression	85.11% (CI: 10.18%)	81.96% (CI: 10.98%)	88.89%	93.33%	70.59%
<i>k</i> -NN	87.23% (CI: 9.54%)	86.18% (CI: 9.87%)	90.00%	90.00%	82.35%
Decision Tree	76.60% (CI: 12.10%)	75.30% (CI: 12.33%)	81.36%	80.00%	70.59%
XGBoost	82.98% (CI: 10.74%)	76.47% (CI: 12.13%)	88.24%	100.00%	52.94%
Random Forest	85.11% (CI: 10.18%)	83.24% (CI: 10.68%)	88.52%	90.00%	76.47%
SVM (rbf)	13.08% (CI: 7.98%)	58.83% (CI: 14.07%)	81.08%	100.00%	17.65%

KCO_SB(%pred), presence of persistent fatigue (after 2-3 weeks after the acute phase), ERV(%pred) and blood test on the Neutrophils level.

The best performing model was *k*-NN, which reached balanced accuracy 86.18% with confidence interval 9.87% and F_1 -score 82.35%. The algorithm was configured as follows: the number of trees was 50, maximal depth of the trees was limited to 5 levels, the learning rate was set to 0.01. The confusion matrix about the XGBoost results is shown in Table VI. In 42 cases, it was correctly predicted the patient would recover with the CS therapy. In 10 cases, it was correctly predicted the patients would not recover without CS medication. There was no false prediction about the patients who will not recover. For two patients, it was predicted they would recover, but in fact, they did not.

The original Czech positional document [5] prioritises not to hurt patients with unnecessary CS medication. This study shows that with relatively high accuracy it can be predicted who of the patients will benefit from the CS and who will not. The *k*-NN model should be preferred to for this purposes.

C. Explainable Recommendation Approach

Although the best results were achieved with *k*-NN, the methodology of how the decision is being made is relatively complex and hard to understand by a human. In order to make it more transparent, we prepared another model, which is based on the decision tree that can be easily interpreted by humans and can be used in clinical practice more easily. This model depends on in total three features, including the age of the patient (feature in the dataset is called “age”), peak expiratory flow (PEF(%pred)), and P.INR. Its accuracy is 85.11%, balanced accuracy is 83.24%, F_1 -score is 88.52%, sensitivity is

90.00%, and specificity is 76.47%. Details about confusion matrix is depicted in the Table VII.

The resulting decision tree model is depicted in figure IV-C. This model was created based on data obtained from 47 patients. In case the age of the patient is over 76, they should not receive corticotherapy. This recommendation has an accuracy 83.33% with a confidence interval (CI) of 10.68%. In case the patient is older than 76 and Prothrombin Time and International Normalized Ratio (P.INR) from blood test is more significant than 1.055, then peak expiratory flow from the functional pulmonary test (PEF(%pred)) should be evaluated. If it is higher than 61.83, then corticotherapy is recommended. The confidence of this decision is 95.83 with an interval +/- 15.03%. Otherwise, the medication is not recommended with confidence of 80.00% and a confidence interval 35.06%. In case P.INR value is between 1.055 and 1.095, corticotherapy is recommended. In case the P.INR is higher than 1.055, then corticotherapy is not recommended. Confidence intervals were computed for confidence level 95%.

The explainable model reaches significantly worse accuracy than the model based on the *k*-NN model. Its advantage is transparency, so the clinicians know exactly reason to involve or exclude the patient into the CS medication. When compared to the Czech positional document [5], the methodology still reaches interesting results and the explainable model should be preferred.

V. LIMITATIONS

The number of patients included in this study (101) is limited. Especially the CS medicated subgroup is significantly imbalanced. This implies that the machine learning models created could be biased and overfitted. Because of this, further validation of the proposed

TABLE VI

CONFUSION MATRIX, *k*-NN, WITHOUT CS MEDICATION.

		Ground truth		Class precision
		Will recover	Will not recover	
Predicted	Will recover	14 (34.00%)	3 (6.00%)	82.35% (CI: 18.12%)
	Will not recover	3 (6.00%)	27 (54.00%)	90.00% (CI: 10.74%)
	Class recall	82.35% (CI: 18.12%)	90.00% (CI: 10.74%)	

TABLE VII

CONFUSION MATRIX, EXPLAINABLE MODEL, WITHOUT CS MEDICATION

		Ground truth		Class precision
		Will recover	Will not recover	
Predicted	Will recover	27 (57.45%)	4 (8.51%)	87.10% (CI: 9.58%)
	Will not recover	3 (6.38%)	13 (27.66%)	81.25% (CI: 11.16%)
	Class recall	90.00% (CI: 8.58%)	76.47% (CI: 12.13%)

Fig. 2. Explainable decision tree model based on Expiratory reserve volume (ERV) (%), age of patient, remaining fatigue and neutrophils percentage from blood test.

methodology is needed in follow-up works. Also, future studies should extend the dataset to create a more robust algorithm. We encourage the other teams to validate the proposed model with their data from their own clinical practice. We also encourage other teams to share the data. Merged datasets can lead to more complex and reliable models.

The results obtained could be biased for a specific group of people. For example, race and ethnicity can influence the variance of the cohort. The data samples were obtained from the Olomouc region, Czech Republic. In some studies, it has been reported that different races can be more vulnerable to the COVID-19 disease [21] so possibly also different parameters could be of other importance in different regions. Validation of whether the findings of this study also fit other races is needed. On the other hand, this study tried to prefer the attributes which are stable and as independent of other factors as possible.

Also, there might be possible selection bias in patients indicated to the examination by the clinicians as the majority of patients sent to the pulmonary department were symptomatic. Another limitation is that only a small proportion of the patients had a lung biopsy, and the treatment was based mainly on clinical status, radiology findings, and pulmonary function tests. Patients with known pre-existing interstitial lung disease had been excluded from the analysis. However, there might be a theoretical chance of possible pre-existing interstitial lung disease in a small proportion of the patients, as in some patients was the treatment of COVID-19 infection was the first contact with the health care system. All the patients remain in our follow-up at least for the next 2 years to confirm the durable resolvment of the lung damage.

It is also not clear what the responses of the human body of people with several comorbidities are. Moreover, the comorbidities could influence the physiological parameters, which could distort the patterns.

VI. CONCLUSION

This paper introduces the first artificial intelligence-based method that predicts who of the patients in the post-COVID phase will benefit from CS and who will not. The data used in this study includes clinical, functional and imaging information about each patient. In

total, 110 patients were enrolled on the study. Every patient suffered from COVID-19 pneumonia in their acute phase. Based on the current Czech national position document for diagnostics and treatment of the lung involvement by COVID-19 [5], these 101 patients received (54) or did not receive CS medication (47). The experiment was performed separately to the 1) subgroup of patients who received CS, 2) who did not receive CS ($86.18\% \pm 9.87\%$ balanced accuracy) and 3) the third experiment was repeated once again with as in 2) but with a simplified model, which can be more easily included into clinical practice. The most significant attributes were identified: EVR(%pred), level of Neutrophils from the blood test, patient age, and persistence of fatigue. The complete dataset was published online.

We conclude that the Czech position document [5] regarding selecting patients for the CS treatment should be used as it is (accuracy 88.89%). The AI-based model reached only 75.00% of balanced accuracy and after an in-depth analysis of the behaviour of the algorithm and consulting the features with clinical experts was concluded that the AI-based model is not reliable.

On the other hand, the patients who are recommended under the Czech position document [5] for no medication, should follow recommendations from k -NN algorithm with the potential for an interesting increase in efficiency of the treatment. For the clinicians who prefer interpretable models, a simplified interpretable model was introduced (see fig. IV-C). The model was also validated by clinicians, and it was concluded the principle makes sense also from clinical practice and are based on attributes that are accessible.

Future work should focus on extending the experiment with more data, which should be of the highest priority. Especially merging the data from various sources has the potential to reach interesting improvements. We encourage other teams to publish the data so the research community can benefit from it. Also, experiments with imaging approaches, including X-ray or HRCT might lead to interesting results. Unfortunately, this would be a bit data demanding. Although this study also included comorbidities, their results cannot be considered statistically significant due to their relatively rare occurrence. Hopefully, they will help in some follow-up studies.

REFERENCES

- [1] Z. M. Afshar, S. Ebrahimpour, M. Javanian, V. Koppolu, V. K. R. Vasigala, A. H. Hasanpour, A. Babazadeh, *et al.*, "Coronavirus disease 2019 (covid-19), mers and sars: Similarity and difference," *Journal of Acute Disease*, vol. 9, no. 5, p. 194, 2020.
- [2] D. V. Parums, "Long covid, or post-covid syndrome, and the global impact on health care," *Medical Science Monitor*, vol. 27, 2021.
- [3] N. M. Dixit, A. Churchill, A. Nsair, and J. J. Hsu, "Post-acute covid-19 syndrome and the cardiovascular system: What is known?," *American heart journal plus: cardiology research and practice*, p. 100025, 2021.
- [4] A. Nalbandian, K. Sehgal, A. Gupta, M. V. Madhavan, C. McGroder, J. S. Stevens, J. R. Cook, A. S. Nordvig, D. Shalev, T. S. Sehrawat, *et al.*, "Post-acute covid-19 syndrome," *Nature medicine*, vol. 27, no. 4, pp. 601–615, 2021.
- [5] M. Sova, M. Doubková, L. Solichová, M. Šterclová, and S. Genzor, "Treatment of pulmonary involvement of patients after covid-19 (coronavirus disease 2019) - position document of the czech pneumological and phthisiological society," *Czech pneumology and phthisiology society*, 2020. [Online; accessed 31-September-2021, <http://www.pneumologie.cz/guidelines/>].
- [6] K. Bieksiene, J. Zaveckiene, K. Malakauskas, N. Vaguliene, M. Zemaitis, and S. Miliuskas, "Post covid-19 organizing pneumonia: The right time to interfere," *Medicina*, vol. 57, no. 3, p. 283, 2021.
- [7] K. J. Myall, B. Mukherjee, A. M. Castanheira, J. L. Lam, G. Benedetti, S. M. Mak, R. Preston, M. Thillai, A. Dewar, P. L. Molyneaux, *et al.*, "Persistent post-covid-19 interstitial lung disease. an observational study of corticosteroid treatment," *Annals of the American Thoracic Society*, vol. 18, no. 5, pp. 799–806, 2021.
- [8] R. C. Group, "Dexamethasone in hospitalized patients with covid-19," *New England Journal of Medicine*, vol. 384, no. 8, pp. 693–704, 2021.
- [9] Á. Estella, J. G. Garmendia, C. de la Fuente, J. M. Casas, M. E. Yuste, R. A. Villar, M. A. Estechea, L. Y. Mateos, M. C. Bulnes, A. Loza, *et al.*, "Predictive factors of six-week mortality in critically ill patients with sars-cov-2: A multicenter prospective study," *Medicina Intensiva*, 2021.
- [10] G. P. Mishra and J. Mulani, "Corticosteroids for covid-19: the search for an optimum duration of therapy," *The Lancet. Respiratory Medicine*, vol. 9, no. 1, p. e8, 2021.
- [11] E. Monreal, S. S. de la Maza, E. Natera-Villalba, Á. Beltrán-Corbellini, F. Rodríguez-Jorge, J. I. Fernández-Velasco, P. Walo-Delgado, A. Muriel, J. Zamora, A. Alonso-Canovas, *et al.*, "High versus standard doses of corticosteroids in severe covid-19: a retrospective cohort study," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 40, no. 4, pp. 761–769, 2021.
- [12] C. Lam, A. Siefkas, N. S. Zelin, G. Barnes, R. P. Dellinger, J.-L. Vincent, G. Braden, H. Burdick, J. Hoffman, J. Calvert, *et al.*, "Machine learning as a precision-medicine approach to prescribing covid-19 pharmacotherapy with remdesivir or corticosteroids," *Clinical therapeutics*, 2021.
- [13] Y. Gao, X. Xiong, X. Jiao, Y. Yu, J. Chi, L. Chen, S. Li, and Q. Gao, "Development and validation of a machine learning model for prediction of response to corticosteroid therapy in covid-19 patients," *Available at SSRN 3834263*, 2021.
- [14] H. Chen, J. Xie, N. Su, J. Wang, Q. Sun, S. Li, J. Jin, J. Zhou, M. Mo, Y. Wei, *et al.*, "Corticosteroid therapy is associated with improved outcome in critically ill coronavirus disease 2019 patients with hyperinflammatory phenotype," *Chest*, 2020.
- [15] M. Whitaker, J. Elliott, M. Chadeau-Hyam, S. Riley, A. Darzi, G. Cooke, H. Ward, and P. Elliott, "Persistent symptoms following sars-cov-2 infection in a random community sample of 508,707 people," *medRxiv*, 2021.
- [16] S.-D. Le Bon, D. Konopnicki, N. Pisarski, L. Prunier, J. R. Lechien, and M. Horoi, "Efficacy and safety of oral corticosteroids and olfactory training in the management of covid-19-related loss of smell," *European Archives of Oto-Rhino-Laryngology*, pp. 1–5, 2021.
- [17] S. K. Tousserkani and A. Ayatollahi, "Oral corticosteroid relieves post-covid-19 anosmia in a 35-year-old patient," *Case Reports in Otolaryngology*, vol. 2020, 2020.
- [18] H. Rhys, *Machine Learning with R, the tidyverse, and mlr*. Simon and Schuster, 2020.
- [19] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.
- [20] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, 2020.
- [21] E. J. Williamson, A. J. Walker, K. Bhaskaran, S. Bacon, C. Bates, C. E. Morton, H. J. Curtis, A. Mehrkar, D. Evans, P. Inglesby, *et al.*, "Factors associated with covid-19-related death using opensafely," *Nature*, vol. 584, no. 7821, pp. 430–436, 2020.