Article

Machine Learning and BMI Improve the Prognostic Value of GAP Index in Treated IPF Patients

Donato Lacedonia [1], Cosimo Carlo De Pace [1,*], Gaetano Rea [2], Ludovica Capitelli [3], Crescenzio Gallo [4], Giulia Scioscia [1], Pasquale Tondo [1] and Marialuisa Bocchino [3]

1 Department of Medical and Surgical Sciences, University of Foggia, 71121 Foggia, Italy
2 Department of Radiology, Monaldi Hospital, AO dei Colli, 80131 Naples, Italy
3 Respiratory Medicine Unit, Department of Clinical Medicine and Surgery, Federico II University of Naples, 80131 Naples, Italy
4 Department of Clinical and Experimental Medicine, University of Foggia, 71121 Foggia, Italy
* Correspondence: cosimo.depace@unifg.it; Tel.: +39-088-173-3037

Abstract: Patients affected by idiopathic pulmonary fibrosis (IPF) have a high mortality rate in the first 2–5 years from diagnosis. It is therefore necessary to identify a prognostic indicator that can guide the care process. The Gender-Age-Physiology (GAP) index and staging system is an easy-to-calculate prediction tool, widely validated, and largely used in clinical practice to estimate the risk of mortality of IPF patients at 1–3 years. In our study, we analyzed the GAP index through machine learning to assess any improvement in its predictive power in a large cohort of IPF patients treated either with pirfenidone or nintedanib. In addition, we evaluated this event through the integration of additional parameters. As previously reported by Y. Suzuki et al., our data show that inclusion of body mass index (BMI) is the best strategy to reinforce the GAP performance in IPF patients under treatment with currently available anti-fibrotic drugs.

Keywords: idiopathic pulmonary fibrosis; GAP index; machine learning; mortality; body mass index; nintedanib; pirfenidone

1. Introduction

Idiopathic pulmonary fibrosis (IPF) [1,2] is a rapidly progressive interstitial lung disease of unknown cause, which usually affects men over 65 years of age with a history of smoking. Diagnosis of IPF is often complex and based on many specialists' experience, not only pulmonologists but also radiologists and histopathologists. High-resolution computed tomography (HRCT) findings, especially after the update in the international guidelines, are crucial in the diagnostic approach, and in case of a radiological usual interstitial pneumonia (UIP) pattern, but also likely in a probable UIP pattern, the diagnosis of idiopathic pulmonary fibrosis is made. A multidisciplinary approach has become more and more relevant, especially in uncertain radiological and clinical conditions and in the absence of histopathologic samples, which requires a multi-specialist presence as evidenced by the newest published guidelines [3].

Symptoms are commonly a dry cough, progressive dyspnea, fatigue, and a progressive decline in capability and independence during daily activities. The most characteristic sign identifiable during the physical examination of the patient is the "velcro-type" crackle during auscultation due to traction bronchiectasis but also, in the advanced stage of the disease, the clubbing of the fingernails due to chronic hypoxemia.

Idiopathic pulmonary fibrosis is associated with a poor prognosis, with high mortality rates in the first 2–5 years from diagnosis in untreated cases. Pharmacological treatment is based exclusively on two approved drugs, which are unfortunately useful only in slowing the evolution of the disease. In 2011, pirfenidone was the first treatment approved [4,5] for idiopathic pulmonary fibrosis for its favorable benefit-risk ratio due to trial results showing

a decrease in disease progression in lung function, exercise tolerance, and improvement in survival compared to a placebo. It is an orally bioavailable antifibrotic agent and has also anti-inflammatory and antioxidant effects. After four years, nintedanib was approved [6,7], the second and, currently, the last pharmacological treatment. It is an intracellular inhibitor of tyrosine kinase that causes an inhibitory effect on fibroblast proliferation and differentiation, delivering significant effects in reducing lung function decline. Adverse effects are possible after the administration of both drugs. Pirfenidone is more often associated with photosensitivity and skin rash while nintedanib is most frequently associated with diarrhea. The two antifibrotic drugs are also associated with liver toxicity, which is why treated patients need to have blood samples taken for transaminase control.

Many trials are in progress to develop newer and more effective drugs to not only slow but also stop the degeneration caused by fibrosis, but until now, unfortunately, there are no effective treatments, which is why the evaluation of the prognosis remains an actual and debated question, as we can also see from the continuous research into more complex biomarkers [8].

Psychological assistance and rehabilitation are central in the treatment of patients affected by IPF, not only due to the poor prognosis but also due to the constant and progressive worsening of the quality of life.

Depression and anxiety are indeed very common in patients affected by IPF [9], regardless of survival or hospitalization, highlighting the need for better detection and treatment of this condition.

An interesting finding is also that patients affected by idiopathic pulmonary fibrosis are often associated with gastro-esophageal reflux, which can cause a worsening of the disease but also several other respiratory diseases such as laryngeal disorders, sinusitis, and chronic coughing [10]. The most frequently used drugs to treat gastro-esophageal reflux are proton pump inhibitors, which seem to be related to a worsening of the psychological status of treated patients [11].

Furthermore, respiratory rehabilitation is a fundamental part of the treatment of IPF, which is useful both to slow the decline of respiratory function and to improve the quality of life [12].

A mention must be made, given the pandemic situation and due to their frail clinical condition, of the importance of early treatment of patients affected by IPF that react positively to the nasopharyngeal swab for research into SARS-CoV-2 [13] and the importance of the creation of adequate paths to post-covid rehabilitation [14] and follow-up [15].

In selected patients, a viable treatment option is lung transplantation [16], which can be unilateral or bilateral, although a recent study [17] has demonstrated no survival advantage with bilateral transplantations. Moreover, unilateral transplantations have reduced wait time, are easier to perform, and it is possible to improve the health of two patients from one donor.

Nowadays the most used prognostic indexes for IPF are the GAP [18] and TORVAN [19] indexes. Developed in 2012, the composite GAP index relies on the assessment of gender (G), age (A), and physiology (P), the latter including two conventional lung function parameters: the forced vital capacity (FVC) and the diffusion lung capacity for CO (DLCO). It assigns a value to each variable (gender: female = 0 vs. male = 1; age: $\leq$60 years = 0, 61–65 years = 1, >65 years = 2; % predicted FVC: >75% = 0, 50–75% = 1, <50% = 2; % predicted DLCO: >55% = 0, 36–55% = 1, $\leq$35% = 2, cannot perform = 3). The final score resulting from the sum of the addenda allows the estimation of the average risk of mortality at 1, 2, and 3 years by disease stage (i.e., score 0–3 = stage I; score 4–5 = stage II; score 6–8 = stage III). First validated in untreated IPF patients, newer publications are evaluating GAP in treated ones [20].

Nevertheless, over the last ten years, the knowledge and treatment of idiopathic pulmonary fibrosis have evolved, and the search for more effective and personalized tools has become a priority to improve the quality of patient care. The prognosis of IPF patients

indeed depends on many variable comorbidities, and clinical, psychological, and physical conditions [21–25].

The TORVAN index was developed in 2019 by Torrisi et al. [19] with the purpose of integrating age, FVC, and DLCO, and also the evaluation of comorbidities, such as systemic hypertension, atrial arrhythmias, diabetes mellitus, major depressive disorder, gastroesophageal reflux, pulmonary hypertension, lung cancer, and valvular heart disease to estimate the TORVAN stage related to the prognosis.

Torrisi et al. did not include gender in their index because they evaluated that sex becomes a less important parameter when considered in relation to comorbidities. In their study, the most important comorbidities resulted from gastroesophageal reflux, pulmonary hypertension, lung cancer, valvular heart disease, and atrial arrhythmias.

The GAP index, however, still remains the most widely used prognostic tool in daily clinical practice, which is why we decided to focus on it, leaving the possibility of a further machine learning (ML) approach to TORVAN.

Recently, Y. Suzuki et al. [26] have shown that the combination of the GAP index with the body mass index (BMI) successfully identified four sub-groups of IPF patients with different prognoses before the initiation of anti-fibrotic therapies.

Machine learning is a branch of artificial intelligence (AI) that focuses on the use of data and algorithms to imitate the way humans learn, automatically improving from experience without being explicitly programmed. ML displays powerful problem-solving capabilities along with the ability to detect the interaction between numerous variables, thus allowing the development of prediction models. In light of this potential, although burdened by some limitations, the application of machine learning has been explored in clinical settings with encouraging results [27].

Given these considerations, the purpose of our study was to validate the GAP index with an ML approach in a retrospective cohort of IPF patients treated in a real-life setting with the two currently available anti-fibrotic drugs, pirfenidone and nintedanib.

As an additional aim, we tested the possibility of integrating the GAP with other clinical variables to improve its prognostic performance.

2. Machine Learning Operation

Artificial intelligence enables, through machine learning, tasks that are too difficult to solve with fixed programs written and designed in a traditional way. From a scientific and philosophical point of view, ML is interesting because the development of its understanding promotes understanding of the principles underlying intelligence.

Machine learning is achieved through the application of special learning mechanisms (algorithms); in particular, an ML algorithm is an algorithm that is able to learn from data [28] through the exploration of correctly predicted data examples that can instruct the algorithm to build an effective predictive model. In order to obtain the best choice of the dataset's variables (attributes) to employ in the predictive model, a ranking of the attributes is performed, which scores them according to their correlation with the value (class) to be predicted.

Feature scoring and ranking can help in understanding the data in supervised settings. This is typically achieved by means of entropy-based [29–31] or correlation-based [32–34] indicators.

3. Materials and Methods

We retrospectively collected the data of 211 patients affected by idiopathic pulmonary fibrosis attending two referral interstitial lung disease (ILD) centers (Universities of Foggia and Federico II of Naples) at the time of first diagnosis.

Variables collected at baseline (before the initiation of anti-fibrotic therapies) included age (69.5 $\pm$ 7.5 years), gender (males = 186; females = 61), BMI (29.5 $\pm$ 4.4 kg/m^2), smoking habit (active or former smokers = 166), type of IPF diagnosis (based on chest imaging = 175; based on lung histology = 36), and disease phenotype (sporadic = 181; familial = 30). Additional study parameters were those related to the lung function, such as FVC (77 $\pm$ 19% of predicted), forced expiratory volume/1st second (FEV1) (83 $\pm$ 20% of predicted), total

lung capacity (TLC) ($62 \pm 16\%$ of predicted), DLCO ($50 \pm 14\%$ of predicted), and the initial and final O2 peripheral saturation (SpO$_2$) recorded at the 6 min walk test (6MWT) ($95 \pm 2\%$ and $87 \pm 6\%$, respectively), along with the distance walked (367 ± 133 m). The GAP scoring was collected as well (3.8 ± 1.4).

The open-source machine data program Orange (https://orange.biolab.si/ (URL accessed on 06 June 2022) was used to design the machine learning approach to validate the GAP index in estimating the 3 year mortality rate in our study cohort.

ML allows us to analyze large amounts of data with the advantage of detecting possible interactions among many different variables, but it is still a too complex predictive model, which is why its use cannot be routine.

The application of ML to our work started with the selection of the most relevant features to be used in the predictive models, which was carried out by means of principal component analysis (PCA) applied to the patient dataset. Subsequent processing was carried out by "feature ranking" in order to further reduce the number of features in the final model. We used the FCBF (fast correlation–based filter) method—an entropy-based measure, which also identifies redundancy due to pairwise correlations between features—choosing the features with a ranking value greater than zero.

Finally, we trained and tested a series of classification models in order to verify our approach to the GAP index predictivity enforcement by means of BMI. In particular, we built three predictive models using the following ML methods:

- Logistic regression (a classification algorithm with LASSO regularization)
- Naive Bayes (a fast and simple probabilistic classifier based on Bayes' theorem with the assumption of feature independence)
- Neural network (a multi-layer perceptron algorithm with backpropagation)

The learning methods showed slightly different classifier performance measures (such as classification accuracy (CA) and AUC (area under the ROC curve)), which we used for validating their performance by means of cross-validation (a sampling method that allows the reduction of overfitting problems), confusion matrix (which shows proportions between the predicted and actual class) and ROC analysis (which plots a true positive rate against a false positive rate of a test, giving a robust assessment of the predictive model).

These validation tools led to approximately the same results in terms of accuracy and F1 score (0.71), with an AUC (area under ROC curve) of 0.74, confirming the soundness of our hypothesis.

A detailed description of the methodology employed to perform the analysis carried out in this work is reported in a previously published paper by C. Mencar et al. [35].

The study was approved by the Ethics Committee of the Azienda Ospedaliera dei Colli, Naples, Italy. Data of interest for the analysis were anonymously collected into a dedicated database. Informed consent was not required due to the retrospective nature of the study.

4. Results

The overall 3 year mortality rate in our study cohort of IPF patients treated with pirfenidone or nintedanib was 36.7% (78 patients have died). This finding was aligned with the mortality prediction based on the GAP index at baseline (Table 1). No differences were observed in the two treatment arms.

Table 1. Comparison between GAP-predicted and observed mortality in our study cohort of IPF patients.

Disease Stage	3 Year Mortality Estimated by GAP	3 Year Observed Mortality
GAP-I	16%	21%
GAP-II	40%	46%
GAP-III	77%	70%

Table 2 shows the accuracy and other ML-derived parameters calculated by means of different methods when evaluating the sole GAP index. Not dissimilar data were obtained by considering the single GAP components (gender, age, FVC, and DLCO) taken individually, suggesting that machine learning did not produce a further improvement. The best result obtained was for logistic regression (AUC = 0.72; CA = 0.69; F1 = 0.67).

Table 2. Validation of the GAP index by means of machine learning AUC = Area Under Curve; CA = Classification Accuracy; F1 = harmonic mean between Precision and Recall.

	AUC	CA	F1
Logistic Regression	0.70	0.71	0.71
Neural Network	0.69	0.71	0.71
Naive Bayes	0.67	0.71	0.71
Stochastic gradient descent	0.68	0.71	0.71
Random Forest	0.68	0.71	0.71

We considered other variables not included in the index such as 6MWT parameters/distance, initial and final SpO_2, type of diagnosis (radiological or histological), type of therapy, TLC%, and Torvan Index. The best results were obtained by considering the final peripheral saturation of oxygen, the type of diagnosis, and type of therapy (Logistic regression: AUC 0.70–F1 0.66, Neural Network: AUC 0.70–F1 0.65).

A clear improvement in the results was obtained by integrating the analysis of the individual GAP parameters with the value of BMI (Table 3).

Table 3. Validation of the GAP index through the individual analysis of its components (gender, age, FVC, and DLCO) in addition to BMI.

	AUC	CA	F1
Naive Bayes	0.78	0.70	0.70
Logistic Regression	0.77	0.70	0.69
Support Vector Machine	0.74	0.72	0.71
Random Forest	0.73	0.70	0.70
K-Nearest Neighbors	0.71	0.72	0.71

5. Discussion

In the present work, we have explored the possibility of improving the prognostic value of the GAP index in idiopathic pulmonary fibrosis patients treated with anti-fibrotics by means of machine learning algorithms employing optimal predictive variables. In order to do this, we measured the performance of various additional parameters from the available dataset through entropy and correlation indicators, which led to the choice of body mass index as the best additional parameter to add to the predicting model.

Although this result may not seem to be of relevant novelty, it highlights the importance of the "machine learning" approach in the prognostic evaluation of diseases (in our case of IPF) for supporting the diagnostic ability of physicians in the execution of their clinical activities which, if otherwise performed "by hand", would often lead to unsatisfactory results.

The machine learning approach was utilized before in relation to idiopathic pulmonary fibrosis, both for diagnosis and prognosis.

Already in 2011, Flietstra B. et al. [36] used ML to automate the analysis of crackles to distinguish them among IPF, congestive heart failure (CHF), and pneumonia. They analyzed distinctive features of IPF crackles separating them from those in patients with pneumonia or CHF, potentially aiming to avoid eventual medication errors and to ease diagnosis.

Furukawa T. et al. [37] tried to utilize ML to detect IPF in a group of ILD patients and to differentiate them. They used a combination of deep learning, used for semantic segmentation of HRCT images, and an ML algorithm to retrospectively analyze HRCT

images of 1068 patients affected by interstitial lung disease and, without data from non-invasive examinations, they reached the target to identify patients affected by IPF in a population with various ILDs.

Pan J. et al. [38] recently published that machine learning may be used to identify novel radiological disease progression in lung CT of IPF patients. In their work, they proceed to answer three main questions.

With respect to the first question, whether it is possible to identify newer lung pattern types associated with radiological disease progression, they affirm this possibility is feasible.

The second question is if patients with different radiological disease progression have different future outcomes, but in this case, the answer is more difficult. They actually found significantly different outcomes, but by extending the evaluation to an external cohort with different scanner manufacturers, they only observed a comparable trend without reaching significance.

Lastly, with respect to the possibility of visualizing the transition pathways of components that correspond to frequently occurring changes in lungs during radiological progression, they specify the analysis shows that is possible to extract these transition networks from patients through follow-up examinations.

This is an important study because they obtained a substantial comparable performance between ML and expert radiologists, even if in a small cohort of patients, opening the way to further informatics improvement also in the radiological prediction of prognosis.

To the best of our knowledge, this is the first attempt to assess whether the application of machine learning can improve the prediction accuracy of the Gender-Age-Physiology index in IPF patients treated with the anti-fibrotic drugs, pirfenidone and nintedanib.

Our results show that ML did not further improve the GAP performance; this observation was such both when considering the GAP as a composite index and when evaluating individually its single components.

We read the validation of the GAP index based on the comparison between mortality in the Korean population [39] and the population selected in the original article by Ley. In this article from 2015, there was reliability in the survival estimation at 1 year, with worsening of the results in the estimations at 2 and 3 years, with lower mortality than predicted, even though these data were limited by differences in the cohorts (better value of FVC and DLCO, lower age). Data in our analysis show higher mortality in the first and second stage at 3 years and, on the contrary, lower mortality in the third stage, especially because all our patients were treated with antifibrotics (no statistical differences were highlighted between pirfenidone and nintedanib in the analysis) while in the article by Ley, no therapy was approved yet, which is why we expected an improvement in the patients' survival in all stages.

This result seems to corroborate the thesis presented by S. Suissa and K. Suissa in their very recent article. They affirm that in many studies, the reduction in mortality after antifibrotic treatment with pirfenidone or nintedanib is a result of immortal time bias that exaggerates the benefits of drugs, even though further studies are requested [40].

During our analysis, we also tried to find a more accurate index as an alternative one in cases where we did not have the data needed to calculate GAP.

As already cited, the best results were obtained with the final peripheral saturation of oxygen, the type of diagnosis, and the type of therapy, but these parameters and machine learning results did not encourage us to proceed in this research, leaving the possibility of pursuing this approach in future studies.

A mild relation is visible in the approach to improving the GAP index attempted by Chandel A. et al. [41], where they integrated into GAP parameters 6MWD and exertional hypoxia.

They propose a Distance-Oxygen-GAP (DO-GAP) in which they suggest that 6MWD is a more reproducible measure of the disease that improves GAP because it assesses overall functional ability and evaluates also the impact of other comorbidities (such as cardiovascular, rheumatologic, or neurologic impact). This is a very interesting result;

unfortunately in our ML approach, with our dataset of patients, we obtained only a mild and not significant improvement.

In addition, we found it really interesting to further classify, as proposed in the article by Suzuki et al., patients considering BMI. In the paper, it is proposed to move to the next stage patients that have a BMI lower than the cut-off (24 kg/m^2), an approach that caused a reclassification, especially of patients with GAP stage II (approximately 85% of their cohort). This theory seems confirmed by our validation through machine learning. Furthermore, as cited in the article, a faster decline in FVC among patients with lower BMI and weight loss > 5% during 52 weeks was also reported in the second analysis of the INPULSIS study.

There are some limitations in this study. We used ML analysis only to evaluate the GAP index. While the TORVAN index certainly allows the inclusion, in the estimation of prognosis, also the presence of comorbidities, as already cited, we preferred to focus on the GAP index because is still the most widely used, leaving the possibility of evaluating TORVAN in a further analysis.

The GAP index was published and tested in an untreated population, so its absence could be another limitation; however, all patients in our dataset are treated with antifibrotics according to guidelines.

Lastly, the dataset is composed of a population geographically limited to Southern Italy; nevertheless, the analyzed patients are from two separate centers and the number of the patient sample is not negligible for a rare disease.

6. Conclusions

In conclusion, the analysis of data, based only on our dataset, demonstrates that even if the GAP index utilizes common variables for patients affected by idiopathic pulmonary fibrosis, making it very simple to measure, is still reliable in terms of prognosis, confirmed by the fact that more complicated calculation systems such as machine learning, utilizing only clinical parameters, does not provide better results. On the other hand, the addition of the parameter BMI, as suggested in the paper by Y. Suzuki et al., seems to further improve the predictive accuracy of GAP.

Author Contributions: Conceptualization, D.L., C.C.D.P. and C.G.; methodology, D.L. and C.G.; software, P.T. and C.G.; validation, C.G. and L.C.; formal analysis, D.L., C.C.D.P. and C.G.; investigation, D.L., C.C.D.P., G.S., G.R. and M.B.; resources, P.T., L.C. and M.B.; data curation, D.L. and C.C.D.P.; writing—original draft preparation, D.L., C.C.D.P. and C.G.; writing—review and editing, L.C., P.T. and M.B.; visualization, P.T., G.R. and C.G.; supervision, D.L. and M.B.; project administration, D.L. and M.B. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was approved by the Ethics Committee of the Azienda Ospedaliera dei Colli, Naples, Italy. Data of interest for the analysis were anonymously collected in a dedicated database.

Informed Consent Statement: Informed consent was not required due to the retrospective nature of the study.

Data Availability Statement: Data available by contacting the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Richeldi, L.; Collard, H.R.; Jones, M.G. Idiopathic pulmonary fibrosis. *Lancet* **2017**, *389*, 1941–1952. [CrossRef] [PubMed]
2. King, T.E., Jr.; Pardo, A.; Selman, M. Idiopathic pulmonary fibrosis. *Lancet* **2011**, *378*, 1949–1961. [CrossRef] [PubMed]
3. Raghu, G.; Remy-Jardin, M.; Richeldi, L.; Thomson, C.C.; Inoue, Y.; Johkoh, T.; Kreuter, M.; Lynch, D.A.; Maher, T.M.; Martinez, F.J.; et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am. J. Respir. Crit. Care Med.* **2022**, *205*, e18–e47. [CrossRef] [PubMed]
4. Noble, P.W.; Albera, C.; Bradford, W.Z.; Costabel, U.; Glassberg, M.K.; Kardatzke, D.; King, T.E., Jr.; Lancaster, L.; Sahn, S.A.; Szwarcberg, J.; et al. CAPACITY Study Group. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): Two randomised trials. *Lancet* **2011**, *377*, 1760–1769. [CrossRef]

5.   Lancaster, L.H.; de Andrade, J.A.; Zibrak, J.D.; Padilla, M.L.; Albera, C.; Nathan, S.D.; Wijsenbeek, M.S.; Stauffer, J.L.; Kirchgaessler, K.-U.; Costabel, U. Pirfenidone safety and adverse event management in idiopathic pulmonary fibrosis. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* **2017**, *26*, 170057. [CrossRef] [PubMed]

6.   Richeldi, L.; du Bois, R.M.; Raghu, G.; Azuma, A.; Brown, K.K.; Costabel, U.; Cottin, V.; Flaherty, K.R.; Hansell, D.M.; Inoue, Y.; et al. INPULSIS Trial Investigators Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **2014**, *370*, 2071–2082. [CrossRef] [PubMed]

7.   Roth, G.J.; Binder, R.; Colbatzky, F.; Dallinger, C.; Schlenker-Herceg, R.; Hilberg, F.; Wollin, S.-L.; Kaiser, R. Nintedanib: From discovery to the clinic. *J. Med. Chem.* **2015**, *58*, 1053–1063. [CrossRef]

8.   Soccio, P.; Moriondo, G.; Lacedonia, D.; Tondo, P.; Quarato, C.M.I.; Foschino Barbaro, M.P.; Scioscia, G. EVs-miRNA: The New Molecular Markers for Chronic Respiratory Diseases. *Life* **2022**, *12*, 1544. [CrossRef]

9.   Lee, Y.J.; Choi, S.M.; Lee, Y.J.; Cho, Y.J.; Yoon, H.I.; Lee, J.H.; Lee, C.T.; Park, J.S. Clinical impact of depression and anxiety in patients with idiopathic pulmonary fibrosis. *PLoS ONE* **2017**, *12*, e0184300. [CrossRef]

10. McDonnell, M.J.; Hunt, E.B.; Ward, C.; Pearson, J.P.; O'Toole, D.; Laffey, J.G.; Murphy, D.M.; Rutherford, R.M. Current therapies for gastro-oesophageal reflux in the setting of chronic lung disease: State of the art review. *ERJ Open Res.* **2020**, *6*, 00190–2019. [CrossRef]

11. Laudisio, A.; Antonelli Incalzi, R.; Gemma, A.; Giovannini, S.; Lo Monaco, M.R.; Vetrano, D.L.; Padua, L.; Bernabei, R.; Zuccalà, G. Use of proton-pump inhibitors is associated with depression: A population-based study. *Int. Psychogeriatr.* **2018**, *30*, 153–159. [CrossRef]

12. Shen, L.; Zhang, Y.; Su, Y.; Weng, D.; Zhang, F.; Wu, Q.; Chen, T.; Li, Q.; Zhou, Y.; Hu, Y.; et al. New pulmonary rehabilitation exercise for pulmonary fibrosis to improve the pulmonary function and quality of life of patients with idiopathic pulmonary fibrosis: A randomized control trial. *Ann. Palliat. Med.* **2021**, *10*, 7289–7297. [CrossRef] [PubMed]

13. Scioscia, G.; De Pace, C.C.; Giganti, G.; Tondo, P.; Foschino Barbaro, M.P.; Lacedonia, D. Real life experience of molnupiravir as a treatment of SARS-CoV-2 infection in vaccinated and unvaccinated patients: A letter on its effectiveness at preventing hospitalization. *Ir. J. Med. Sci.* **2022**, 1–3. [CrossRef] [PubMed]

14. Steinmetz, A.; Bahlmann, S.; Bergelt, C.; Bröker, B.M.; Ewert, R.; Felix, S.B.; Flöel, A.; Fleischmann, R.; Hoffmann, W.; Holtfreter, S.; et al. The Greifswald Post COVID Rehabilitation Study and Research (PoCoRe)-Study Design, Characteristics and Evaluation Tools. *J. Clin. Med.* **2023**, *12*, 624. [CrossRef] [PubMed]

15. Lacedonia, D.; Scioscia, G.; De Pace, C.C.; Laricchiuta, A.; Tondo, P.; Sabato, R.; Foschino Barbaro, M.P. How Are We Handling the Post-COVID Patients? The Dance of Uncertainties. *Respir. Int. Rev. Thorac. Dis.* **2022**, *101*, 210–213. [CrossRef]

16. Glass, D.S.; Grossfeld, D.; Renna, H.A.; Agarwala, P.; Spiegler, P.; DeLeon, J.; Reiss, A.B. Idiopathic pulmonary fibrosis: Current and future treatment. *Clin. Respir. J.* **2022**, *16*, 84–96. [CrossRef]

17. Li, D.; Liu, Y.; Wang, B. Single versus bilateral lung transplantation in idiopathic pulmonary fibrosis: A systematic review and meta-analysis. *PLoS ONE* **2020**, *15*, e0233732. [CrossRef]

18. Ley, B.; Ryerson, C.J.; Vittinghoff, E.; Ryu, J.; Tomassetti, S.; Lee, J.S.; Poletti, V.; Buccioli, M.; Elicker, B.M.; Jones, K.D.; et al. A multidimensional index and staging system for idiopathic pulmonary fibrosis. *Ann. Intern. Med.* **2012**, *156*, 684–691. [CrossRef]

19. Torrisi, S.E.; Ley, B.; Kreuter, M.; Wijsenbeek, M.; Vittinghoff, E.; Collard, H.R.; Vancheri, C. The added value of comorbidities in predicting survival in idiopathic pulmonary fibrosis: A multicentre observational study. *Eur. Respir. J.* **2019**, *53*, 1801587. [CrossRef]

20. Zhang, X.; Ren, Y.; Xie, B.; Wang, S.; Geng, J.; He, X.; Jiang, D.; He, J.; Luo, S.; Wang, X.; et al. External validation of the GAP model in Chinese patients with idiopathic pulmonary fibrosis. *Clin. Respir. J.* 2022; *early view*. [CrossRef]

21. Jouneau, S.; Rousseau, C.; Lederlin, M.; Lescoat, A.; Kerjouan, M.; Chauvin, P.; Luque-Paz, D.; Guillot, S.; Oger, E.; Vernhet, L.; et al. Malnutrition and decreased food intake at diagnosis are associated with hospitalization and mortality of idiopathic pulmonary fibrosis patients. *Clin. Nutr.* **2022**, *41*, 1335–1342. [CrossRef]

22. Mann, J.; Goh, N.S.L.; Holland, A.E.; Khor, Y.H. Cough in Idiopathic Pulmonary Fibrosis. *Front. Rehabil. Sci.* **2021**, *2*, 751798. [CrossRef] [PubMed]

23. Horio, Y.; Takihara, T.; Takahashi, F.; Enokida, K.; Nakamura, N.; Tanaka, J.; Tomomatsu, K.; Niimi, K.; Tajiri, S.; Hayama, N.; et al. Prognosis of acute exacerbation in idiopathic pulmonary fibrosis with pulmonary emphysema: A retrospective cohort study in Japan. *BMJ Open* **2022**, *12*, e062236. [CrossRef] [PubMed]

24. Badenes-Bonet, D.; Rodó-Pin, A.; Castillo-Villegas, D.; Vicens-Zygmunt, V.; Bermudo, G.; Hernández-González, F.; Portillo, K.; Martínez-Llorens, J.; Chalela, R.; Caguana, O.; et al. Predictors and changes of physical activity in idiopathic pulmonary fibrosis. *BMC Pulm. Med.* **2022**, *22*, 340. [CrossRef] [PubMed]

25. Caminati, A.; Lonati, C.; Cassandro, R.; Elia, D.; Pelosi, G.; Torre, O.; Zompatori, M.; Uslenghi, E.; Harari, S. Comorbidities in idiopathic pulmonary fibrosis: An underestimated issue. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* **2019**, *28*, 190044. [CrossRef]

26. Suzuki, Y.; Mori, K.; Aono, Y.; Kono, M.; Hasegawa, H.; Yokomura, K.; Naoi, H.; Hozumi, H.; Karayama, M.; Furuhashi, K.; et al. Combined assessment of the GAP index and body mass index at antifibrotic therapy initiation for prognosis of idiopathic pulmonary fibrosis. *Sci. Rep.* **2021**, *11*, 18579. [CrossRef]

27. Scioscia, G.; Tondo, P.; Foschino Barbaro, M.P.; Sabato, R.; Gallo, C.; Maci, F.; Lacedonia, D. Machine learning-based prediction of adherence to continuous positive airway pressure (CPAP) in obstructive sleep apnea (OSA). *Inform. Health Soc. Care* **2022**, *47*, 274–282. [CrossRef]

28. Mitchell, T.M. *Machine Learning*; McGraw-Hill: New York, NY, USA, 1997.
29. Duch, W.; Wieczorek, T.; Biesiada, J.; Blachnik, M. Comparison of feature ranking methods based on information entropy. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), Budapest, Hungary, 25–29 July 2004; Volume 2, pp. 1415–1419.
30. Gradojevic, N.; Caric, M. Predicting systemic risk with entropic indicators. *J. Forecast.* **2017**, *36*, 16–25. [CrossRef]
31. Aremu, O.O.; Cody, R.A.; Hyland-Wood, D.; McAree, P.R. A relative entropy based feature selection framework for asset data in predictive maintenance. *Comput. Ind. Eng.* **2020**, *145*, 106536. [CrossRef]
32. Pramokchon, P.; Piamsanga, P. An unsupervised, fast correlation-based filter for feature selection for data clustering. In Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013); Springer: Singapore, 2014; pp. 87–94.
33. Gopika, N.; Me, A.M.K. Correlation based feature selection algorithm for machine learning. In Proceedings of the 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 15–16 October 2018; pp. 692–695.
34. Du, S.; Wang, S. An overview of correlation-filter-based object tracking. In *IEEE Transactions on Computational Social Systems*; Browse Journals & Magazines: Piscataway, NJ, USA, 2021.
35. Mencar, C.; Gallo, C.; Mantero, M.; Tarsia, P.; Carpagnano, G.E.; Foschino Barbaro, M.P.; Lacedonia, D. Application of machine learning to predict obstructive sleep apnea syndrome severity. *Health Inform. J.* **2020**, *26*, 298–317. [CrossRef]
36. Flietstra, B.; Markuzon, N.; Vyshedskiy, A.; Murphy, R. Automated analysis of crackles in patients with interstitial pulmonary fibrosis. *Pulm. Med.* **2011**, *2011*, 590506. [CrossRef] [PubMed]
37. Furukawa, T.; Oyama, S.; Yokota, H.; Kondoh, Y.; Kataoka, K.; Johkoh, T.; Fukuoka, J.; Hashimoto, N.; Sakamoto, K.; Shiratori, Y.; et al. A comprehensible machine learning tool to differentially diagnose idiopathic pulmonary fibrosis from other chronic interstitial lung diseases. *Respirology* **2022**, *27*, 739–746. [CrossRef]
38. Pan, J.; Hofmanninger, J.; Nenning, K.H.; Prayer, F.; Röhrich, S.; Sverzellati, N.; Poletti, V.; Tomassetti, S.; Weber, M.; Prosch, H.; et al. Unsupervised machine learning identifies predictive progression markers of IPF. *Eur. Radiol.* **2022**, *33*, 925–935. [CrossRef]
39. Kim, E.S.; Choi, S.M.; Lee, J.; Park, Y.S.; Lee, C.-H.; Yim, J.-J.; Yoo, C.-G.; Kim, Y.W.; Han, S.K.; Lee, S.-M. Validation of the GAP score in Korean patients with idiopathic pulmonary fibrosis. *Chest* **2015**, *147*, 430–437. [CrossRef] [PubMed]
40. Suissa, S.; Suissa, K. Antifibrotics and Reduced Mortality in Idiopathic Pulmonary Fibrosis: Immortal Time Bias. *Am. J. Respir. Crit. Care Med.* **2023**, *207*, 105–109. [CrossRef] [PubMed]
41. Chandel, A.; Pastre, J.; Valery, S.; King, C.S.; Nathan, S.D. Derivation and validation of a simple multidimensional index incorporating exercise capacity parameters for survival prediction in idiopathic pulmonary fibrosis. *Thorax* **2022**. [CrossRef] [PubMed]