



# Patients with idiopathic pulmonary fibrosis have poor clinical outcomes with COVID-19 disease: a propensity matched multicentre research network analysis

Syeda Fatima Naqvi ,<sup>1</sup> Dhairya A Lakhani ,<sup>2</sup> Amir Humza Sohail,<sup>3</sup> James Maurer,<sup>3</sup> Sarah Sofka,<sup>4</sup> Arif Sarwari,<sup>5</sup> Yousaf B Hadi<sup>6</sup>

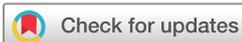
**To cite:** Naqvi SF, Lakhani DA, Sohail AH, *et al.* Patients with idiopathic pulmonary fibrosis have poor clinical outcomes with COVID-19 disease: a propensity matched multicentre research network analysis. *BMJ Open Res Res* 2021;**8**:e000969. doi:10.1136/bmjresp-2021-000969

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjresp-2021-000969>).

SFN and DAL contributed equally.

SFN and DAL are joint first authors.

Received 23 April 2021  
Accepted 15 June 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Yousaf B Hadi;  
yousaf.hadi@gmail.com

## ABSTRACT

**Introduction** Outcomes of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in patients with pre-existing idiopathic pulmonary fibrosis (IPF) remain understudied, and it is unknown if IPF is an independent predictor of worse disease course. Herein, we report the clinical outcomes in a large cohort of 251 patients with COVID-19 in the setting of known IPF. Outcomes were compared with a propensity matched cohort of patients with COVID-19 without IPF.

**Methods** Analysis of a federated multicentre research network TriNetX was performed including patients more than 16 years of age diagnosed with SARS-CoV-2 infection. Outcomes in patients diagnosed as positive for SARS-CoV-2 infection with concurrent IPF were compared with a propensity matched cohort of patients without IPF.

**Results** A total of 311 060 patients with SARS-CoV-2 infection on the research network were identified, 251 patients (0.08%) carried a diagnosis of IPF. Mean age of patients with IPF was 68.30±12.20 years, with male predominance (n=143, 56.97%). Comorbidities including chronic lower respiratory diseases, diabetes mellitus, ischaemic heart disease and chronic kidney disease were more common in patients with IPF when compared with the non-IPF cohort. After propensity matching, higher rates of composite primary outcome (death or mechanical ventilation) at 30 and 60 days, as well as need for hospitalisation, critical care, and acute kidney injury were observed in the IPF cohort.

**Conclusion** Poor outcomes of COVID-19 disease were observed in patients with IPF after robust matching of confounders. Our data confirm that patients with IPF constitute a high-risk cohort for poor outcomes related to COVID-19 disease.

## INTRODUCTION

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) mainly affects the lung parenchyma. Interstitial lung disease (ILD) is a group of fibroinflammatory diseases which affects the alveolar interstitium. The

## Key messages

- Are patients with idiopathic pulmonary fibrosis (IPF) at increased risk of worse outcomes with COVID-19?
- COVID-19 has worse outcomes in patients with IPF compared with those without IPF.
- Outcomes of COVID-19 in patients with IPF are not yet reported in current literature. In our large multicentre research network analysis, we have compared outcomes of COVID-19 in patients with IPF to a propensity matched control cohort.

prevalence of ILD is under 1 per 1000 people and annual incidence of 20 per 100 000 people in Europe.<sup>1</sup> Idiopathic pulmonary fibrosis (IPF) is a fibrotic ILD with an incidence of 12 per 100 000 in UK.<sup>2</sup> IPF tends to affect older population and is associated with other comorbid conditions which includes diabetes, hypertension and ischaemic heart disease.<sup>3</sup> These comorbid conditions have shown association with severe forms of COVID-19 disease.<sup>4,5</sup>

Literature have shown that patients with concomitant pulmonary disease have worse outcomes with COVID-19 disease. Therefore, infected patients with any underlying chronic lung disease, especially with ILD such as IPF and sarcoidosis, would be theoretically at higher risk of severe complications. Data on SARS-CoV-2 infection outcomes in patients with pre-existing IPF is currently limited. Further, there is no controlled study of COVID-19 disease in IPF available in the current literature. Hence it is unknown if IPF independently portends a worse prognosis in COVID-19 disease.

Herein, we report the clinical outcomes among a large cohort of 251 patients with COVID-19 in the setting of known IPF.



Outcomes were compared with a propensity matched cohort of patients with COVID-19 without IPF.

## STUDY DESIGN AND METHODS

### Description of data source

Retrospective cohort study was performed using multi-institutional research network TriNetX (Cambridge, Massachusetts, USA) platform. TriNetX is a federated research network which includes more than 40 healthcare organisations in USA. This platform provides real-time access to the healthcare records and includes patients from the healthcare organisations (more than 40 million), in a de-identified fashion. The data are directly retrieved from the electronic health record management systems (EHRs) of the participating organisations. These organisations are large academic centres that operate both tertiary care and satellite outpatient office locations.

Clinical variables (facts) are derived directly from EHRs and through a built-in Natural Language Processing system which extracts variables from clinical documents. Robust quality assurance is achieved at the time of extraction from EHRs before inclusion in this database, in a systemic and standardised format. This interface only provides aggregate counts and statistical summaries to protect patient health information and ensures that the data remains de-identified at all levels of data retrieval and dissemination.

West Virginia University Clinical and Translational Science Institute manages the TriNetX platform at our institution and provides access to the end-users. This platform obfuscates patient counts less than 10 to ensure anonymity.

### Study participants

A real-time search and analysis was conducted on the TriNetX platform and updated through 18 January 2021. All patients with SARS-CoV-2 infection on TriNetX platform who were more than 16 years of age at diagnosis of COVID-19 were included. Diagnosis of SARS-CoV-2 infection or documentation of positive SARS-CoV-2 test between 20 January 2020 and 15 November 2020 were required for inclusion in the study cohort. The first reported case of SARS-CoV-2 infection in the USA was on 20 January 2020 and hence was chosen as the start date for enrolment. Primary study endpoint included outcomes at 60 days from diagnosis, and hence 15 November 2020 was chosen as the end date for enrolment.

The selection and identification criteria were based on diagnostic codes and positive laboratory confirmation using standard COVID-19 terminology recommended by the Centers for Disease Control and Prevention and endorsed by TriNetX platform for use in studies conducted on this research network and is detailed in the online supplemental appendix 1. Patients with SARS-CoV-2 infection identified by these criteria were

then divided in cohorts based on previous diagnosis of IPF.

Patients with IPF were identified using the International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification code J84.1.

### Patient and public involvement

Our study was conducted on a multicentre research network which has de-identified patient information. There was no direct interaction with any patients and therefore they were not involved in study design, recruitment or conduct of the study.

### Study outcomes

Clinical and laboratory outcomes were compared between patients with and without IPF. Primary study outcome was defined as a composite of death or requirement for mechanical ventilation in 30-day and 60-day period from the index event. Either the time of SARS-CoV-2 infection or first SARS-CoV-2 positive test result date, whichever occurred first, was defined as the index event for the purposes of the study. Other study outcomes included death, hospitalisation and need for mechanical ventilation in the 30-day and 60-day period from SARS-CoV-2 diagnosis. All laboratory outcomes were considered up to 7 days from the index event.

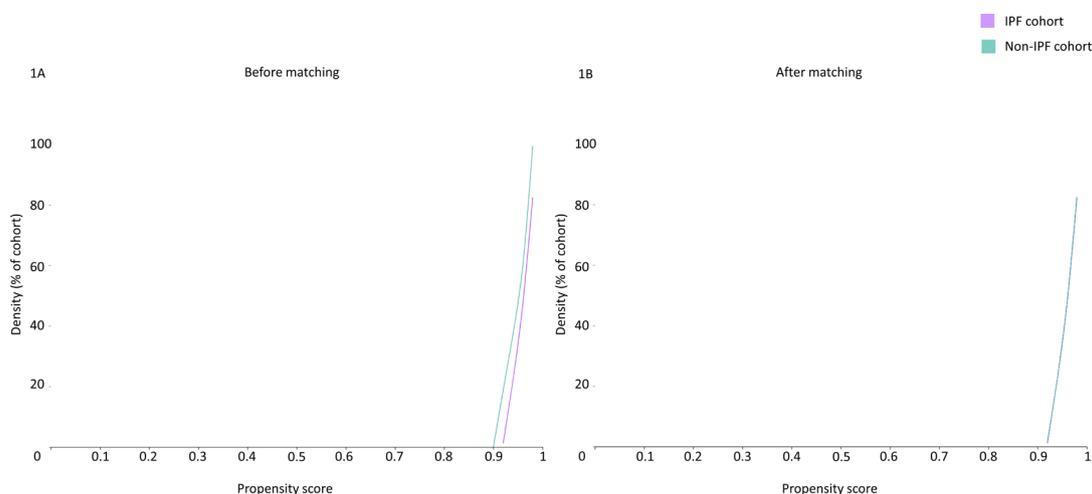
### Statistical analysis

All statistical analyses were conducted using TriNetX platform. Baseline characteristics of study cohorts were described using means, SD and proportions as appropriate. Univariate analysis was conducted using  $\chi^2$  tests for categorical data and independent sample t-tests for continuous data. Subsequently, propensity score matching of the two study cohorts was conducted within the platform. One-to-one (1:1) propensity score matching was performed with age, race, chronic kidney disease, diabetes, hypertension, chronic lung diseases, nicotine dependence, heart failure, ischaemic heart disease, body mass index (BMI), and gender as covariates and thus a matched control group of patients without IPF was identified. TriNetX platform uses input matrices of the user identified covariates to conduct logistic regression analysis to obtain propensity scores for all individual subjects. One-to-one (1:1) matching was performed based on the propensity scores thus generated by using greedy nearest neighbour algorithms using a calliper width of 0.1 pooled SD. TriNetX randomises the order of rows in order to eliminate bias resulting from nearest neighbour algorithms.

A two-sided alpha of less than 0.05 was defined a priori for statistical significance. Balance on covariates in the matched cohorts was assessed using standardised mean difference, and absolute values >0.1 were considered positive for residual imbalance. Risk ratios with 95% CIs were calculated for all analyses between cohorts.

**Table 1** Comparison of clinical characteristics in patients with COVID-19 with and without idiopathic pulmonary fibrosis (IPF) before and after propensity score matching

Variable	Before propensity matching				After propensity matching				P value	Standard difference	P value	
	IPF cohort (n=251)		Non-IPF cohort (n=310809)		IPF cohort (n=251)		Non-IPF cohort (n=251)					
	Number	Per cent/SD	Number	Per cent/SD	Number	Per cent/SD	Number	Per cent/SD				
<b>Demographics</b>												
Age (mean, SD)	68.30	12.20	45.80	18.90	1.41	<0.01	68.30	12.20	690	12.90	0.06	0.52
Male	143	56.97	138248	44.48	0.25	<0.01	143	56.97	142	56.57	0.01	0.93
Female	108	43.03	171351	55.13	0.24	<0.01	108	43.03	109	43.43	0.01	0.93
Body mass index (30 and above)	99	39.44	50100	16.12	0.54	<0.01	99	39.44	92	36.65	0.06	0.52
African American	37	14.74	57982	18.66	0.11	0.11	37	14.74	37	14.74	<0.01	1
Caucasian	175	69.72	179702	57.82	0.25	<0.01	175	69.72	185	73.71	0.09	0.32
Hispanic or Latino	38	15.14	48701	15.67	0.01	0.82	38	15.14	20	7.97	0.01	0.23
<b>Comorbidities</b>												
Hypertension	200	79.68	81038	26.07	1.27	<0.01	200	79.68	201	80.08	0.01	0.91
Chronic kidney disease	99	39.44	17761	5.71	0.88	<0.01	99	39.44	92	36.65	0.06	0.52
Chronic lower respiratory diseases	185	73.71	45324	14.58	1.48	<0.01	185	73.71	185	73.71	<0.01	1
Diabetes mellitus	127	50.60	41489	13.35	0.87	<0.01	127	50.60	126	50.20	0.01	0.93
Ischaemic heart disease	143	56.97	23683	7.62	1.24	<0.01	143	56.97	152	60.56	0.07	0.41
Nicotine dependence	26	10.36	20169	6.49	0.14	0.01	26	10.36	28	11.16	0.03	0.77



**Figure 1** Propensity score density graph in the unmatched and matched IPF (purple) and non-IPF (green) cohorts. IPF, idiopathic pulmonary fibrosis.

## RESULTS

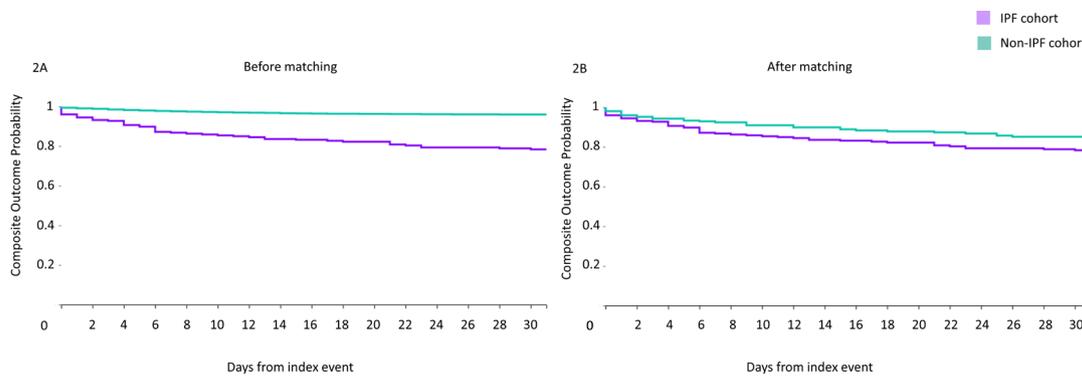
### Study population

Using our inclusion criteria, we identified a total of 311060 patients with SARS-CoV-2 infection on the research network during the study period. Of these identified patients, 251 patients (0.08%) carried a diagnosis of IPF, while the rest 310809 patients (99.92%) were included in the non-IPF cohort.

Mean age at the time of COVID-19 diagnosis in patients with IPF was  $68.30 \pm 12.20$  years, with male predominance. Majority of the participants were Caucasian ( $n=175$ , 69.72%). Comorbidities including chronic lower respiratory diseases, diabetes mellitus, nicotine dependence, ischaemic heart disease, hypertension and chronic kidney disease were more common in patients with IPF when compared with the non-IPF cohort (all p

**Table 2** Clinical outcomes in the two cohorts of patients with COVID-19 with and without idiopathic pulmonary fibrosis (IPF) before and after propensity score matching

Outcome	IPF group (n=251)	Percentage	Non-IPF group (n=310809)	Percentage	Risk ratio	95% CI lower	95% CI upper
<b>Before propensity score matching</b>							
Mortality within 30 days	40	15.94	5342	1.72	9.27	6.97	12.33
Mortality within 60 days	46	18.33	6087	1.96	9.36	7.20	12.17
Inpatient services	111	44.22	26929	8.66	5.10	4.44	5.87
Critical care	55	21.91	9120	2.93	7.47	5.91	9.44
Mechanical ventilation	29	11.55	5606	1.80	6.41	4.54	9.03
30-day composite outcome	50	19.92	8740	2.81	7.08	5.52	9.09
60-day composite outcome	54	21.51	9240	2.97	7.24	5.71	9.17
Acute renal failure	65	25.90	11 803	3.80	6.82	5.53	8.41
Need for renal replacement therapy	11	4.38	1311	0.42	10.39	5.82	18.56
Outcome	IPF group (n=251)	Percentage	Non-IPF group (n=251)	Percentage	Risk ratio	95% CI lower	95% CI upper
<b>After propensity score matching</b>							
Mortality within 30 days	40	15.94	28	11.16	1.43	0.91	2.24
Mortality within 60 days	46	18.33	30	11.95	1.53	1	2.35
Inpatient services	111	44.22	62	24.70	1.79	1.39	2.32
Critical care	55	21.91	20	7.97	2.75	1.70	4.45
Mechanical ventilation	29	11.55	13	5.18	2.23	1.19	4.19
30-day composite outcome	50	19.92	32	12.75	1.56	1.04	2.35
60-day composite outcome	54	21.51	34	13.55	1.59	1.07	2.35
Acute renal failure	65	25.90	42	16.73	1.55	1.10	2.19



**Figure 2** Kaplan-Meier plots of composite endpoint (mortality and mechanical ventilation combined) in SARS-CoV-2 infected patients with IPF (purple) and without IPF (green), before (figure 2A) and after propensity matching (figure 2B). IPF, idiopathic pulmonary fibrosis.

values < 0.01). Baseline demographic characteristics and comorbid conditions are described in [table 1](#).

Of the patients with IPF, 59 patients (23.51%) were prescribed antifibrotic agents in the 6 months before index event; 34 patients (14%) were using nintedanib while pirfenidone was used by 25 patients (10%).

### CLINICAL OUTCOMES

Cohort one-to-one (1:1) matching was performed for age, BMI, chronic kidney disease, diabetes mellitus, hypertension, chronic lung diseases, ethnicity, nicotine dependence, heart failure, ischaemic heart disease and gender to identify a propensity matched control group of patients without IPF. The cohorts were well matched after propensity score matching (mean standard difference < 0.1 for all covariates) as described in [table 1](#) and illustrated in [figure 1](#).

In the 30-day and 60-day period post SARS-CoV-2 infection, 40 (15.94%) and 60 (18.33%) deaths were reported in the IPF group, respectively. The composite outcome (death or mechanical ventilation) was reached by 50 (19.92%) and 54 (21.51%) patients in the IPF group at 30-day and 60-day follow-up, respectively ([table 2](#)).

In the crude unmatched analysis, significantly higher rates of mortality, mechanical ventilation, need for hospitalisation, critical care, renal failure and requirement for renal replacement therapy were observed in the IPF cohort when compared with patients without IPF as illustrated in [table 2](#). Patients with IPF were seven times more likely to reach composite endpoint of death or mechanical ventilation at 30 days (risk ratio: 7.08, 95% CI, 5.52 to 9.09).

After propensity matching, higher rates of composite endpoint 30 and 60 days, as well as need for hospitalisation, critical care and acute renal failure were observed in the IPF cohort, as shown in [table 2](#).

Kaplan-Meier survival curves were plotted for the primary study outcome (composite of death or mechanical ventilation) ([figure 2](#)). Lower survival probability was observed in the IPF cohort at follow-up compared with matched patients with non-IPF (logrank survival p value 0.045).

Subgroup analysis of propensity matched hospitalised patients only, revealed no difference in clinical outcomes in the two groups ([table 3](#)).

**Table 3** Clinical outcomes in the two cohorts of hospitalised patients with COVID-19 with and without idiopathic pulmonary fibrosis (IPF) before and after propensity score matching

Outcome	IPF group (n=111)	Percentage	Non-IPF group (n=111)	Percentage	Risk ratio	95% CI lower	95% CI upper
Mortality within 30 days	31	27.93	24	21.62	1.29	0.81	2.05
Mortality within 60 days	34	30.63	27	24.32	1.26	0.82	1.94
Critical care	47	42.34	38	34.23	1.24	0.88	1.73
Mechanical ventilation	19	17.12	13	11.71	1.46	0.76	2.81
30-day composite outcome	37	33.33	31	27.93	1.19	0.80	1.78
60-day composite outcome	39	35.14	33	29.73	1.18	0.81	1.73
Acute renal injury	33	29.73	34	30.63	0.97	0.65	1.45

**Table 4** Laboratory values in the two cohorts of patients with COVID-19 with and without idiopathic pulmonary fibrosis (IPF) before and after propensity score matching

Outcome	Before propensity matching				After propensity matching					
	IPF cohort (n=251)		Non-IPF cohort (n=310809)		IPF cohort (n=251)		Non-IPF cohort (n=251)			
	Patients with outcome	Mean (SD)	Patients with outcome	Mean (SD)	P value	Patients with outcome	Mean (SD)	Patients with outcome	Mean (SD)	P value
C reactive protein (mg/L)	80	70.11 (77.96)	33 169	63.50 (77.20)	0.44	80	70.11 (77.96)	70	66.40 (77.49)	0.77
Lactate dehydrogenase (U/L)	72	400.88 (212.88)	26 623	411.63 (508.58)	0.86	72	400.88 (212.88)	58	389.78 (348.61)	0.82
Erythrocyte sedimentation rate (mm/hour)	24	44.67 (31.32)	9903	51.92 (32.42)	0.27	24	44.67 (31.32)	17	42.53 (31.84)	0.83
Alanine aminotransferase (U/L)	109	47.84 (130.06)	54 603	49.69 (155.41)	0.90	109	47.84 (130.06)	98	33.15 (39.90)	0.28
Aspartate aminotransferase (U/L)	108	47.82 (103.17)	53 037	57.06 (358.16)	0.79	108	47.82 (103.17)	96	36.66 (23.07)	0.30
Bilirubin (mg/dL)	107	0.65 (0.53)	54 338	0.63 (1.06)	0.80	107	0.65 (0.53)	99	0.59 (0.34)	0.32
Ferritin (ng/mL)	70	1019.95 (1643.80)	30 142	974.95 (2785.47)	0.89	70	1019.95 (1643.80)	62	688.53 (915)	0.16
Creatinine (ng/dL)	123	1.47 (1.32)	64 487	1.18 (1.38)	0.02	123	1.47 (1.32)	115	1.59 (1.77)	0.56

## LABORATORY VALUES

Table 4 describes mean values of C reactive protein (mg/L), lactate dehydrogenase (U/L), erythrocyte sedimentation rate (mm/hour), alanine aminotransferase (U/L), aspartate aminotransferase (U/L), serum bilirubin (mg/dL) and serum ferritin (ng/mL). No difference in studied laboratory parameters were found in the two cohorts (table 4).

## DISCUSSION

Management of ILD, especially IPF, is challenging during the current pandemic. Patients with IPF have underlying parenchymal fibrotic disease, there is a propensity of disease exacerbations in these patients in the setting of superimposed pulmonary disorders. Further, the use of immunomodulating therapy may be theorised to influence the disease course of COVID-19 in these patients. Viral illnesses may also be associated with IPF exacerbations.<sup>6</sup> Recently, some case series and retrospective data on outcomes of COVID-19 in patients with ILD have been reported in scientific literature.<sup>7-9</sup>

Three major reports of COVID-19 outcomes in patients with ILD are available in literature. In a multi-centre case-control study, Esposito *et al* described a fourfold increase in adjusted odds of death in patients with COVID-19 with comorbid ILD compared with those without ILD.<sup>7</sup> However, their cohort only included six patients with IPF, and 74% of patients with ILD were derived from a hospitalised cohort. In another recent report of 123 patients with ILD with COVID-19, Gallay *et al* included 20 patients with IPF.<sup>8</sup> They found a mortality rate of 35% for patients with fibrotic ILD, however outcomes were not described for IPF separately. Most of their cohort (>80%) was derived from hospitalised patients as well. As disease pathogenesis and pulmonary parenchymal phenotype is markedly different in fibrotic and non-fibrotic ILD, authors postulated that outcomes and disease manifestations may not be the same in different ILDs, and thus ILD population may be a heterogeneous entity as pertains to COVID-19 disease.

Drake *et al*, in their multinational registry, reported outcomes of 161 patients with ILD hospitalised with COVID-19 and compared them to a non-ILD cohort; their cohort included 68 patients with IPF. Survival was significantly poorer in patients with IPF when compared with those without any ILD (HR, 1.74; 95% CI, 1.16 to 2.60;  $p=0.007$ ).<sup>9</sup> Case fatality rate was 49.1% for patients with ILDs.

Our analysis of a large propensity matched cohort of patients adds important and novel data on the disease course of COVID-19 in patients with IPF. This is the first report of COVID-19 outcomes in both hospitalised and non-hospitalised patients with IPF. As previous data were limited to hospitalised patients, this likely resulted in omission of the less severe cases, and thus explains the lower overall mortality rate of 15.94% noted in our cohort. We found that 44% of patients with COVID-19 and comorbid

IPF require hospitalisation. In the subgroup analysis of hospitalised patients with IPF only, mortality rate of one in three was observed, closer to earlier smaller reports, and thus provides a sensitivity analysis to our hypothesis. Regardless, a significantly higher composite poor outcome, critical care need and hospitalisation rate was observed in patients with IPF after matching, with lower survival. Rates of acute kidney injury were also higher in the IPF cohort. Thus, COVID-19 disease carries poor outcomes in patients with coexisting IPF.

The main limitations of our study include the biases inherent to all studies conducted on electronic health-care record data. Care was taken by the study authors to limit study outcomes and variables to those that are less likely to suffer from such bias. Furthermore, we included patients based on COVID-19 test results and diagnosis. Patients who remain asymptomatic during their disease course may go untested and undiagnosed, and thus some of these asymptomatic patients may have remained uncaptured in our cohort. Thus, our analysis may, to some degree, reflect a more severe end of the COVID-19 disease spectrum.

## CONCLUSION

In conclusion, worse outcomes of COVID-19 disease were observed after robust matching of confounders. Our data confirm that patients with IPF constitute a high-risk cohort for poor outcomes related to COVID-19 disease. Therefore, we suggest that patients with IPF should be advised to strictly adhere to prevention guidelines. Consideration for prioritisation of these patients in preventative efforts like vaccination may also be considered.

### Author affiliations

- Section of Pulmonary and Critical Care Medicine, Department of Medicine, West Virginia University, Morgantown, West Virginia, USA
- Radiology, West Virginia University, Morgantown, West Virginia, USA
- General Surgery, NYU Langone Health, New York, New York, USA
- Internal Medicine, West Virginia University, Morgantown, West Virginia, USA
- Section of Infectious Disease, Department of Medicine, West Virginia University, Morgantown, West Virginia, USA
- Department of Medicine, West Virginia University, Morgantown, West Virginia, USA

**Twitter** Dhairya A Lakhani @dhairyalakhani

**Contributors** Acquisition of the data: YH, SFN and DAL. Analysis of the data: YH and AHS. Manuscript writing: YH, SFN, DAL, SS, AHS, JM and AS. Project supervision: AS.

**Funding** This work was supported in part by West Virginia Clinical and Translational Science Institute via USA National Institute of Health (NIH) grants U54 GM104942.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** TriNetX has received a waiver from Western Institutional Review Board as it only provides aggregate accounts.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Data can be made available upon reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been



peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Syeda Fatima Naqvi <http://orcid.org/0000-0001-7098-0321>  
Dhairya A Lakhani <http://orcid.org/0000-0001-7577-1887>

#### REFERENCES

- 1 Duchemann B, Annesi-Maesano I, Jacobe de Naurois C, *et al*. Prevalence and incidence of interstitial lung diseases in a multi-ethnic county of greater Paris. *Eur Respir J* 2017;50. doi:10.1183/13993003.02419-2016. [Epub ahead of print: 03 08 2017].
- 2 Olson AL, Gifford AH, Inase N, *et al*. The epidemiology of idiopathic pulmonary fibrosis and interstitial lung diseases at risk of a progressive-fibrosing phenotype. *Eur Respir Rev* 2018;27:180077.
- 3 Hyldgaard C, Hilberg O, Bendstrup E. How does comorbidity influence survival in idiopathic pulmonary fibrosis? *Respir Med* 2014;108:647–53.
- 4 Docherty AB, Harrison EM, Green CA, *et al*. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985.
- 5 Zhou F, Yu T, Du R. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
- 6 Azadeh N, Limper AH, Carmona EM, *et al*. The role of infection in interstitial lung diseases: a review. *Chest* 2017;152:842–52.
- 7 Esposito AJ, Menon AA, Ghosh AJ, *et al*. Increased odds of death for patients with interstitial lung disease and COVID-19: a case-control study. *Am J Respir Crit Care Med* 2020;202:1710–3.
- 8 Gallay L, Uzunhan Y, Borie R. Risk factors for mortality following COVID-19 in patients with pre-existing interstitial lung disease. *Am J Respir Crit Care Med* 2020;203.
- 9 Drake TM, Docherty AB, Harrison EM, *et al*. Outcome of hospitalization for COVID-19 in patients with interstitial lung disease. An international multicenter study. *Am J Respir Crit Care Med* 2020;202:1656–65.

## e-Appendix 1: Additional details of study methodology

### Patient selection:

COVID-19 patients were identified using CDC COVID-19 coding guidance. Patients were identified by using International Classification of Diseases, Ninth Revision and Tenth Revision, Clinical Modification (ICD-10-CM) codes as well as LOINC codes for positive laboratory tests discussed below. This search strategy has been previously used in published data from the TriNetX database.

ICD-10-CM codes U07.1 (COVID-19, virus identified), OR B34.2 (Coronavirus infection, unspecified), OR B97.29 (Other coronavirus as the cause of diseases classified elsewhere) OR J12.81 (Pneumonia due to SARS-associated coronavirus). Patients were excluded if they had diagnosis code 079.89 (Other specified viral infection). This code is mapped to ICD-10 code B34.2 and B97.2, and it was excluded to prevent false positives because it is used as a catch all code sometimes for many viral infections.

The following LOINC codes were also used to identify COVID-19 patients with positive COVID-19 test results.

94533-7: SARS coronavirus 2 N gene [Presence] in Respiratory specimen by NAA with probe detection

94534-5: SARS coronavirus 2 RdRp gene [Presence] in Respiratory specimen by NAA with probe detection

41458-1 SARS coronavirus RNA [Presence] in Unspecified specimen by NAA with probe detection

94309-2 SARS coronavirus 2 RNA [Presence] in Unspecified specimen by NAA with probe detection

94531-1 SARS Coronavirus 2 RNA panel - Respiratory specimen by NAA with probe detection

94506-3 SARS coronavirus 2 IgM Ab [Units/volume] in Serum or Plasma by Immunoassay

94500-6 SARS coronavirus 2 RNA [Presence] in Respiratory specimen by NAA with probe detection

94315-9 SARS coronavirus 2 E gene [Presence] in Unspecified specimen by NAA with probe detection.

94316-7 SARS-CoV-2 (COVID19) N gene [Presence] in Unspecified specimen by NAA with probe detection

94502-2 SARS-related coronavirus RNA [Presence] in Respiratory specimen by NAA with probe detection

Study duration: Enrolment between January 20, 2020, and November 15, 2020 were required for inclusion in the study cohort. January 20, 2020 was chosen as it was the date of diagnosis of the first case of SARS-CoV-2 infection in the USA. November 15, 2020 was chosen to ensure 60 days of follow up for all included patients, since the primary study endpoint was a composite outcome at 60 days from diagnosis

Patient identification period was limited from January 20, 2020 to November 15, 2020. January 20, 2020 was chosen as the initial date as it was the date of diagnosis of the first case of COVID-19 in the USA. November 15, 2020 was chosen as the last date so that all patients had at least 60 days follow up data available, as the primary study endpoint included a composite outcome at 60 day follow up.

Patients with ages 16 years or more at index event were included.

Idiopathic Pulmonary Fibrosis (IPF) patients were identified using the International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification (ICD-10-CM) codes given below:

**D86.** IPF code J84.1.

#### **Study definitions for data aggregation:**

##### **Index event:**

Diagnosis of COVID 19 per criteria defined above was considered the index event for the purposes of our study.

##### **Time windows:**

Study outcomes were assessed 30 days and 60 post index event and included the day of index event. All laboratory values were assessed within 7 days from COVID-19 diagnosis. Baseline characteristics for patients were considered until the day of diagnosis with COVID-19 or first positive COVID-19 test.