**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 COVD-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.