

Title: Outcomes of SARS-CoV-2 infection in patients with pulmonary sarcoidosis: A multicenter retrospective research network study

Authors: Yousaf B Hadi^{1*}, Dhairya A Lakhani^{2*}, Syeda F Z Naqvi³, Shailendra Singh¹, Justin T Kupec¹

*Denotes equal contribution

Affiliation: ¹Department of Internal Medicine; ²Department of Radiology; ³Section of Pulmonary and Critical Care Medicine, Department of Medicine; West Virginia University; Morgantown, WV.

Correspondence:

Justin T Kupec, MD
Department of Medicine
1 Medical Center Drive, Morgantown, WV 26506
jkupec@hsc.wvu.edu

Abbreviation: BMI: Body Mass Index; COVID-19: Coronavirus disease 2019; ICD-10-CM: International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; HCOs: Health Care Organizations; EHRs: Electronic health record management systems

Keywords: Coronavirus; COVID-19; interstitial lung disease; sarcoidosis; SARS-CoV-2.

Conflicts of interest: All authors have no relevant disclosures.

Funding source: Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number 5U54GM104942-05. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health

Guarantor: YBH

Author's contribution: Acquisition of the data: YBH, DAL, SFN; Analysis of the data: YBH, DL, SFN; Interpretation of the data: YBH, DAL, SFN; Drafting the work or revising it critically for important intellectual content: YBH, DAL, SFN, SS, JTK; Final approval of the version to be published: YBH, DAL, SFN, SS, JTK

e-Appendix: 1

Table: 3

Figure: 1

Word count: Main Text: 1862; Abstract: 250

Abstract

Background: Analyses of COVID-19 infection outcomes in patients with preexisting pulmonary sarcoidosis are lacking and are limited to case reports or small case series with the largest study reporting outcomes of 37 patients.

Research Question: Retrospective cohort study to assess clinical outcomes of 945 patients with pulmonary sarcoidosis, presenting with COVID 19, compared to a propensity matched cohort of patients without sarcoidosis.

Study Design and Methods: Analysis of a multi-center research network TriNETX was performed including patients more than 16 years of age diagnosed with COVID-19. Outcomes in COVID-19 positive patients with concurrent pulmonary sarcoidosis were compared with a propensity score matched cohort of patients without pulmonary sarcoidosis.

Results: A total of 278,271 patients with COVID-19 on the research network were identified, 954 patients (0.34%) carried a diagnosis of pulmonary sarcoidosis. Mean age of patients with sarcoidosis was 56.3 ± 13.2 years, with female predominance (n=619, 64.89%). 49.69% of the participants were African American (n=474). Co-morbidities including hypertension, chronic lower respiratory diseases, diabetes mellitus, ischemic heart disease, nicotine dependence, and chronic kidney disease were more common in patients with pulmonary sarcoidosis when compared to the non-pulmonary sarcoidosis cohort (all p values <0.01). In unmatched analysis, pulmonary sarcoidosis group had higher mortality, increased risk for hospitalization, intubation and need for renal replacement therapy. After propensity score matching, no difference in any of the outcome measures was observed.

Interpretation:

Crude COVID-19 mortality and other clinical outcome measures are poor in pulmonary sarcoidosis cohort; however, propensity-matched analyses revealed no difference in outcomes, showing that higher mortality is driven by higher burden of comorbidities.

Main Text

Background:

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) mainly affects the lung parenchyma. Sarcoidosis is an autoinflammatory disease characterized by diffuse granulomatous inflammation in the lungs and other organs.¹ For patients with pulmonary or autoimmune disorders, and those on immune modulating therapies, COVID-19 pandemic represents a therapeutic challenge as the outcome of SARS-CoV-2 infection in these populations continue to be understudied.²

Analyses of SARS-CoV-2 infection outcomes in patients with preexisting interstitial lung disease (ILD) or pulmonary sarcoidosis is lacking but registries are ongoing. Data on COVID-19 disease in patients with sarcoidosis are currently limited to case reports or small case series with the largest study reporting outcomes of 37 patients.³ It is previously known that patients with concomitant pulmonary disease have worse outcomes.^{2,4} Therefore, infected patients with any underlying chronic lung disease, especially with pulmonary fibrosis and sarcoidosis, would be theoretically at higher risk of severe complications. Some preliminary data also suggest that patients with sarcoidosis are at risk of severe complications,⁴ however, the small numbers of included patients do not allow for comparisons with patients without sarcoidosis.

Here, we studied the clinical outcomes of 945 patients with pulmonary sarcoidosis presenting with COVID 19 and compared them with a propensity matched cohort of patients without pulmonary sarcoidosis.

Study Design and Methods:

Description of data source

A retrospective cohort study was conducted using the multi-institutional research network TriNETX (Cambridge, MA, USA) platform. This platform is a federated research network of more than 40 Health Care Organizations (HCOs) in the United States of America. TriNETX provides real time access to the healthcare records of includes patients from the healthcare organizations (> 40 million), that are de-identified and retrieved directly from the electronic health record management systems (EHRs) of participating organizations. Participating organizations are generally large academic centers that operate both tertiary care and satellite secondary or primary office locations.

Clinical variables (referred as facts) are retrieved both directly from EHRs, as well as through processing clinical documents through a built-in Natural Language Processing system. Quality assurance is achieved at the time of data extraction from EHRs before inclusion in the federated network's database in a systemic, standardized format. Only aggregate counts and statistical summaries are provided by the interface, and no protected health information is

released by maintaining de-identification of subjects at all levels of data retrieval and dissemination.

Western IRB has granted a waiver to TriNetX as it only provides de-identified data. At our institution, West Virginia University Clinical and Translational Science Institute (WVCTSI) manages the TriNetX platform and provides access to end-users. TriNetX obfuscates patient counts less than 10 to ensure anonymity.

Study participants:

We conducted a real time search on the TriNetX platform and updated through January 06, 2021. All patients with COVID 19 on this platform who were more than 16 years of age at diagnosis were identified and included. Patients were included if they received a positive SARS-CoV-2 test or COVID-19 diagnosis between January 20, 2020, and October 30, 2020. January 20, 2020 was chosen as it was the date of diagnosis of the first case of COVID-19 in the USA. October 30, 2020 was chosen so that all included patients had 60 days of follow up available, as the primary study endpoint was a composite outcome at 60 days from diagnosis.

We based the selection and identification criteria on diagnostic codes and positive laboratory confirmation using standard terminology recommended by the Centers for Disease Control and Prevention and purported by TriNetX platform, which is detailed in the e-Appendix-1. Patients with COVID-19 identified through this terminology were categorized in groups based on previous diagnosis of pulmonary sarcoidosis. Patients with pulmonary sarcoidosis were identified using the International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification (ICD-10-CM) code D86.0.

Study outcomes:

Clinical and laboratory outcomes were then compared between patients with and without a diagnosis of pulmonary sarcoidosis. Primary study outcome was a composite of death or need of mechanical ventilation in the 30-day and 60-day period from index event. Index event was defined as either the time of COVID-19 diagnosis or first COVID-19 positive test result date, whichever occurred first. Secondary study outcomes included mortality, hospitalization, and need for mechanical ventilation in the 30-day period from COVID-19 diagnosis and mortality in 60 days after diagnosis. Laboratory outcomes were considered up to 7 days from COVID-19 diagnosis.

Statistical analysis:

All statistical analyses were conducted utilizing TriNetX platform in real-time. Cohort characteristics were described using mean, standard deviation and proportions whenever appropriate. For univariate analysis, chi square test and independent sample T tests employed for categorical and continuous data respectively, as appropriate. Subsequently, the TriNetX

platform was used for propensity score matching of the two study cohorts. 1:1 propensity score matching was conducted using race, age, diabetes, hypertension, chronic lung diseases, nicotine dependence, heart failure, ischemic heart disease, body mass index (BMI), chronic kidney disease and gender in order to match a control group of patients without sarcoidosis. The platform utilizes the input matrices of the identified and above-mentioned covariates to conduct logistic regression analysis that generates and assigns propensity scores to subjects. One-to-one matching was then performed based on the propensity scores thus generated using greedy nearest neighbor algorithms with a caliper width of 0.1 pooled standard deviations. In order to eliminate bias resulting from nearest neighbor algorithms, the orders of rows are randomized.

We defined a two-sided alpha of less than 0.05 a priori for statistical significance. We assessed the balance on covariates in the propensity score matched cohorts using standardized mean difference, and absolute values > 0.1 were considered reflective of residual imbalance in the two groups. For comparison between cohorts, risk ratios with 95% confidence intervals were calculated.

Results:

Study population:

Using our inclusion criteria, we identified a total of 278,271 patients with COVID-19 on the research network. Of these identified patients, 954 patients (0.34%) carried a diagnosis of pulmonary sarcoidosis, while the rest 277,317 patients (99.66%) were included in the non-pulmonary sarcoidosis cohort. Mean age at the time of COVID-19 diagnosis in patients with pulmonary sarcoidosis was 56.3 ± 13.2 years, with female predominance (n= 619, 64.89%) and 49.69% (n=474) were African American. Co-morbidities including hypertension, chronic lower respiratory diseases, diabetes mellitus, ischemic heart disease, nicotine dependence, and chronic kidney disease were more common in patients with pulmonary sarcoidosis when compared to the non-pulmonary sarcoidosis cohort (all p values <0.01). Baseline demographic characteristics and comorbid conditions are described in Table 1. Further, no patients with pulmonary sarcoidosis were lost in 1:1 propensity score matching analysis (Table 1).

Of the patients with pulmonary sarcoidosis, 544 patients (57%) were prescribed glucocorticoids in the 6 months before index event, and 617 (65%) were using bronchodilator medications. Methotrexate was used by 57 patients (6%). Geographic distribution of the participants with sarcoidosis is as follows: Southern USA (n=485, 51%), Mid-west USA (n= 197, 21%), Northern USA (n=191, 20%), Western USA (n=70, 7%), and other/unknown location (n=11, <1%).

Clinical outcomes:

All the covariates used for matching in the two groups were similar after propensity score matching (mean standard difference < 0.1) and are illustrated in Table 1. In the 30 and 60-day period post COVID-19 diagnosis, 33 (3.46%) and 41 (4.30%) deaths were reported in the sarcoidosis group, respectively. The composite outcome (death or mechanical ventilation) was reached by 58 (6.08%) and 62 (6.50%) patients in the pulmonary sarcoidosis group at 30 and 60 day follow up respectively (Figure 1, Table 2). In the crude, unmatched analysis, mortality and composite endpoint at 30 days was higher in the pulmonary sarcoidosis group (RR 1.92, 95%CI 1.37-2.68) and (RR 2.04, 95%CI 1.59-2.61) respectively. Similar trends were observed for mortality and composite endpoint at 60 days, as well as need for hospitalization, critical care need, acute kidney injury, and requirement for renal replacement therapy (Table 2). However no statistically significant difference was observed following propensity matching, for mortality, composite endpoint, acute kidney injury, requirement for renal replacement therapy, inpatient admission and critical care need in the two groups, as illustrated in Table 2.

Laboratory values:

Table 3 describes mean values of C-reactive protein (mg/L), Lactate dehydrogenase (U/L), Erythrocyte sedimentation rate (mm/hr), Alanine aminotransferase (U/L), Aspartate aminotransferase (U/L), Serum Bilirubin (mg/dL) and Serum Ferritin (ng/ml). We observed statistically significant difference in mean values of C-Reactive protein and Alanine Aminotransferase, both values were significantly higher in non-sarcoidosis group (Table 3).

Discussion:

In our large cohort of propensity score matched analysis, previous diagnosis of pulmonary sarcoidosis was not independently associated with poor COVID-19 outcomes. We have noted a COVID-19 mortality rate of 4.3% in patients with pulmonary sarcoidosis. Interestingly, our findings are in stark contrast to previously available limited data on risk of severe COVID-19 disease in pulmonary sarcoidosis. Some previous reports of COVID-19 in patients with sarcoidosis pointed toward an increased risk of severe disease, however, no matched analysis was available to compare outcomes with a non-pulmonary sarcoidosis cohort.^{5,6} Morgenthau et al in their cohort of 37 patients with sarcoidosis reported a mortality rate of 16.2% (n=6).³ Similarly, in a French multicenter registry of COVID19 patients that included 36 patients with sarcoidosis, 14% mortality (n=5) was reported.⁷ The markedly higher reported mortality in earlier smaller studies was likely secondary to inclusion of only hospitalized patients, who are expected to carry higher disease related morbidity. We observed that a small proportion of pulmonary sarcoidosis patients with COVID-19 (less than 1 in 5) require hospitalization, with most being managed as outpatients. Inclusion of both hospitalized patients as well as out-patients in our cohort explains the lower mortality observed in our report and paints a truer 'real world' picture of COVID-19 related morbidity in pulmonary sarcoidosis patients.

Nonetheless, in this large cohort, the crude rates of mortality and critical illness in patients with pulmonary sarcoidosis were higher when compared to those without pulmonary

sarcoidosis. Higher numbers of these patients required hospitalization and mechanical ventilation and developed complications including renal injury requiring renal replacement therapy. Thus, it can be inferred that patients with pulmonary sarcoidosis comprise a high-risk cohort for COVID-19 related severe disease. The fact that the increase in risk is not sustained following robust propensity matching implies that this excess risk is derived from higher burden of comorbid diseases and other risk factors for severe COVID-19 prevalent in the pulmonary sarcoidosis group.

Strengths of our study include the large-included cohort of patients with sarcoidosis and robust control of confounders. This is also the first report that captures both inpatient and ambulatory sarcoidosis patients with COVID-19 and allows for comparison with a non-pulmonary sarcoidosis control cohort.

Propensity score was employed for patient matching. In this method, a single propensity score is assigned to each patient summarizing all measured confounders and then cases and controls are matched. This provides a simple tool for matched analysis, to conventional independent matching on all of the covariates.⁸ Limitations to this technique include: i) The technique cannot balance unmeasured characteristics and confounders; ii) The analysis does not “fix” any other potential methodologic biases that may exist in retrospective studies and studies conducted on electronic healthcare record data.⁸ Care was taken by the study authors to limit study outcomes and variables to those that are less likely to suffer from such bias. We also were not able to study the effect of immunosuppressant use, presence of malignancy, pregnancy and other immunosuppressive states like organ transplantation on outcomes in sarcoidosis cohort. These limitations could be addressed by prospective studies. Further, since only patients with pulmonary sarcoidosis were included, excluding multi-organ sarcoidosis, the outcomes in those with respect to COVID-19 may be different.

In conclusion, in a large research network study, we found that previous diagnosis of sarcoidosis does not independently increase risk of worse COVID-19 outcomes, however there is high COVID-19 related morbidity in these patients owing to higher prevalence of risk factors of severe COVID. These data emphasize the importance of adherence to preventative measures in this high-risk cohort.

Tables:

Table 1. Baseline demographic and clinical characteristics of COVID-19 patients with and without pulmonary sarcoidosis before and after propensity score matching.

Table 2. Outcomes in COVID-19 patients with and without pulmonary sarcoidosis before and after propensity score matching.

Table 3. Laboratory values after COVID 19 episodes in the two matched and unmatched groups

Figure Legends:

Figure 1: Kaplan Meier plots of composite endpoint (mortality and mechanical ventilation combined) in SARS-CoV-2 infected patients with pulmonary sarcoidosis (purple) and without pulmonary sarcoidosis (green), before (Figure 1A) and after propensity matching (Figure 1B).

e-Appendix-1: Description of diagnostic codes used to identify COVID-19 patients.

Reference:

1. Calender A, Israel-Biet D, Valeyre D, Pacheco YJTil. Modeling potential autophagy pathways in COVID-19 and sarcoidosis. 2020;41(10): 856-859.
2. Györfi AH, Kopp M, May M, et al. Glucocorticoid-induced relapse of COVID-19 in a patient with sarcoidosis. 2020.
3. Morgenthau AS, Levin MA, Freeman R, Reich DL, Klang EJJ. Moderate or severe impairment in pulmonary function is associated with mortality in sarcoidosis patients infected with SARS-CoV-2. 2020;198(5): 771-775.
4. Southern BJCCJoM. Patients with interstitial lung disease and pulmonary sarcoidosis are at high risk for severe illness related to COVID-19. 2020.
5. Gianfrancesco M, Hyrich KL, Al-Adely S, et al. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: data from the COVID-19 Global Rheumatology Alliance physician-reported registry. 2020.
6. Sweiss NJ, Korsten P, Syed H, et al. When the game changes: Guidance to adjust sarcoidosis management during the COVID-19 pandemic. 2020.
7. Jeny F, Lhote R, Lorillon G, et al. Correspondence on 'Glucocorticoid-induced relapse of COVID-19 in a patient with sarcoidosis'. 2020.
8. Okoli GN, Sanders RD, Myles P. Demystifying propensity scores. Br J Anaesth. 2014;112(1):13-15.

Table 1. Baseline demographic and clinical characteristics of COVID-19 patients with and without pulmonary sarcoidosis before and after propensity score matching.

| | Before propensity matching | | | | | After propensity matching | | | | |
|-----------------------------------|----------------------------|----------------|-----------------------------------|----------------|-------------------|----------------------------|----------------|--------------------------------|----------------|-------------------|
| | Sarcoidosis cohort (n=954) | | Non-sarcoidosis cohort (n=277317) | | p-value/ SDiff | Sarcoidosis cohort (n=954) | | Non-sarcoidosis cohort (n=954) | | p-value/ SDiff |
| Variable | Number/Mean | Percent/ SD | Number | Percent/ SD | | Number/Mean | Percent/ SD | Number | Percent/ SD | |
| Demographics | | | | | | | | | | |
| Age (Mean) | 56.3 | 13.2 | 45.8 | 19 | <0.001/ 0.64 | 56.3 | 13.2 | 57.4 | 14.9 | 0.11 / 0.07 |
| Male | 333 | 34.91 | 123793 | 44.64 | <0.001/0.20 | 333 | 34.91 | 322 | 33.75 | 0.59 / 0.02 |
| Female | 619 | 64.89 | 152513 | 54.99 | <0.001/0.20 | 619 | 64.89 | 632 | 6.25 | 0.53 / 0.03 |
| BMI (30 and above) | 310 | 32.50 | 41049 | 14.80 | <0.001 / 0.42 | 310 | 32.50 | 315 | 33.02 | 0.81 / 0.01 |
| Black or African American | 474 | 49.69 | 53122 | 19.16 | <0.001 / 0.68 | 474 | 49.68 | 475 | 49.79 | 0.96 / 0.002 |
| White | 410 | 42.98 | 157160 | 56.67 | <0.001/0.28 | 410 | 42.98 | 405 | 42.45 | 0.82 / 0.01 |
| Hispanic or Latino | 58 | 6.08 | 44958 | 16.21 | <0.001/0.32 | 58 | 6.08 | 66 | 6.91 | 0.45 / 0.03 |
| Comorbidities | | | | | | | | | | |
| Hypertension | 639 | 66.98 | 72443 | 26.12 | <0.001 / 0.89 | 639 | 66.98 | 650 | 68.13 | 0.59 / 0.02 |
| Chronic lowerrespiratory diseases | 467 | 48.95 | 39927 | 14.39 | <0.001 /0.8 | 467 | 48.95 | 466 | 48.85 | 0.96 / 0.002 |
| Diabetes mellitus | 368 | 38.57 | 37447 | 13.50 | <0.001 / 0.59 | 368 | 38.57 | 348 | 36.48 | 0.34 / 0.04 |
| Ischemic heart disease | 261 | 27.35 | 21228 | 7.66 | <0.001 / 0.54 | 261 | 27.36 | 252 | 26.42 | 0.64 / 0.02 |
| Nicotine dependence | 135 | 14.15 | 17943 | 6.47 | <0.001 / 0.25 | 135 | 14.15 | 130 | 13.63 | 0.74 / 0.02 |
| Chronic kidney disease | 184 | 19.28 | 16125 | 5.82 | <0.001 / 0.42 | 184 | 19.29 | 170 | 17.82 | 0.41 / 0.04 |

Table Legend: Abbreviation: BMI: Body Mass Index; SDiff: Standard Difference; SD: Standard Deviation

Table 2. Outcomes in COVID-19 patients with and without pulmonary sarcoidosis before and after propensity score matching.

| Outcome | Sarcoidosis group (n=954) | Percentage | Non-Sarcoidosis group (n=277317) | Percentage | Risk Ratio | 95 % CI Lower | 95 % CI Upper |
|---|------------------------------|------------|-------------------------------------|------------|------------|------------------|------------------|
| Before propensity score matching | | | | | | | |
| Mortality within 30 days | 33 | 3.46 | 5004 | 1.80 | 1.92 | 1.37 | 2.68 |
| Mortality within 60 days | 41 | 4.30 | 5704 | 2.06 | 2.09 | 1.55 | 2.82 |
| Inpatient services | 185 | 19.40 | 24563 | 8.86 | 2.19 | 1.92 | 2.49 |
| Critical Care | 66 | 6.92 | 8455 | 3.05 | 2.27 | 1.80 | 2.87 |
| Mechanical ventilation | 42 | 4.40 | 5366 | 1.94 | 2.28 | 1.69 | 3.06 |
| 30-day composite outcome | 58 | 6.08 | 8282 | 2.99 | 2.04 | 1.59 | 2.61 |
| 60-day composite outcome | 62 | 6.50 | 8746 | 3.15 | 2.06 | 1.62 | 2.62 |
| Acute Kidney Injury | 97 | 10.17 | 10956 | 3.95 | 2.57 | 2.13 | 3.11 |
| Renal replacement therapy | 18 | 1.89 | 1192 | 0.43 | 4.39 | 2.77 | 6.96 |
| Outcome | Sarcoidosis group (n=954) | Percentage | Non-Sarcoidosis group (n=954) | Percentage | Risk Ratio | 95 % CI Lower | 95 % CI Upper |
| After propensity score matching | | | | | | | |
| Mortality within 30 days | 33 | 3.46 | 42 | 4.40 | 0.79 | 0.50 | 1.23 |
| Mortality within 60 days | 41 | 4.30 | 45 | 4.72 | 0.91 | 0.60 | 1.38 |
| Inpatient services | 185 | 19.39 | 166 | 17.4 | 1.11 | 0.92 | 1.35 |
| Critical Care | 66 | 6.92 | 62 | 6.50 | 1.07 | 0.76 | 1.49 |
| Mechanical ventilation | 42 | 4.40 | 38 | 3.98 | 1.11 | 0.72 | 1.70 |
| 30-day composite outcome | 58 | 6.08 | 62 | 6.50 | 0.94 | 0.66 | 1.32 |
| 60-day composite outcome | 62 | 6.50 | 64 | 6.71 | 0.97 | 0.69 | 1.36 |
| Acute Kidney Injury | 97 | 10.17 | 88 | 9.22 | 1.10 | 0.84 | 1.45 |
| Renal replacement therapy | 18 | 1.89 | 13 | 1.36 | 1.39 | 0.68 | 2.81 |

Table 3. Laboratory values after COVID 19 episodes in the two matched and unmatched groups

| Outcome | After propensity matching | | | | |
|---|-------------------------------|---------------------|------------------------------------|---------------------|---------|
| | Sarcoidosis cohort (n=954) | | Non- Sarcoidosis cohort (n=954) | | p-value |
| | Patients with Outcome | Mean (SD) | Patients with Outcome | Mean (SD) | |
| C Reactive protein (mg/L) | 201 | 54.53 (68.45) | 189 | 78.56 (91.36) | < 0.001 |
| Lactate dehydrogenase (U/L) | 158 | 389.22 (234.90) | 164 | 388.56 (217.52) | 0.98 |
| Erythrocyte sedimentation rate (mm/hr) | 78 | 57.49 (30.83) | 57 | 56.60 (30.03) | 0.86 |
| Alanine aminotransferase (U/L) | 297 | 30.94 (25.41) | 279 | 40.32 (74.52) | 0.04 |
| Aspartate aminotransferase (U/L) | 296 | 38.03 (45.53) | 278 | 51.79 (143.25) | 0.12 |
| Serum Bilirubin (mg/dL) | 299 | 0.60 (0.48) | 279 | 0.66 (1.30) | 0.43 |
| Serum Ferritin (ng/ml) | 165 | 824.44 (1177.06) | 166 | 965.01 (1201.80) | 0.28 |

Table Legend: Abbreviation: SD: Standard Deviation

e-Appendix-1: Description of diagnostic codes used to identify COVID-19 patients.

Center of Disease Control and Prevention (CDC) coronavirus (COVID-19) coding guidance was used to identify COVID-19 cases. International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification (ICD-10-CM) codes and Logical Observation Identifiers Names and Codes (LOINC) codes for positive laboratory tests were used. This search strategy is previously used in published data from the TriNetX.

ICD-10-CM Codes Strategy:

OR B34.2 (Coronavirus infection, unspecified), ICD-10-CM codes U07.1 (COVID-19, virus identified), OR J12.81 (Pneumonia due to SARS-associated coronavirus), OR B97.29 (Other coronavirus as the cause of diseases classified elsewhere).

079.89 (Other specified viral infection) were excluded. This code is mapped to ICD-10 code B34.2 and B97.2, which were excluded to prevent false positives as it captures all code for many viral infections.

LOINC codes strategy:

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 definitions for data aggregation:

Index event: Diagnosis of COVID 19 by criteria defined above.

Time windows: Study outcomes were assessed at 30 days and at 60 from index event and included the day of index event. Laboratory values were assessed within 7 days from COVID-19 diagnosis. Baseline characteristics for patients were considered until the day of COVID-19 diagnosis or first positive COVID-19 test result.