High Dimensional Characterization of Post-acute Sequalae of COVID-19: analysis of health outcomes and clinical manifestations at 6 months

Ziyad Al-Aly (zalaly@gmail.com)
Washington University School of Medicine https://orcid.org/0000-0002-2600-0434
Yan Xie
VA Saint Louis Health Care System
Benjamin Bowe
VA Saint Louis Health Care System

Biological Sciences - Article

Keywords: severe acute respiratory syndrome coronavirus 2, COVID-19, post-acute sequalae

Posted Date: January 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-150398/v1

Version of Record: A version of this preprint was published at Nature on April 22nd, 2021. See the published version at https://doi.org/10.1038/s41586-021-03553-9.

Abstract

The coronavirus disease 2019 (COVID-19) is a viral illness caused by the severe acute respiratory syndrome coronavirus 2. The acute clinical manifestations of COVID-19 are well characterized. The post-acute sequalae of COVID-19 have not been comprehensively described. Here, we use the national healthcare databases of the US Department of Veterans Affairs to undertake a high dimensional approach to comprehensively identify 6-months outcomes of incident clinical manifestations including diagnoses, medication use, and laboratory abnormalities in people who survived the first 30 days of COVID-19. We first describe the clinical manifestations in people with COVID-19 compared to users of the Veterans Affairs healthcare system. We then provide a comparative evaluation of the post-acute sequalae in 30-day survivors who were hospitalized for COVID-19 vs. seasonal influenza. We show that beyond the first 30 days of illness, people with COVID-19 are at higher risk of death and health resource utilization. Our approach identifies incident clinical manifestations in the respiratory system and several other manifestations including the nervous system and neurocognitive disorders, mental health disorders, metabolic disorders, cardiovascular disorders, gastrointestinal disorders, and signs and symptoms related to poor generalized wellbeing including malaise, fatigue, musculoskeletal pain, and anemia. There was increased incident use of pain medications (opioids and non-opioids), antidepressants, anxiolytics, antihypertensives, antihyperlipidemic, insulin, and several other medication classes. The findings show that beyond the first 30 days of illness, substantial burden of health loss — spanning pulmonary and several extrapulmonary organ systems — is experienced by COVID-19 survivors. The results provide a roadmap to inform health system planning and development of multidisciplinary care strategies aimed at reducing chronic health loss and optimizing wellness among COVID-19 survivors.

Main

The coronavirus disease 2019 (COVID-19) is a viral illness caused by the severe acute respiratory syndrome coronavirus. The acute clinical manifestations of COVID-19 are well characterized and involve both pulmonary and extrapulmonary systemic manifestations[1]. Emerging reports of “long haulers” suggest that beyond the acute setting, some patients with COVID-19 may experience persistent long-lasting clinical manifestations. However, the post-acute sequelae of COVID-19 are not yet clear. A better of the post-acute sequelae of COVID-19 and their burdens will inform care strategies to optimize wellness and reduce the burden of chronic and permanent health loss among people who had COVID-19.

Here we leveraged the breadth and depth of the US Department of Veterans Affairs electronic health databases to undertake a high dimensional approach to comprehensively identify the 6-months outcomes of incident diagnoses (from 379 diagnostic categories), incident medication use (from 380 medication classes), and incident laboratory abnormalities (from 62 laboratory tests) in people who survived the first 30 days of COVID-19.

Comparative evaluation of 6-months outcomes of COVID-19 vs. all users of the Veteran Health Administration

The cohort included 51,376 users of the Veteran Health Administration (VHA) with COVID-19 who survived at least the first 30 days after COVID-19 diagnosis, and 5,051,832 VHA users who did not have COVID-19 (Extended data Fig. 1a and 2a). The median follow-up and interquartile range were 152 (111, 188) and 153 (113, 191) days in the COVID-19 and VHA user groups (Table 1). Beyond the first 30 days of illness, COVID-19 survivors had increased risk of death (HR 2.17 (1.98, 2.38)). We also estimated the excess burden due to COVID-19 per 1000 persons at 6-months based on the difference of estimated incidence rate between COVID-19 and all users of VHA. The excess death was estimated at 65.42 (64.72, 66.06) per 1000 COVID-19 patients at 6-months. Risk of hospital admission was also higher among people with COVID-19 (HR 1.55 (1.47, 1.64); excess burden 86.28 (84.47, 87.99)). Those with COVID-19 had a higher risk of outpatient care encounter (HR 1.28 (1.26, 1.29); excess burden 120.48 (116.98, 124.03)) and at a greater frequency (1.01 (0.96, 1.05) additional encounters every 30 days) (Table 2).

In evaluating the risk of incident 379 diagnoses (categorized from ICD-10 codes based on Clinical Classifications Software Refined (CCSR)), 380 medication classes, and 62 laboratory tests beyond the first 30 days, several conditions in almost every organ system exhibited an adjusted hazard ratio greater than 1, and a p value lower than $6.57x10^{-5}$ (significance level adjusting for multiple comparisons). The hazard ratio and burden for all outcomes are presented in Fig. 1a-c and supplementary information Table 1–3. The result for outcomes that were positively associated with COVID-19 are presented in extended data Fig. 3. Conditions with excess burden higher than 10 per 1000 persons at 6-months are reported in Fig. 2a-c and extended data Table 1 and are detailed below:

Respiratory conditions

At six months following a COVID-19 infection, excess burden of respiratory conditions was most common and included respiratory signs and symptoms (46.18 (43.63, 48.57) per 1000 COVID-19 patients at 6-months), respiratory failure, insufficiency, arrest (24.73 (23.89, 25.47)), and lower respiratory disease (11.68 (10.80, 12.44)). There was also evidence of high burden of incident use of bronchodilators (36.26 (34.48, 37.89)), antitussives and expectorants (22.71 (21.3, 23.98)), anti-asthmatics (17.08 (15.69, 18.32)), and glucocorticoids (14.87 (12.54, 17.03)).

Diseases of the Nervous system

Excess burden of nervous system disorders was evident including nervous system signs and symptoms (excess burden 31.39 (28.82, 33.81) per 1000 COVID-19 patients at 6-months), neurocognitive disorders (11.87 (10.77, 12.84)), and other nervous system disorders (14.08 (12.56, 15.45)).

Mental health burden

The results showed excess burden of sleep wake disorders (22.66 (19.09, 26.01)) per 1000 COVID-19 patients at 6-months), depressive disorders (12.40 (9.48, 15.12)), anxiety and fear-related disorders (10.69 (8.33, 12.88)), and trauma and stress related disorders (16.45 (13.75, 18.95)). These findings were coupled with evidence of excess burden of incident use of non-opioid analgesics (45.07 (41.99, 47.96)),

opioid analgesics (19.01 (16.35, 21.49)), antidepressants (18.36 (15.79, 22.23)) and benzodiazepines and sedatives (10.76 (9.29. 12.08)).

Metabolic disorders

Excess burden of several metabolic disorders was evident including disorders of lipid metabolism (23.96 (18.96, 28.69) per 1000 COVID-19 patients at 6-months), diabetes (14.00 (11.73, 16.07)), and obesity (13.04 (10.67, 15.22)). These was also evidence of excess burden of incident use of antilipemic agents (19.14 (15.79, 22.23)), and insulin (16.19 (14.95, 17.29)) and excess burden of elevated levels of triglycerides (10.51 (6.67, 14.14)).

Poor general wellbeing

Survivors of COVID-19 exhibited excess burden of poor general wellbeing including malaise and fatigue (27.22 (25.54, 28.77) per 1000 COVID-19 patients at 6-months), muscle disorders (18.55 (17.19, 19.77)), musculoskeletal pain (15.70 (10.85, 20.32)) and anemia (18.55 (16.91, 20.05)). These diagnoses were coupled with laboratory evidence of excess burden of anemia (decreased hemoglobin (51.35 (47.79, 54.72), and decreased hematocrit levels (53.25 (49.44, 56.86)).

Cardiovascular conditions

Excess burden of cardiovascular conditions included hypertension (26.5 (21.99, 30.72) per 1000 COVID-19 patients at 6-months), cardiac dysrhythmias (17.53 (16.05, 18.86)), circulatory signs and symptoms (15.23 (13.43, 16.9)), chest pain (14.94 (13.16, 16.56)), coronary atherosclerosis (11.37 (9.59, 12.99)), and heart failure (11.22 (9.97, 12.34)). There was also evidence of excess burden of incident use of beta blockers (20.42 (18.49, 22.17)), calcium channel blockers (16.69 (14.85, 18.34)), and loop diuretics (13.57 (12.24, 14.77)).

Gastrointestinal system: There was evidence of excess burden of the following conditions: esophageal disorders (16.69 (13.88, 19.32) per 1000 COVID-19 patients at 6-months), gastrointestinal disorders (14.07 (12.26, 15.72)), dysphagia (12.77 (11.59, 13.83)), abdominal pain and other digestive/abdomen signs and symptoms (12.22 (9.75, 14.51)). These were coupled with evidence of increased use of laxatives (22.27 (20.22, 24.18)), gastric medications (26.70 (23.90, 29.32)), and antihistamines (10.99 (8.79, 13.00)). Laboratory abnormalities included increased risk of incident high levels of alanine aminotransferase (13.99 (11.28, 15.82)).

Other sequalae

There was also evidence of excess burden in incident use of anticoagulants (44.31 (42.48, 46.01) per 1000 COVID-19 patients at 6-months). Other conditions included excess burden of acute renal failure (15.55 (14.39, 16.59) per 1000 COVID-19 patients at 6-months), infections (including urinary tract infections (12.54 (11.27, 13.67)), and genitourinary signs and symptoms (10.03 (8.25, 11.67)) (Fig. 2a-c, extended data Table 1 and supplementary information Table 1–3).

Comparative evaluation of 6-months outcomes of hospitalized patients with COVID-19 vs. hospitalized patients with seasonal influenza

To gain a better understanding of the spectrum of clinical manifestations in survivors of COVID-19 who got hospitalized, we undertook a comparative evaluation in a cohort of hospitalized individuals with COVID-19 vs. those hospitalized with seasonal influenza (a well-known, well characterized respiratory viral illness). This approach will allow the identification of the post-acute manifestations that may be occurring at higher rates among hospitalized patients with COVID-19 than among hospitalized patients with seasonal influenza.

The hospitalized cohort included 8783 people with COVID-19 and 14053 people with influenza who survived at least 30 days after hospital admission (Extended data Fig. 1b and 2b). The median follow-up and interquartile range were 155 (112, 213) and 290 (290, 290) days in the COVID-19 and influenza groups (Table 1). Beyond the first 30 days of illness, COVID-19 survivors who had been hospitalized for COVID-19 had increased risk of death (HR = 1.45 (1.30, 1.62)); excess death was estimated at 25.37 (18.75, 31.33) per 1000 persons at 6-months. Those with COVID-19 exhibited a higher risk of outpatient care encounter (HR 1.16 (1.13, 1.20), excess burden 7.88 (6.73, 8.92))) and with greater frequency (1.97 (1.77, 2.18) additional encounters every 30 days) (Table 2).

Compared to those hospitalized with seasonal influenza, and beyond the first 30 days of illness, COVID-19 survivors who had been hospitalized for COVID-19 had a higher burden of a broad array of pulmonary and extrapulmonary systemic manifestations including neurologic disorders (nervous system signs and symptoms (35.09 (23.09, 46.14) per 1000 hospitalized COVID-19 patients) and neurocognitive disorders (22.83 (16.42, 28.39)), mental health disorders (e.g. anxiety and fear-related disorders (15.72 (9.11, 21.48)), metabolic disorders (43.17 (26.97, 57.93)), cardiovascular disorders (e.g. circulatory signs and symptoms (17.80 (9.90, 24.85)), gastrointestinal disorders (18.31 (10.26, 25.52)), coagulation disorders (14.25 (9.67, 18.13)) and pulmonary embolism (15.33 (12.67. 17.43)), and other disorders (Fig. 1d-f, Fig. 2d-f, extended data Table 2, extended data Fig. 4 and supplementary information table 4–6).

Discussion

In this study, we use a high dimensional approach to identify the spectrum of clinical abnormalities (incident diagnoses, incident medication use, and incident laboratory abnormalities) experienced by COVID-19 survivors beyond the first 30 days of illness. The results suggest that beyond the first 30 days of illness, people with COVID-19 are at higher risk of death, health care resource utilization, and exhibit a broad array of incident pulmonary and extrapulmonary clinical manifestations including nervous system and neurocognitive disorders, mental health disorders, metabolic disorders, cardiovascular disorders, gastrointestinal disorders, and signs and symptoms related to poor general wellbeing including malaise, fatigue, musculoskeletal pain, and anemia. Increased risk of incident use of several medication classes was also observed including pain medications (opioids and non-opioids), antidepressants, anxiolytics, antihypertensives, antihyperlipidemic, insulin, and other medication classes.

Our report provides a comprehensive identification of the 6-months clinical manifestations of COVID-19. Our approach to evaluate risk of COVID-19 vs. VHA users provides a more general picture of the overall post-acute clinical manifestations encountered by people infected with COVID-19. Our comparative evaluation of hospitalized patients with COVID-19 vs. seasonal influenza identifies risks in a sicker population (hospitalized individuals) and benchmarks it against a well characterized serious respiratory viral illness (seasonal influenza)[2]. Both approaches illuminate our understanding of the post-acute sequalae of COVID-19. We note that while both groups (COVID-19 and non-COVID-19) were balanced at baseline, following the first 30 days of COVID-19 infection, the exposure groups exhibited remarkably divergent health trajectories. People with COVID-19 exhibited increases in health resource utilization and excess burden of a broad array of clinical conditions. While some of the clinical manifestations may resolve with treatment and time, it is evident that some sequalae may be chronic consequences of long COVID-19 that will require long term management (e.g. incident mental health conditions, metabolic disorders, and cardiovascular conditions). The constellation of evidence suggests substantial burden of health loss and highlights the need for a holistic and integrated multidisciplinary long-term care of COVID-19 survivors.

The implications of our results for health systems are clear. As the number of COVID-19 cases continues to climb across the globe, health systems face the dual challenge of coping with surge in acute infections, and caring for COVID-19 survivors (accounting for more than 1% of the global population and growing) who will also likely require substantial care to mitigate permanent health loss. This will place additional demands on already strained health systems. Governments and health systems around the globe are actively devising plans to address the tide of COVID-19 survivors in need of post COVID-19 care. For example, the National Institute for Health and Care Excellence (NICE), the Scottish Intercollegiate Guidelines Network, and the Royal College of General Practitioners have developed a rapid living guideline (which will be updated as more evidence becomes available) for managing post-acute COVID-19[3]. NHS England released guidance for post COVID-19 assessment clinics[4]. Both NICE and NHS England noted the dearth of evidence in this area; they and other stakeholders highlighted the urgent need for research to better characterize the clinical phenotypes of post COVID-19 and the burdens of the medical conditions that need to be cared for[5,6]. Such evidence is needed to inform capacity planning and optimal composition of post COVID-19 clinics[7]. Our studies illuminate this important area and may be used to project need for health resources and inform the optimal multidisciplinary composition of post COVID-19 clinics.

The mechanism which underlies the post-acute and chronic manifestations of COVID-19 are not entirely clear. Some of the manifestations may be driven by a direct effect of the viral infection and may be putatively explained by several hypotheses including persistent virus in immune-privileged sites, aberrant immune response, hyperactivation of the immune system, or autoimmunity[8]. Indirect effects including changes in social (e.g. reduced social contact and loneliness), economic (e.g. loss of employment), and behavioral conditions (e.g. changes in diet and exercise) that may be differentially experienced by people with COVID-19 may also shape health outcomes in COVID-19 survivors and may be responsible drivers of

some of the clinical manifestations reported here[9–13]. Some of the burden reported here may not be related to COVID-19 but may be diagnosed during the course of receiving care for COVID-19 — this may represent a previously unrealized disease that was brought to prominence (and realized) as a consequence of the higher frequency of interaction with the health care system that is experienced by COVID-19 survivors. A better delineation of the direct and indirect effects and a deeper understanding of the underlying biologic mechanisms and epidemiologic drivers of the multifaceted long-term consequences of COVID-19 is needed[14].

Here we report estimates of 6-month excess burden of disease due to COVID-19 and refer to it as post-acute sequalae of COVID-19. Other terms used in the literature include “long COVID-19”, “post-COVID-19 syndrome” and “post-acute COVID-19 syndrome”, and people with symptoms and clinical manifestations beyond the acute phase have been referred to in the lay vernacular as “long haulers”[15]. We recognize that patient groups with lived experiences of post COVID-19 and other stakeholders hold a broad range of views regarding these terminologies[16,17]. Identification of scientifically accurate and culturally sensitive terms to describe the illness beyond the acute phase will be an important step not only to standardize scientific communications globally but also to support clear and consistent public health messaging about the long-term consequences of COVID-19.

Strengths And Weaknesses:

To our knowledge, this is the largest post-acute COVID-19 study to date involving 51,376 patients with COVID-19, and 5,051,832 controls (corresponding to 2,173,245.13 person years of follow-up), and 8738 hospitalized patients with COVID-19, and 14,053 patients hospitalized with seasonal influenza (corresponding to 14,740.07 person years of follow-up). We leveraged the breadth and depth of the US Department of Veterans Affairs national health care databases — the largest nationally integrated healthcare delivery system in the US — to undertake a comprehensive approach to identify the 6-months health outcomes and clinical manifestations in COVID-19 patients who survived the first 30 days of COVID-19 illness. The high dimensional nature and depth of VA data facilitate ascertainment of medical conditions not only through diagnostic codes but also through medications, and laboratory data. Our comparative approach to estimate risk (and burden) in people with COVID-19 vs users of the VHA and in people hospitalized COVID-19 vs. those hospitalized seasonal influenza facilitates the estimation of excess risk (and burden) relative to the control groups. The simultaneous examination of comprehensive set of incident diagnoses, incident medication use, and incident laboratory abnormalities in the same analytic framework allows the comparative evaluation of risks and ranking of burdens of these conditions — providing health care providers, health system planners, public health officials, and the public at large with a rank list of the post-acute clinical conditions encountered in COVID-19 survivors. For each outcome examined, we built a cohort free of the related outcome at baseline to identify the risk of incident outcome during follow up — this approach allows the identification of incident clinical manifestations and abnormalities following COVID-19 infection. We used the overlap weighting method to balance exposure groups using several pre-defined covariates as well as a high dimensional variable selection algorithm to enhance the ability to establish a comparative evaluation of risk. While we

conducted survival analyses to estimate the risk of each outcome examined, we — for each outcome — also estimated the excess burden per 1000 persons due to COVID-19.

This study has several limitations. While our approach identifies the incident clinical manifestations in COVID-19 survivors, some of these manifestations may be direct or indirect consequences of COVID-19 infection, and some may be clinical manifestations that are unrelated to COVID-19 but discovered during the course of receiving care post COVID-19 illness. We examined 379 diagnostic categories, 380 medication classes, and 62 laboratory tests; our identification of the clinical manifestations is necessarily limited by the granularity of the diagnosis and medication grouping. Furthermore, our data shows increased risk of both outpatient care and inpatient admission among COVID-19 patients, but our approach does not detail whether the identified clinical manifestations are diagnosed during an inpatient or outpatient encounter. Finally, COVID-19 patients were enrolled in our cohorts from March 01 to October 15, 2020 and followed until December 15, 2020; as the COVID-19 global pandemic continues to evolve, and as treatment strategies improve, new variants of the virus emerge, and vaccine availability increases, it is likely that the epidemiology, short term, and long term outcomes of COVID-19 will likely also change over time.

In conclusion, the findings show that beyond the first 30 days of illness, substantial burden of health loss — spanning pulmonary and several extrapulmonary organ systems — is experienced by COVID-19 survivors. Our results inform the global discussion on the post-acute manifestations of COVID-19; the findings provide a roadmap to inform health system planning and development of care strategies aimed at reducing chronic and permanent health loss and optimizing wellness among COVID-19 survivors.

Methods

Setting:

Users of the US Veteran Health Administration (VHA) were selected from US Department of Veterans Affairs (VA) electronic health care databases. The VHA provides health care to discharged veterans of the US armed forces and operates the largest nationally integrated healthcare system in the United States, with 1,255 health care facilities, including 170 VA Medical Centers and 1,074 outpatient sites located across the United States. Veterans enrolled have access to the Department of Veterans Affairs comprehensive medical benefits package including inpatient hospital care; outpatient services; preventive, primary, and specialty care; prescriptions; mental healthcare; home healthcare; geriatric and extended care; medical equipment; and prosthetics. VA electronic health care databases are update daily.

Cohort:

The cohort was constructed from 5,808,018 participants who had encountered the VHA between January 01, 2019 and December 31, 2019. Within those alive on March 01, 2020 (N = 5,606,309), a COVID-19 group was selected as those with a COVID-19 positive test between March 01, 2020 and October 15, 2020 (n = 55,181). To examine post-acute outcome, we then selected from the COVID-19 group those alive at 30th day after their positive test (COVID-19 participants n = 51,376). To generate a comparison group that

had a similar length of follow up, we then conducted matching without replacement, matching each COVID-19 participant with 100 VHA users who did not have COVID-19 positive test during the period of time (control group n = 5,137,600). In the VHA users group we similarly selected those who were alive 30 days after the date of positive test in matched COVID-19 patient (control group n = 5,051,832) (Extended data Fig. 1a and 2a). For the matched participants in COVID-19 group and VHA users group, the date of positive test was considered date of enrollment and December 15, 2020 was the end of follow up.

To compare post-acute outcomes of hospitalized COVID-19 and hospitalized seasonal influenza participants, 10,246 COVID-19 participants who were admitted to a hospital within 30 days after or 5 days before their first positive test were selected from the 55,181 patients with positive COVID-19 test between March 01, 2020 and October 15, 2020. Similarly, 62,909 patients with their first positive seasonal influenza test between October 01, 2016 and February 29, 2020 who encountered the VHA at least once in the calendar year before the test were collected. Within them, 14,948 seasonal influenza patients were admitted to a hospital within 30 days after or 5 days before their first positive influenza test. The hospitalized cohort was further restricted to those alive at 30th day after hospital admission (COVID-19 n = 8738 and seasonal influenza n = 14,212). For 159 patients in both hospitalized COVID-19 and seasonal influenza group, only their COVID-19 hospitalizations were used in the analyses (Extended data Fig. 1b and 2b). In this cohort, participants were considered enrolled at the time of hospitalization. COVID-19 hospitalized patients were followed until December 15, 2020 and seasonal influenza hospitalized patients were followed until 290 days after hospital admission.

Data Sources:

Electronic health records from VA Corporate Data Warehouse (CDW) were used in this study[18–21]. The CDW Outpatient Encounters domains provided information related to outpatient encounters and Inpatient Encounters domains provided information between hospital admission and discharge[22]. The CDW Outpatient Pharmacy domain and CDW Bar Code Medication Administration domain were used to collect medication data and CDW Patient domain was used to collect demographic information. The CDW Laboratory Results domain was used to collect laboratory test information, and the COVID-19 Shared Data Resource was used to collect COVID-19 test and demographic information for COVID-19 patients. In addition, the Area Deprivation Index (ADI), a composite measure of income, education, employment, and housing was obtained from the University of Wisconsin[23].

Post-acute Health Resource Utilization And Death:

Outcomes which occurred after 30 days of cohort enrollment including death, incident admission to hospital, incident outpatient encounter and frequency of outpatient encounter were examined in both cohorts. Frequency of outpatient encounters was computed based on the number of days with outpatient encounter over days of follow up after 30 days and was reported as number of outpatient encounters per 30 days.

High Dimensional Post-acute Clinical Characteristics:

Diagnoses:

All ICD-10 diagnosis codes from cohort participants from days 30 after COVID-19 diagnosis until end of follow-up were used to define the post-acute diagnosis outcomes. More than 70,000 ICD-10 diagnosis codes were classified into 540 diagnostic categories based on the Clinical Classifications Software Refined (CCSR) version 2021.1, which is developed as part of the Healthcare Cost and Utilization Project sponsored by the Agency for Healthcare Research and Quality[24–26]. We only examined diagnostic categories that may be plausibly be considered post-acute sequalae of COVID-19 in the adult population. Diagnostic categories including external causes of morbidity, injury, poisoning and certain other consequences of external causes, congenital malformations, deformations and chromosomal abnormalities, certain conditions originating in the perinatal period or outcome from pregnancy, childbirth and the puerperium were not examined, yielding 379 diagnostic categories.

Medication Use:

Cohort participants' prescription records from day 30 after COVID-19 diagnosis until end of follow-up were used to define the post-acute medication use. 3425 medications were classified based on the US Department of Veterans Affairs Drug Classification system into 543 medication classes[27,28]. After removing items in the medication group of investigational agents or prosthetics, supplies and devices, in total 380 different medication outcomes were examined.

Laboratory Abnormalities:

In total 62 laboratory test abnormalities from 38 laboratory measurements from day 30 after COVID-19 diagnosis until end of follow-up were examined including absolute T cell count, alanine aminotransferase, aspartate aminotransferase, blood urea nitrogen, brain natriuretic peptide, C reactive protein, carbon dioxide, CD4/CD8 ratio, direct bilirubin, estimated glomerular filtration rate, ferritin, hematocrit, hemoglobin, hemoglobin A1c, high density lipoprotein cholesterol, high-sensitivity C-reactive protein, international normalized ratio, low density lipoprotein cholesterol, microalbumin/creatinine ratio, partial thromboplastin time, platelet count, pro B natriuretic peptide, prothrombin time, serum albumin, serum alkaline phosphatase, serum calcium, serum chloride, serum creatinine, serum phosphate, serum potassium, serum sodium, serum total protein, total bilirubin, total cholesterol, total white blood cell count, triglycerides, troponin I and troponin T were identified based on Logical Observation Identifiers Names and Codes (LOINC). Each laboratory test result was classified into abnormally high or abnormally low based on whether results were above the upper normal range or below the lower normal range, in the instance where for a given lab a high, or low, result might be clinically possible. The definition of the abnormality for each laboratory test is presented in supplementary information table 3 and 6.

Covariates:

Predefined covariates for analyses included demographics such as age, race (white, black, and other), and sex; proxies of healthcare utilization such as number of outpatient encounters, number of hospital admissions, number of outpatient prescriptions and number of outpatient eGFR measurements in the

year before enrollment. In addition, area deprivation index at patients' residency address as a summary measurement of socio-economic deprivation was included. To address for the potential non-linear association, all continuous variables were adjusted as restricted cubic spline functions.

To most optimally further adjust the models, we leveraged the multidimensionality of the VA's electronic health care databases to algorithmically identify covariates (potential confounders) spanning multiple domains (diagnoses, pharmacy records, laboratory tests) that showed evidence of difference in prevalence between the comparison groups[29]. In COVID-19 vs. VHA users' cohort, and separately in hospitalized COVID-19 vs. influenza cohort, high dimensional covariates were ascertained within one year before the date of enrollment. Within all diagnoses, medication classes and laboratory tests, we first selected variables that occurs in at least 10 patients in both groups. We then estimated the unadjusted relative risk of each variables with being in the COVID-19 or comparator group. The top 100 high dimensional variables with strongest association with group membership were used with predefined covariates in analyses.

Statistical Analyses:

Characteristics of the COVID-19 positive VHA users, VHA users without COVID-19, hospitalized COVID-19 participants and hospitalized seasonal influenza participants were described. The flowchart of the overall analytic approach is presented in extended data Fig. 5.

We estimated the risk of health resource utilization, death, and risk of each diagnosis, medication use and laboratory abnormality between COVID-19 and all VHA users, and separately, between those who had been hospitalized for COVID-19 and seasonal influenza. To estimate the risk of each incident outcome, we built a cohort of participants without a history of the outcome being examined (for example, risk of insulin use was estimated within a cohort of participants without history of insulin use in the year prior to cohort enrollment). For each cohort, propensity score based on predefined variables and high dimensional algorithmically selected variables was estimated. The propensity score was then used to compute the overlap weight, which is the probability of membership in the non-observed exposure group (one minus the propensity of in the observed group)[30,31].

Risks of health resource utilizations including hospital admission, outpatient encounter and death between COVID-19 and all VHA users, and between COVID-19 hospitalization and influenza hospitalization were estimated from COX survival model weighted by overlap weights, where death was considered as a competing risk in the evaluation of health resource utilizations. Frequency of outpatient encounter was modeled based on weighted linear regression. Hazard ratios for each of the outcomes including incident diagnoses, incident medication use, and incident laboratory abnormalities were estimated from cause specific hazard models weighted by overlap weights, where occurrence of death was considered as a competing risk. Event rates per 1000 participants at 6-months (180 days) of follow up in each group, and the excess burden based on the differences between two groups were estimated. Models were only built for outcomes occurring in at least ten participants from each group. Bonferroni

correction was applied in consideration of multiple hypotheses testing for high dimensional outcomes. A P-values of less than $6.57x10^{-5}$ was considered statistically significant. Results are additionally presented with a focus on common post-acute sequalae of COVID-19, where we selected those with hazard ratio greater than 1, P-values of less than $6.57x10^{-5}$ and excess burden more than 10 per 1000 COVID-19 patients at 6-months.

All analyses were done using SAS Enterprise Guide version 7.1 (SAS Institute, Cary, NC). Data visualizations were performed in R 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria). The study was approved by the Institutional Review Board of the Department of Veterans Affairs St. Louis Health Care System, St. Louis, MO.

Declarations

Data Availability

The data that support the findings of this study are available from the US Department of Veterans Affairs.

Code Availability

All SAS and R programing codes will be made available upon request.

Acknowledgements: This study used data from the VA COVID-19 Shared Data Resource. Miao Cai developed the data visualization and Andrew K. Gibson provided technical and editorial assistance.

Author Contributions: ZAA, YX, and BB contributed to the development of the study concept and design. YX and BB contributed to data acquisition. YX, BB, and ZAA contributed to data analysis and interpretation. YX and BB contributed to statistical analysis. ZAA and YX drafted the manuscript. YX, BB, and ZAA contributed to critical revision of the manuscript. ZAA provided administrative, technical, and material support, as well as supervision and mentorship. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved. All authors approved the final version of the report. The corresponding author attests that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted.

Funding: This research was funded by the United States Department of Veterans Affairs and the Institute for Public Health at Washington University in Saint Louis, Missouri, USA (for ZAA), and two American Society of Nephrology and KidneyCure Pre-doctoral fellowship awards (for BB and YX). The funders of this study had no role in study design; collection, analysis, and interpretation of data; writing the report; and the decision to submit the report for publication.

Competing Interest: The authors report no competing interests

Ethical approval: This research project was reviewed and approved by the Institutional Review Board of the Department of Veterans Affairs Saint Louis Health Care System.

Materials & Correspondence: Correspondence and request for materials should be addressed to ZAA.

Additional Information: Supplementary information is available for this paper.

References

1 Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *Jama*, doi:10.1001/jama.2020.1585 (2020).

2 Xie, Y., Bowe, B., Maddukuri, G. & Al-Aly, Z. Comparative evaluation of clinical manifestations and risk of death in patients admitted to hospital with covid-19 and seasonal influenza: cohort study. *Bmj*, doi:http://dx.doi.org/10.1136/bmj.m4677 (2021).

3 NICE. Covid-19 rapid guideline: managing the long-term effects of covid-19. . (2020).

4 NHS England. National guidance for post-covid syndrome assessment clinics. . (2020).

5 Yelin, D. *et al.* Long-term consequences of COVID-19: research needs. *Lancet Infect Dis* **20**, 1115-1117, doi:10.1016/S1473-3099(20)30701-5 (2020).

6 Meeting the challenge of long COVID. *Nat Med* **26**, 1803, doi:10.1038/s41591-020-01177-6 (2020).

7 Sivan, M. & Taylor, S. NICE guideline on long covid. *Bmj* **371**, m4938, doi:10.1136/bmj.m4938 (2020).

8 Long-term immunological health consequences of COVID-19. *British Society for Immunology* (2020).

9 Figueroa, J. D. *et al.* Distinguishing between direct and indirect consequences of covid-19. *Bmj* **369**, m2377, doi:10.1136/bmj.m2377 (2020).

10 Townsend, E. COVID-19 policies in the UK and consequences for mental health. *Lancet Psychiatry* **7**, 1014-1015, doi:10.1016/S2215-0366(20)30457-0 (2020).

11 Knipe, D., Evans, H., Marchant, A., Gunnell, D. & John, A. Mapping population mental health concerns related to COVID-19 and the consequences of physical distancing: a Google trends analysis. *Wellcome open research* **5**, 82, doi:10.12688/wellcomeopenres.15870.2 (2020).

12 Raker, E. J., Zacher, M. & Lowe, S. R. Lessons from Hurricane Katrina for predicting the indirect health consequences of the COVID-19 pandemic. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 12595-12597, doi:10.1073/pnas.2006706117 (2020).

13 Mahase, E. Covid-19: Mental health consequences of pandemic need urgent research, paper advises. *Bmj* **369**, m1515, doi:10.1136/bmj.m1515 (2020).

14 Del Rio, C., Collins, L. F. & Malani, P. Long-term Health Consequences of COVID-19. *Jama*, doi:10.1001/jama.2020.19719 (2020).

15 Greenhalgh, T., Knight, M., A'Court, C., Buxton, M. & Husain, L. Management of post-acute covid-19 in primary care. *Bmj* **370**, m3026, doi:10.1136/bmj.m3026 (2020).

16 Gorna, R. *et al.* Long COVID guidelines need to reflect lived experience. *Lancet*, doi:10.1016/S0140-6736(20)32705-7 (2020).

17 The, L. Facing up to long COVID. *Lancet* **396**, 1861, doi:10.1016/S0140-6736(20)32662-3 (2020).

18 Xie, Y. *et al.* Proton Pump Inhibitors and Risk of Incident CKD and Progression to ESRD. *Journal of the American Society of Nephrology : JASN* **27**, 3153-3163, doi:10.1681/ASN.2015121377 (2016).

19 Xie, Y. *et al.* Risk of death among users of Proton Pump Inhibitors: a longitudinal observational cohort study of United States veterans. *BMJ Open* **7**, e015735, doi:10.1136/bmjopen-2016-015735 (2017).

20 Xie, Y. *et al.* Long-term kidney outcomes among users of proton pump inhibitors without intervening acute kidney injury. *Kidney international* **91**, 1482-1494, doi:10.1016/j.kint.2016.12.021 (2017).

21 Xie, Y. *et al.* Higher blood urea nitrogen is associated with increased risk of incident diabetes mellitus. *Kidney international* **93**, 741-752, doi:10.1016/j.kint.2017.08.033 (2018).

22 Vincent, B. M., Wiitala, W. L., Burns, J. A., Iwashyna, T. J., Prescott, H. C. Using veterans affairs corporate data warehouse to identify 30-day hospital readmissions. *Health Services and Outcomes Research Methodology* **18**, 143-154 (2018).

23 Kind, A. J. & Buckingham, W. R. Making Neighborhood-Disadvantage Metrics Accessible—The Neighborhood Atlas. *New England Journal of Medicine* **378**, 2456-2458 (2018).

24 Wei, Y. *et al.* Short term exposure to fine particulate matter and hospital admission risks and costs in the Medicare population: time stratified, case crossover study. *Bmj* **367**, l6258, doi:10.1136/bmj.l6258 (2019).

25 Aubert, C. E. *et al.* Best Definitions of Multimorbidity to Identify Patients With High Health Care Resource Utilization. *Mayo Clin Proc Innov Qual Outcomes* **4**, 40-49, doi:10.1016/j.mayocpiqo.2019.09.002 (2020).

26 HCUP CCSR. Healthcare cost and utilization project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD.

27 Olvey, E. L., Clauschee, S. & Malone, D. C. Comparison of critical drug-drug interaction listings: the Department of Veterans Affairs medical system and standard reference compendia. *Clin Pharmacol Ther* **87**, 48-51, doi:10.1038/clpt.2009.198 (2010).

28 Greene, M., Steinman, M. A., McNicholl, I. R. & Valcour, V. Polypharmacy, drug-drug interactions, and potentially inappropriate medications in older adults with human immunodeficiency virus infection. *Journal of the American Geriatrics Society* **62**, 447-453, doi:10.1111/jgs.12695 (2014).

29 Xie Y *et al.* Comparative Effectiveness of SGLT2 Inhibitors, GLP1 Receptor Agonists, DPP4 Inhibitors and Sulfonylureas on Risk of Kidney Outcomes: emulation of a target trial using healthcare databases. *Diabetes care*, doi:https://doi.org/10.2337/dc20-1890.

30 Thomas, L. E., Li, F. & Pencina, M. J. Overlap Weighting: A Propensity Score Method That Mimics Attributes of a Randomized Clinical Trial. *Jama* **323**, 2417-2418, doi:10.1001/jama.2020.7819 (2020).

31 Li, F., Morgan, K. L. & Zaslavsky, A. M. Balancing Covariates via Propensity Score Weighting. *Journal of the American Statistical Association* **113**, 390-400, doi:10.1080/01621459.2016.1260466 (2018).

Tables

Table 1
Characteristics of (1) People with COVID-19 and users of the Veterans Health Administration (VHA), and (2) people hospitalized with COVID-19 and people hospitalized seasonal influenza.

		COVID-19 vs. VHA users		Hospitalized COVID-19 vs. Seasonal Influenza	
		COVID-19 N = 51,376	VHA users N = 5,051,832	Hospitalized COVID-19 N = 8738	Hospitalized seasonal influenza N = 14,053
Age (IQR)		62.31 (49.35, 72.12)	66.78 (52.02, 73.98)	69.42 (59.88, 75.24)	70.14 (62.98, 77.03)
Race (%)	White	31,976 (62.24)	3,873,680 (76.68)	4721 (54.03)	10,277 (73.13)
	Black	16,727 (32.56)	941,620 (18.64)	3444 (39.41)	3137 (22.32)
	Other	2675 (5.21)	236,532 (4.68)	573 (6.56)	639 (4.55)
Gender (%)	Male	45,440 (88.45)	4,572,449 (90.51)	8194 (93.77)	13,261 (94.36)
	Female	5936 (11.55)	479,383 (9.49)	544 (6.23)	792 (5.64)
Number of outpatient encounter (IQR) [a]		4 (2, 6)	2 (1, 4)	8 (5, 11)	9 (5, 13)
Number of hospital admission (IQR) [a]		0 (0, 0)	0 (0, 0)	0 (0, 1)	0 (0, 1)
Number of prescriptions received (IQR) [a]		8 (4, 14)	6 (3, 11)	13 (7, 21)	16 (9, 23)
Number of outpatient eGFR measurements (IQR) [a]		2 (1, 3)	1 (0, 2)	3 (2, 7)	5 (2, 10)
Area deprivation index (IQR)		54.17 (42.87, 62.54)	53.86 (41.89, 62.61)	53.64 (42.87. 61.28)	52.70 (40.08. 61.31)
Follow up days (IQR)		152 (111, 188)	153 (113, 191)	155 (112, 213)	290 (290, 290)

	COVID-19 vs. VHA users		Hospitalized COVID-19 vs. Seasonal Influenza	
Total Person-years (Sum)	21,562.88	2,151,682.25	3974.69	10765.38
a. Data collected within one year before the cohort enrollment				

Table 2

Risk of death and health resource utilization in people with COVID-19 compared to users of the Veterans Health Administration (VHA), and in people who had been hospitalized with COVID-19 compared to people who had been hospitalized with seasonal influenza.

Cohort	Outcomes [a]	Hazard ratio (95% confidence interval) [b]	Incident rate per 1000 at 6-months in COVID-19 group (95% confidence interval) [b]	Incident rate per 1000 at 6-months in comparison group (95% confidence interval) [b, c]	Excess burden per 1000 at 6-months (95% confidence interval) [b]
COVID-19 vs. VHA users	Death	2.17 (1.98, 2.38)	72.80 (66.71, 79.42)	7.38 (6.74, 8.07)	65.42 (64.72, 66.06)
	Admission	1.55 (1.47, 1.64)	119.64 (113.78, 125.79)	33.37 (31.66, 35.17)	86.28 (84.47, 87.99)
	Outpatient encounter	1.28 (1.26, 1.29)	984.15 (983.28, 984.98)	863.67 (860.12, 867.17)	120.48 (116.98, 124.03)
	Number of any outpatient encounter per 30 days		In COVID-19 group (95% confidence interval) [a]	In VHA users (95% confidence interval) [a]	Excess encounters (95% confidence interval) [a]
			4.06 (4.02, 4.10)	3.05 (3.04, 3.06)	1.01 (0.96, 1.05)
Hospitalized COVID-19 vs. Hospitalized Seasonal Influenza	Death	1.45 (1.30, 1.62)	84.19 (75.75, 84.19)	58.82 (52.86, 65.44)	25.37 (18.75, 31.33)
	Admission	0.97 (0.91, 1.03)	250.76 (237.64, 264.46)	257.76 (244.34, 271.77)	-7.00 (-21.01, 6.42)
	Outpatient encounter	1.16 (1.13, 1.20)	991.91 (990.75, 992.94)	984.02 (982.09, 985.79)	7.88 (6.73, 8.92)

Cohort	Outcomes [a]	Hazard ratio (95% confidence interval) [b]	Incident rate per 1000 at 6-months in COVID-19 group (95% confidence interval) [b]	Incident rate per 1000 at 6-months in comparison group (95% confidence interval) [b, c]	Excess burden per 1000 at 6-months (95% confidence interval) [b]
	Number of any outpatient encounter per 30 days		In COVID-19 group (95% confidence interval) [a]	In seasonal influenza (95% confidence interval) [a]	Excess encounters (95% confidence interval) [a]
			7.73 (7.55, 7.90)	5.76 (5.65, 5.87)	1.97 (1.77, 2.18)

a. For analyses of people with COVID-19 vs. VHA users, outcomes were ascertained from 30 days after COVID-19 diagnosis. For analyses of people who had been hospitalized with COVID-19 vs. people who had been hospitalized with seasonal influenza, outcomes were ascertained from 30 days after hospital admission

b. Results based on survival models adjusted through overlap weighting

c. For analyses of people with COVID-19 vs. VHA users, VHA users served as the referent category. For analyses of people hospitalized with COVID-19 vs. people hospitalized with seasonal influenza, people who had been hospitalized with seasonal influenza served as the referent category

Figures

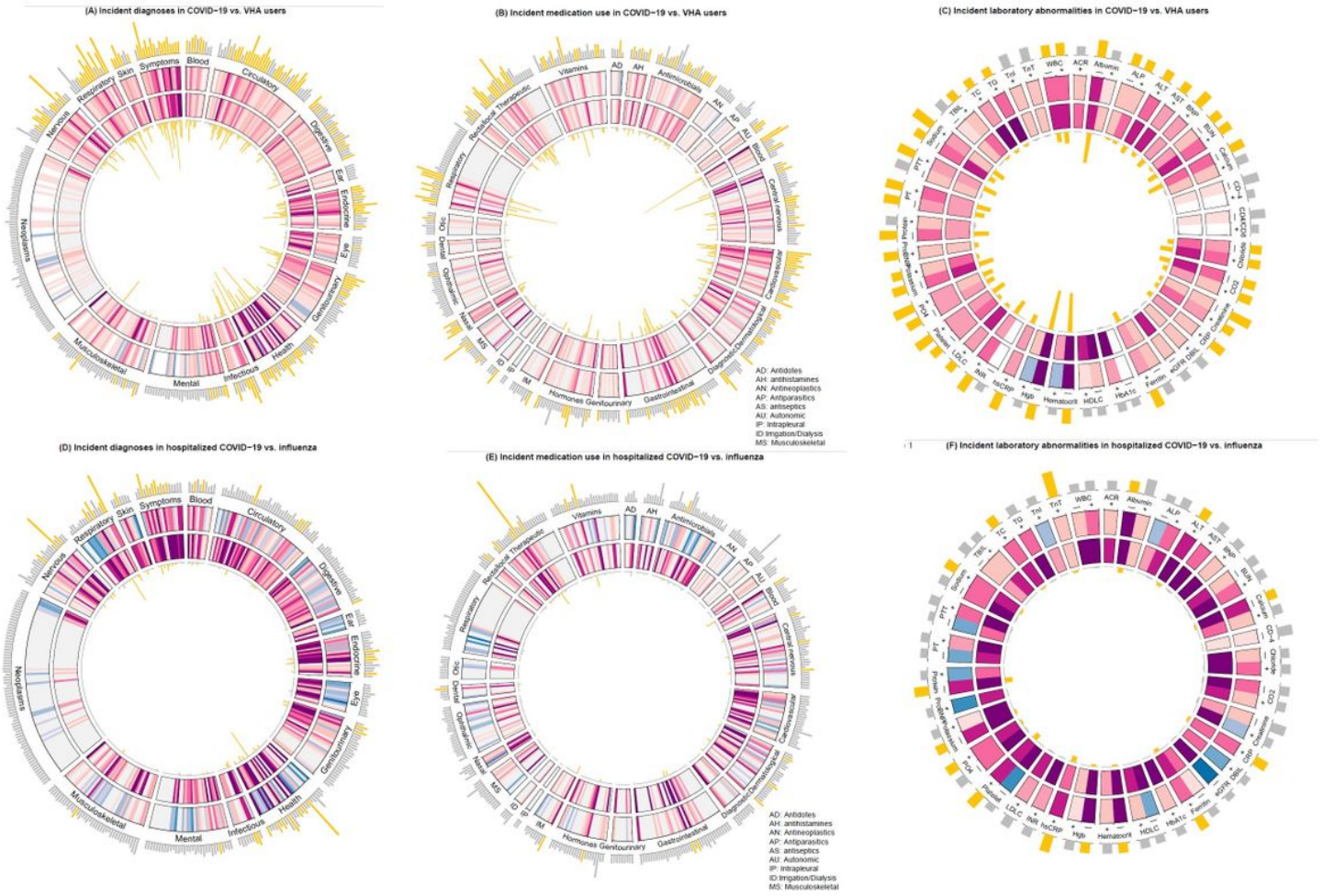

Figure 1

High dimensional identification of the incident post-acute sequalae of COVID-19. (A, D) Incident diagnoses, (B, E) Incident medication use and (C, F) Incident laboratory abnormalities. (A, B, C) COVID-19 vs. all VHA users. (D, E, F) hospitalized COVID-19 vs. hospitalized influenza. Post-acute sequalae were ascertained from 30 days after infection until end of follow-up. Beginning from the outside ring, the first ring represents hazard ratios for the post-acute sequalae of COVID-19. A higher bar indicates larger hazard ratio. Hazard ratios with point estimate larger than one and statistically significant were colored in yellow. The second ring represents excess burden per 1000 COVID-19 patients at 6-months. Color of the cell indicates value of the excess burden, where deeper shades of red indicate higher excess burden and deeper shades of blue indicate greater reduced burden. The third ring represents the baseline incident rate in the control group, where deeper shades of red indicate higher incident rate. The fourth ring represents negative log of the P value, where a higher bar indicates smaller P value and yellow bar indicate statistically significant.

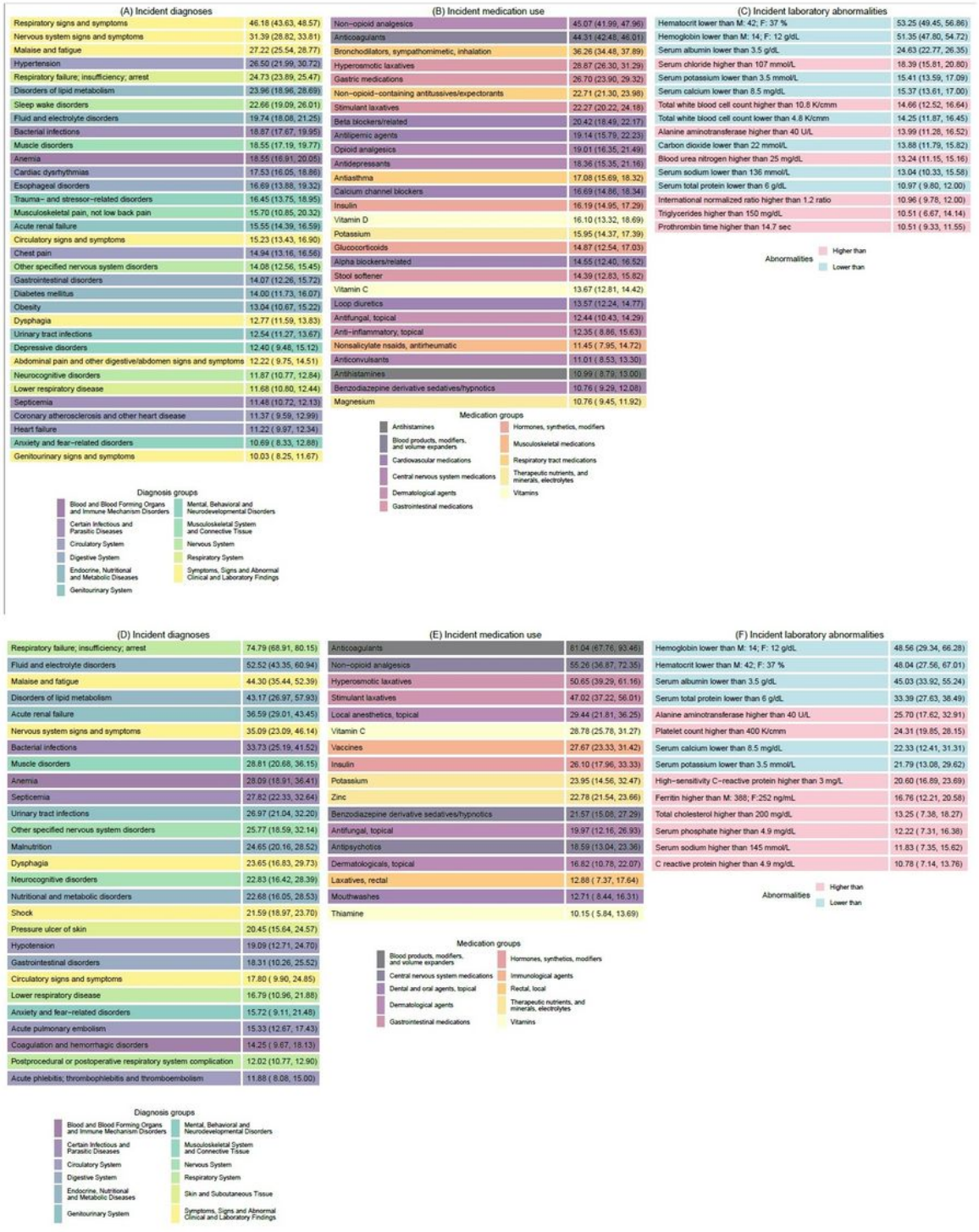

Figure 2

Common post-acute sequalae of COVID-19; post-acute sequalae due to COVID-19 with excess burden greater than 10 per 1000 persons at 6-month are presented. (A, D) Incident diagnoses, (B, E) Incident medication use and (C, F) Incident laboratory abnormalities. (A, B, C) COVID-19 vs. all VHA users. (D, E, F) hospitalized COVID-19 vs. hospitalized influenza. Post-acute sequalae were ascertained from 30 days after infection until end of follow-up. Sequalae were selected based on hazard ratio larger than one, P

value less than 6.57x10-5 and excess burden larger than 10 per 1000 COVID-19 patients at 6-months. Excess burdens per 1000 COVID-19 patients at 6-months are presented. Within each domain, outcomes are ranked based on excess burden from high to low. Diagnoses are colored based on diagnosis group, medications are colored based on medication class and laboratory abnormalities are colored based on higher or lower than normal range.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Combinedextendeddatafile.pdf
- SItable1COVIDvsallVHAdiagnosis.xlsx
- SItable2COVIDvsallVHAmedication.xlsx
- SItable3COVIDvsallVHAlab.xlsx
- SItable4HosCOVIDvsFLUdiagnosis.xlsx
- SItable5HosCOVIDvsFLUmedication.xlsx
- SItable6HosCOVIDvsFLUlab.xlsx