Outcomes of SARS-CoV-2 Reinfection

Ziyad Al-Aly (zalaly@gmail.com)
Washington University School of Medicine https://orcid.org/0000-0002-2600-0434
Benjamin Bowe
VA Saint Louis Health Care System
Yan Xie
VA Saint Louis Health Care System https://orcid.org/0000-0002-2457-9382

Article

Keywords:

Posted Date: June 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1749502/v1

Abstract

First infection with SARS-CoV-2 is associated with increased risk of acute and post-acute death and sequelae in the pulmonary and extrapulmonary organ systems. However, whether reinfection adds to the risk incurred after the first infection is not clear. Here we use the national health care databases of the US Department of Veterans Affairs to build a cohort of people with first infection (n = 257,427), reinfection (2 or more infections, n = 38,926), and a non-infected control group (n = 5,396,855) to estimate risks and 6-month burdens of all-cause mortality, hospitalization, and a set of pre-specified incident outcomes. We show that compared to people with first infection, reinfection contributes additional risks of all-cause mortality, hospitalization, and adverse health outcomes in the pulmonary and several extrapulmonary organ systems (cardiovascular disorders, coagulation and hematologic disorders, diabetes, fatigue, gastrointestinal disorders, kidney disorders, mental health disorders, musculoskeletal disorders, and neurologic disorders); the risks were evident in those who were unvaccinated, had 1 shot, or 2 or more shots prior to the second infection; the risks were most pronounced in the acute phase, but persisted in the post-acute phase of reinfection, and most were still evident at 6 months after reinfection. Compared to non-infected controls, assessment of the cumulative risks of repeated infection showed that the risk and burden increased in a graded fashion according to the number of infections. The constellation of findings show that reinfection adds non-trivial risks of all-cause mortality, hospitalization, and adverse health outcomes in the acute and post-acute phase of the reinfection. Reducing overall burden of death and disease due to SARS-CoV-2 will require strategies for reinfection prevention.

Introduction

A large body of evidence suggests that first infection with SARS-CoV-2 is associated with increased risk of acute and post-acute death and sequelae in the pulmonary and broad array of extrapulmonary organ systems[1–9]. However, many people around the globe are experiencing repeat SARS-CoV-2 infections (reinfections). Whether reinfection adds to the risk incurred after the first infection is not clear. And whether reinfection contributes to increased risk of post-acute sequelae is also not known. Addressing these questions has broad public health implications as it will inform whether strategies to prevent or reduce risk of repeat infections should be implemented.

In this work, we use the vast electronic health care databases of the US Department of Veterans Affairs to address the question of whether SARS-CoV-2 reinfection adds to the health risks associated with a first SARS-CoV-2 infection. We characterize the risks and 6-month burdens of a panel of pre-specified outcomes in a cohort of people who experience a SARS-CoV-2 reinfection compared to those with only the first infection, characterize the risks of acute and post-acute outcome in people with reinfection, and finally estimate the cumulative risks and 6-month burdens associated with one, two, and three or more infections compared to a non-infected control cohort.

Results

There were 257,427 cohort participants with first SARS-CoV-2 infection and 38,926 participants who had SARS-CoV-2 reinfection (two or more infections); 5,396,855 participants with no record of positive SARS-CoV-2 infection were in the control group. Among those with reinfection, 36,417 (12.29%) people had two infections, 2,263 (0.76%) people had three infections, and 246 (.08%) people had four or more infections. The median distribution of time between the first and second infection was 79 days (IQR: 48–119), and between the second and third was 65 (43–97). The demographic and health characteristics of those with first infection, reinfection (2 or more infections), and the non-infected control group are presented in Supplementary Table 1.

Sequelae of SARS-CoV-2 reinfection

To gain a better understanding of whether reinfection adds risk, we first conducted analyses to examine risks of all-cause mortality, hospitalization, and a set of pre-specified outcomes in people with reinfection compared to those with first infection.

We provide two measures of risk: (1) we estimated the adjusted hazard ratios of a set of incident pre-specified outcomes comparing people with reinfection versus first SARS-CoV-2 infection; (2) we estimated the adjusted excess burden of each outcome per 1,000 persons 6-months after SARS-CoV-2 reinfection on the basis of the difference between the estimated incidence rate in individuals with reinfection and first SARS-CoV-2 infection. Assessment of standardized mean differences of participant characteristics (from data domains including diagnoses, medications, and laboratory test results) after application of weighting showed they are well balanced in each analysis of incident outcomes (Supplementary Table 2, Supplementary Fig. 1).

Compared to those with first infection, those with reinfection exhibited an increased risk of all-cause mortality (Hazard Ratio (HR) 2.14; 95% confidence interval (CI): 1.97, 2.33) and excess burden of all-cause mortality estimated at 23.8 (95% CI: 18.9, 29.2) per 1000 persons at 6 months; all burden estimates represent excess burden and are given per 1,000 persons at 6 months (Fig. 1, Supplemental Table 3). People with a reinfection also had an increased risk of hospitalization (HR 2.98 (2.83, 3.12); burden of 95.47 (89.17, 102.03)), and having at least one sequela of SARS-CoV-2 infection (HR 1.82 (1.78, 1.88); burden of 196.2 (186.57, 205.87)) (Fig. 1, Supplemental Table 3).

Compared to those with first infection, those with reinfection exhibited increased risk of sequelae in the pulmonary (HR 2.49 (2.34, 2.65); burden 50.35 (45.64, 55.32)), and several extrapulmonary organ systems including cardiovascular disorders (HR 2.36 (2.23, 2.51); burden 49.83 (45.07, 54.86)), coagulation and hematologic disorders (HR 2.22 (2.05, 2.41); burden 25.93 (22.51, 29.6)), , fatigue (HR 2.4 (2.22, 2.58); burden 33.39 (29.41, 37.65)), gastrointestinal disorders (HR 1.69 (1.58, 1.8); burden 27.45 (23.24, 31.93)), kidney disorders (HR 1.70 (1.45, 2.46); burden 14.31 (10.69, 18.34)), mental health disorders (HR 1.97 (1.9, 2.04); burden 121.05 (113, 129.31)), diabetes (HR 1.62 (1.49, 1.76); burden 21.98 (17.56, 26.75)), musculoskeletal disorders (HR 1.29 (1.2, 1.38); burden 12.97 (9.04, 17.17)), and neurologic disorders (HR 1.39 (1.32, 1.46); burden 36.02 (30.16, 42.12)). Risks and excess burdens of reinfection are provided in Fig. 1 and Supplementary Table 3. Analyses examining whether the duration of

time from initial infection to reinfection might modify the association between reinfection and the risks of all-cause mortality, hospitalization, and at least one sequela suggested no interaction on the multiplicative scale (p-value for effect modification of 0.247, 0.301, and 0.706, respectively).

Analyses of prespecified subgroups based on vaccination status prior to the reinfection (no vaccination, 1 shot, or 2 or more shots) showed that reinfection (compared to first infection) was associated with a higher risk of all-cause mortality, hospitalization, at least one sequala, and sequelae in the different organ systems (Fig. 2, Supplementary Table 4) in people with no prior vaccination, one vaccine shot, or two or more vaccine shots.

Acute and post-acute sequelae of SARS-CoV-2 reinfection

We examined whether the risk of sequelae of SARS-CoV-2 reinfection was present in the acute and post-acute phase of the reinfection. We conducted analyses examining risk and burden starting from time of reinfection up to 180 days later in 30-day increments. Compared to those with first infection, those with reinfection exhibited increased risk and excess burden of all-cause mortality, hospitalization, and at least one sequela in the acute phase and the post-acute phase of the reinfection. The risks and excess burdens of all-cause mortality, hospitalization, and at least on sequela during the post-acute phase gradually attenuated over time but remained evident even six months after reinfection (Fig. 3, Supplementary Table 5). Examination of sequelae by organ system suggested an increased risk and excess burden in all organ systems during the acute phase (Fig. 4, Supplementary Table 5); the risks and burdens persisted in the post-acute phase of reinfection, and most were still evident at 6 months after reinfection.

Cumulative risk and burden of one, two, and three or more SARS-CoV-2 infections

To better understand the cumulative risks incurred by people with multiple infections, we estimated the risk and cumulative risk and burden of a set of pre-specified outcomes in those who during the 6-month period after the acute phase of the first infection did not have a reinfection (had only one infection), had two infections, and had three or more infections, compared to a non-infected control group. Cohort characteristics before and after weighting are provided in Supplementary Fig. 2 and Supplementary Tables 6-7. There was a graded association in that the risks of adverse health outcomes increased as the number of infections increased. Compared to the non-infected control group, those who only had one infection had an increased risk of at least one sequela (HR 1.35 (1.4, 1.36); burden 84.13 (82.03, 86.24)); the risk was higher in those who had two infections (HR 2.11 (2.07, 2.15); burden 234.58 (227.08, 241.92)), and highest in those with three or more infections (HR 3.00 (2.71, 3.31); burden 362.82 (326.37, 398.08)). In pairwise comparison of those with two infections vs first infection, those with two infections had an increased risk of at least one sequela (HR 1.57 (1.53, 1.60); burden 150.36 (142.95, 157.79)); in pairwise comparison of those with three or more infections vs those with only two infections, those with three or more infections had a higher risk of at least one sequela (HR 1.42 (1.28, 1.57); burden 128.33 (91.88, 164.31)). Results were consistent in examination of all-cause mortality, hospitalization, and sequelae by organ system (Fig. 5, Supplementary Tables 8-13).

Positive and negative outcome controls

We conducted a positive outcome control analysis to examine whether our approach reproduced prior established knowledge, testing whether the association of a SARS-CoV-2 infection (irrespective of reinfection) was associated with risk of fatigue (a well characterized, cardinal post-acute sequela of COVID-19, where a positive association would be expected based on prior evidence). Results showed compared to a non-infected control group, those with a SARS-CoV-2 infection exhibited an increased risk of fatigue (HR 2.02 (1.92-2.13)).

We then conducted a set of negative outcome control analyses to test for potential presence of spurious associations using the same data sources, cohort construction processes, covariate selections and definitions (including predefined and algorithmically selected high dimensional covariates), covariate balance methods, and result interpretations as those of our primary analysis. Results examining the risk of atopic dermatitis and neoplasms (negative outcome controls), where there was no prior biologic or epidemiologic evidence to suggest an association should be expected, showed no significant association (HR 1.00 (0.89, 1.13) and 1.02 (0.97, 1.06), respectively).

Discussion

In this study of 5,693,208 million people including 257,427 people with first infection, 38,926 people with reinfection, and 5,396,855 non-infected controls, we show that compared to people with first infection, people with reinfection exhibited increased risks of all-cause mortality, hospitalization, and several pre-specified outcomes. The risks were evident in subgroups including those who were unvaccinated, had 1 shot, or 2 or more shots prior to the second infection. The risks were most pronounced in the acute phase, but persisted in the post-acute phase of reinfection, and risks for most sequelae were still evident at 6 months. Compared to non-infected controls, assessment of the cumulative risks of repeated infection showed that the risk and burden of all-cause mortality and the prespecified health outcomes increased in a graded fashion according to the number of infections (that is risks were lowest in people with 1 infection, increased in people with 2 infections, and highest in people with 3 or more infections). Altogether, the findings show that reinfection adds non-trivial risks of all-cause mortality and adverse health outcomes in the acute and post-acute phase of the reinfection. The findings highlight the consequences of reinfection and emphasize the importance of preventing re-infection SARS-CoV-2. Surveillance systems should be de

Estimates suggest that more than half a billion people around the globe have been infected with SARS-CoV-2 at least once[10]. For the large and growing number of people who encountered a first infection, the question of whether a second infection carries additional risk is important. In this work, we show that reinfection adds risk of all-cause mortality and adverse health outcomes in both the acute phase and the post-acute phase of reinfection – suggesting that for people who already infected once, continued vigilance to reduce risk of reinfection may be important to reduce overall risk to one's health.

Given the likelihood that SARS-CoV-2 will remain a threat for years if not decades, we urgently need to develop public health measures that would be embraced by the public and could be sustainably implemented in the long-term to protect people from re-infection. Pharmaceutical interventions to lessen both the risk of reinfection and its adverse health consequences are also urgently needed.

Questions have been raised whether reinfection increases the risk of Long Covid – the umbrella term encompassing the post-acute sequelae of SARS-CoV-2 infection. Our results show that beyond the acute phase, reinfection with SARS-CoV-2 contributes substantial additional risks of all-cause mortality, hospitalization, and post-acute sequelae in the pulmonary and broad array of pulmonary organ systems.

The mechanisms underpinning the increased risks of death and adverse health outcomes in reinfection are not completely clear. Prior exposure to the virus may be expected to hypothetically reduce risk of reinfection and its severity; however, SARS-CoV-2 is mutating rapidly, and new variants are replacing older ones every few months. Evidence suggests that the reinfection risk is especially higher with the Omicron variant which was shown to have a marked ability to evade immunity from prior infection[11]. And any protection from infection also wanes over time[11]; protection from reinfection declined as the time increased since the last immunity-conferring event in people who had previously been infected with SARS-CoV-2 (regardless of vaccination status)[12]. Furthermore, impaired health as a consequence of the first infection might result in increased risk of adverse health consequences upon reinfection. Our results expand this evidence base and show that reinfection adds risk in both the acute and post-acute phase and that this was evident even among fully vaccinated people – suggesting that even combined natural immunity (from prior infection) and vaccine-induced immunity does not abrogate risk of adverse health effects following reinfection. The totality of evidence suggests that prevention strategies of reinfection might benefit people regardless of prior history of infection and vaccination status.

This study has several strengths. To our knowledge, this is the first study to characterize the health risks of reinfection. We used the national healthcare databases of the US Department of Veterans Affairs (the largest nationally integrated healthcare delivery system in the US) to undertake the analyses. We used advanced statistical methodologies and adjusted through weighting for a battery of predefined covariates selected based on prior knowledge and algorithmically selected covariates from high dimensional data domains including diagnoses, prescription records, and laboratory test results. Because the virus is mutating over time, and because different variants may have different effects on outcomes, we further adjusted our analyses for measures of space and time, and additionally for proportion of variant by region at time of infection. We evaluated both acute and post-acute outcomes of reinfection and examined risks according to vaccination status prior to reinfection. We evaluated the rigor of our approach by testing positive and negative outcome controls to determine whether our approach would produce results consistent with pre-test expectations.

The study also has several limitations. The cohorts of people with 1, 2, 3 or more infections included those that had a positive test for SARS-CoV-2 and did not include those who may have had an infection with SARS-CoV-2 but were not tested, if present in large numbers this may have resulted in

misclassification of exposure. Although the VA population is comprised of mostly men, it includes 10% women which across the groups in our study included 566,020 female participants. Although we balanced the exposure groups (through weighting using a battery of predefined and algorithmically selected covariates), we cannot completely rule out residual confounding. The COVID-19 pandemic is a highly dynamic global event that is still unfolding in real time; as various epidemiologic drivers of this pandemic change over time (including emergence of new variants, increase in vaccine uptake, and waning vaccine immunity), it is likely that the epidemiology of reinfection and its health consequences may also change over time.

In sum, in this study of 5,693,208 we provide evidence that reinfection contributes additional health risks beyond those incurred in the first infection including all-cause mortality, hospitalization, and sequelae in the pulmonary and broad array of extrapulmonary organ systems. The risks were evident in the acute and post-acute phase of reinfection. The evidence suggests that for people who already had a first infection, prevention of a second infection may protect from additional health risks. Prevention of infection and reinfection with SARS-CoV-2 should continue to be the goal of public health policy.

Declarations

Acknowledgements: This study used data from the VA COVID-19 Shared Data Resource. This research was funded by the United States Department of Veterans Affairs (ZAA) and two American Society of Nephrology and KidneyCure fellowship awards (BB and YX). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The contents do not represent the views of the US Department of Veterans Affairs or the US government.

Author Contributions: ZAA, BB, and YX contributed to the development of the study concept and design. ZAA, BB, and YX contributed to data analysis and interpretation. ZAA and BB drafted the manuscript. Critical revision of the manuscript was contributed to by ZAA, BB, and YX. ZAA provided administrative, technical, and material support. ZAA provided supervision and mentorship. ZAA is the guarantor of the work. 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.

Competing interests: The authors declare no conflict of interest.

References

1. Al-Aly, Z., Xie, Y. & Bowe, B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature **594**, 259–264 (2021).

2. Cohen, K., *et al.* Risk of persistent and new clinical sequelae among adults aged 65 years and older during the post-acute phase of SARS-CoV-2 infection: retrospective cohort study. BMJ **376**, e068414 (2022).
3. Bull-Otterson L, B.S., Saydah S, et al. Post–COVID Conditions Among Adult COVID-19 Survivors Aged 18–64 and ≥ 65 Years — United States, March 2020–November 2021. MMWR Morb Mortal Wkly Rep 2022;71:713–717. DOI: http://dx.doi.org/10.15585/mmwr.mm7121e1external.
4. Daugherty, S.E., *et al.* Risk of clinical sequelae after the acute phase of SARS-CoV-2 infection: retrospective cohort study. BMJ **373**, n1098 (2021).
5. Ayoubkhani, D., *et al.* Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. BMJ **372**, n693 (2021).
6. Carfì, A., Bernabei, R., Landi, F. & Gemelli Against, C.-P.-A.C.S.G. Persistent Symptoms in Patients After Acute COVID-19. JAMA **324**, 603–605 (2020).
7. Xie, Y., Bowe, B. & Al-Aly, Z. Burdens of post-acute sequelae of COVID-19 by severity of acute infection, demographics and health status. Nat Commun **12**, 6571 (2021).
8. Al-Aly, Z., Bowe, B. & Xie, Y. Long COVID after breakthrough SARS-CoV-2 infection. Nature Medicine (2022).
9. Taquet, M., *et al.* Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. PLoS Med **18**, e1003773 (2021).
10. Johns Hopkins Coronavirus Resource Center: COVID-19 map. (2022).
11. Pulliam, J.R.C., *et al.* Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa. Science **376**, eabn4947 (2022).
12. Goldberg, Y., *et al.* Protection and Waning of Natural and Hybrid Immunity to SARS-CoV-2. New England Journal of Medicine **386**, 2201–2212 (2022).

Methods

All participants that were eligible for this study were enrolled; no a-priori sample size analyses were conducted to guide enrollment. All analyses were observational, and investigators were aware of participant exposure and outcome status.

Setting

Participants were selected from the United States Veterans Health Administration (VHA) electronic health databases. The VHA delivers healthcare to discharged Veterans of the US armed forces in a network of nationally integrated healthcare systems including more than 1,415 healthcare facilities. Veterans enrolled for care in the VHA have access to extensive medical benefits such as inpatient and outpatient services, preventative, primary and specialty care, mental health services, geriatric care, long term and home healthcare, medications, and medical equation and prosthetics. VHA electronic health databases are updated daily.

Cohorts

A flowchart of cohort construction is provided in Supplementary Fig. 3. We first identified users of the VHA with at least one positive SARS-CoV-2 test between March 1, 2020 and September 4, 2021 (N=310,223), enrolling these participants at date of first positive test (set as T_0). Use of the VHA was defined as having record of use of outpatient or inpatient service, receipt of medication, or use of laboratory service with the VHA health care system in the two-years prior to enrollment. We excluded those who died during the first 30 days after the first positive SARS-CoV-2 test (N=296,353). We then further selected participants who experienced reinfection, defined as positive SARS-CoV-2 test more than 30 days after the first infection[13,14]. There were 38,926 participants who had a reinfection within the 6 months following 30 days after T_0, while 257,427 participants had only the first infection.

We then constructed a non-infected control group. We first identified 5,714,736 VHA users between March 1, 2020 and September 4, 2021 with no record of a positive SARS-CoV-2 test. We then randomly assigned a T_0 to each participant in the group on the basis of the distribution of the T_0 dates in those with at least one positive SARS-CoV-2 test between March 1, 2020 and September 4, 2021. We excluded those who died in the first 30 days after their T_0, yielding a control cohort of 5,396,855. All cohort participants were followed until April 4th, 2022.

Data sources

Participant data was obtained from the VHA Corporate Data Warehouse (CDW). The SPatient and Patient domains provided data on demographic characteristics. The Outpatient and Inpatient Encounters domains provided health characteristic information including details on date and place of encounter with the healthcare system, as well as diagnostic and procedural information. The Pharmacy and Bar Code Medication Administration domains provided medication records, while the Laboratory Results domain provided laboratory rest results for tests conducted in both inpatient and outpatient settings[7,15]. Information about SARS-CoV-2 tests and vaccinations were obtained from the COVID-19 Shared Data Resource (CSDR). The 2019 Area Deprivation Index (ADI) at each cohort participant residential address was used as a contextual measure of socioeconomic disadvantage[16]. Information from the US Center for Disease Control and Prevention (CDC) provided portion of SARS-CoV-2 variant by week in each Health and Human Services region.

Outcomes

Outcomes were pre-specified on the basis of on prior evidence[1-9,15,17-22]. Outcomes included all-cause mortality, hospitalization, having at least one sequela, as well as organ system disorders including cardiovascular disorders, coagulation and hematologic disorders, diabetes, fatigue, gastrointestinal disorders, kidney disorders, mental health disorders, musculoskeletal disorders, neurologic disorders, and pulmonary disorders. Organ system disorders were defined at date of first incident sequela in that system during follow-up. A list of individual sequelae by organ system are provided in Supplementary Table 14.

The outcome of “at least one sequela” was defined at the time of occurrence of first incident sequela among all individual sequelae. For a participant, for a given outcome, each individual sequala was included in the assessed outcome only when there was no record of that health condition in the two years prior to T_0. Participants were excluded from the analysis of an outcome if they had prior history of all the individual sequalae that contributed to the outcome being examined. Hospitalization was defined as first inpatient admittance during follow-up. In analyses of kidney disorders, participants with a prior history of end stage kidney disease (ESKD) were excluded, and follow-up was censored at time of ESKD (Supplementary Table 14).

Covariates

Covariates included a set of variables that were predefined based on prior knowledge[4-7,15,17,19-21,23-29] and a set of variables that were algorithmically selected. Predefined covariates included demographic information (age, race, and sex), contextual information (ADI), and measures of the healthcare utilization in the two years prior to T_0, which included the number of outpatient visits, inpatient visits, unique medication prescriptions, routine laboratory blood panels, and utilization of Medicare services, as well as a prior history of receiving an influenza vaccination. Smoking status was also included as a covariate. Characteristics of the participants health history included record of anxiety, cancer, cardiovascular disease, cerebrovascular disease, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression, type 2 diabetes mellitus, estimated glomerular filtration rate, immunocompromised status, peripheral artery disease, as well as systolic and diastolic blood pressure and body mass index (BMI). Immunocompromised status was defined according to CDC definitions by a history of organ transplantation, advanced kidney disease (an estimated glomerular filtration rate less than 15 $ml/min/1.73m^2$ or end stage renal disease), cancer, HIV, or conditions with prescriptions of more than 30 days use of corticosteroids or immunosuppressants including systemic lupus erythematosus and rheumatoid arthritis.

We also included a set of covariates related to the acute phase of the first infection: severity of the acute phase of the disease, defined in mutually exclusive groups of non-hospitalized, hospitalized, and admitted to the ICU during the acute phase, and whether the participant received a SARS-CoV-2 treatment (antivirals, antibodies, and steroids). We also included — as measures of spatiotemporal differences — the calendar week of enrollment and geographic region of receipt of care defined by Veterans Integrated Services Networks. We also adjusted for vaccination status, which was defined as receiving 0, 1, or 2 or more Janssen [Johnson & Johnson] (Ad26.COV2.S) vaccination, Pfizer-BioNTech (BNT162b2) or Moderna (mRNA-1273) vaccination shots. In consideration of the dynamicity of the pandemic, additional covariates included hospital system capacity (the total number of inpatient hospital beds), and inpatient bed occupancy rates (the percentage of hospital beds that were occupied), as well as a measure of the proportions of SARS-CoV-2 variants by Health and Human Services region[29]. These measures were ascertained for each participant in the week of cohort enrollment at the location of the health care system they received care at.

In addition to the predefined covariates, we leveraged the high dimensionality of VA electronic health records by employing a high dimensional variable selection algorithm to identify additional covariates that may potentially confound the examined associations[30]. We used the diagnostic classifications system from the Clinical Classifications Software Refined (CCSR) version 2021.1, available from the Healthcare Cost and Utilization Project sponsored by the Agency for Healthcare Research and Quality, to classify more than 70,000 ICD-10 diagnoses codes in the two years prior to T_0 for each participant into 540 diagnostic categories[31-33]. Using the VA national drug classification system, we also classified 3,425 different medications into 543 medication classes[34,35]. Finally, on the basis of Logical Observation Identifiers Names and Codes, we classified laboratory results from 38 different laboratory measurements into 62 laboratory test abnormalities, defined by being above or below the corresponding reference ranges. Of the high dimensional variables that occurred at least 100 times in participants in each group, we selected the top 100 variables with the highest relative risk for differences in group membership in first infection or reinfection.

Statistical Analysis

Mean (standard deviation) and frequency (percentage) of characteristics are reported for those with first SARS-CoV-2 infection, SARS-CoV-2 reinfection, and the non-infected control group, where appropriate. We provide information on the distribution of frequency of reinfections and time between infections.

All associations were estimated based on weighting approaches combined with survival analyses. We conducted a primary analysis to evaluate the risk and burden of reinfection in comparison to first infection (Supplementary Fig. 3), where reinfection was compared to those with first infection at the same time since the first infection. To achieve this, we constructed six sets of sub cohorts by 30-day time periods starting from 30 days after T_0 where within each period, participants were assigned to the reinfection or first infection group dependent on having a reinfection during that period. Those with a reinfection prior to the period were excluded. Participants with multiple reinfections were not censored at time of third plus infection. Participants, by period, were followed from time of reinfection (T_1) up to death, six months, or administrative censoring. To enhance comparisons, within each sub-cohort the distribution of time from the initial infection for the first infection group was randomly assigned on the basis of the distribution in those with reinfection.

For each sub cohort, logistic regressions were constructed to estimate the propensity score of group membership. A reference cohort of the overall infected cohort at T_0 was used as the target population. Inverse probability weighting was then used to balance of covariates. Differences in duration of follow-up were adjusted for. Cohorts across periods were pooled to estimate the average risk difference between those with and without a repeat infection using a weighted Cox survival model conditional on period. Standard errors were estimated by applying the robust sandwich variance estimator method. Covariate balance among all predefined and high dimensional variables were assessed for each group/period pair through the standardized mean difference (SMD), where a difference <0.1 was taken as evidence of balance. We estimated the incidence rate difference (referred to as excess burden) between groups per

1,000 participants at 6 months after the start of follow-up based on the difference in survival probability between the groups. These analyses were repeated by subgroups on the basis of the number of vaccination shots received (0, 1, or 2+) before reinfection. To test whether the risk on the multiplicative scale differed between the periods, a model with a linear interaction term between reinfection status and period was constructed, and the corresponding p-value is reported for the outcomes of all-cause mortality, hospitalization, and having at least one sequela.

To examine whether risks associated with a reinfection were present in the acute and post-acute phase of the reinfection, we conducted analyses to examine risks in 30-day time intervals starting at time of reinfection up to 180 days after reinfection. Hazard ratios and 30-day burdens were estimated independently for each 30-day time interval. During each 30-day interval outcomes were defined at time of first occurrence within this interval in those that did not have that outcome in the two years prior to the first infection.

We then examined the risk and cumulative burden of sequelae associated with first, two, and three or more infections versus a non-infected control (Supplementary Fig. 4). A third or more infection was defined as a positive test at least 30 days after the second infection. Number of infections and outcomes were assessed in the 180 days following T_0 + 30 days. Because participants with three or more infections must have not died during the follow-up period to have that third (or more) infection, we did not examine the outcome of all-cause mortality due to immortal time bias.

Positive and negative controls

We examined, as positive outcome controls, the risk of fatigue in those with a SARS-CoV-2 infection compared to the non-infected control as a means of testing whether our approach would reproduce established knowledge[4,5,19-21].

The application of negative outcome control may help detect both suspected and unsuspected sources of spurious biases. We, therefore, examined the difference in risks of atopic dermatitis and neoplasms between those with reinfection and the first infection– where no prior knowledge suggests an association should be expected. The testing of positive outcome control and negative outcome controls may lessen, though not eliminate, concerns about biases related to study design, covariate selection, analytic approach, outcome ascertainment, unmeasured confounding, and other potential sources of latent biases[36,37].

All analyses were two-sided. In all analyses, a 95% confidence interval that excluded unity was considered evidence of statistical significance. All analyses were conducted in SAS Enterprise Guide 8.2, and all figures were generated in R 4.0.4. This study was approved the VA St. Louis Health Care System Institutional Review Board (protocol number 1606333).

Data availability: The data that support the findings of this study are available from the US Department of Veterans Affairs. VA data are made freely available to researchers behind the VA firewall with an

approved VA study protocol. For more information, please visit https://www.virec.research.va.gov or contact the VA Information Resource Center (VIReC) at VIReC@va.gov

Code availability: The analytic code is available at https://github.com/BcBowe3

References

1. Al-Aly, Z., Xie, Y. & Bowe, B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* **594**, 259-264 (2021).

2. Cohen, K.*, et al.* Risk of persistent and new clinical sequelae among adults aged 65 years and older during the post-acute phase of SARS-CoV-2 infection: retrospective cohort study. *BMJ* **376**, e068414 (2022).

3. Bull-Otterson L, B.S., Saydah S, et al. Post–COVID Conditions Among Adult COVID-19 Survivors Aged 18–64 and ≥65 Years — United States, March 2020–November 2021. MMWR Morb Mortal Wkly Rep 2022;71:713–717. DOI: http://dx.doi.org/10.15585/mmwr.mm7121e1external.

4. Daugherty, S.E.*, et al.* Risk of clinical sequelae after the acute phase of SARS-CoV-2 infection: retrospective cohort study. *BMJ* **373**, n1098 (2021).

5. Ayoubkhani, D.*, et al.* Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. *BMJ* **372**, n693 (2021).

6. Carfi, A., Bernabei, R., Landi, F. & Gemelli Against, C.-P.-A.C.S.G. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* **324**, 603-605 (2020).

7. Xie, Y., Bowe, B. & Al-Aly, Z. Burdens of post-acute sequelae of COVID-19 by severity of acute infection, demographics and health status. *Nat Commun* **12**, 6571 (2021).

8. Al-Aly, Z., Bowe, B. & Xie, Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nature Medicine* (2022).

9. Taquet, M.*, et al.* Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *PLoS Med* **18**, e1003773 (2021).

13. Adrielle Dos Santos, L.*, et al.* Recurrent COVID-19 including evidence of reinfection and enhanced severity in thirty Brazilian healthcare workers. *J Infect* **82**, 399-406 (2021).

14. Michlmayr, D.a.A., Michael Asger and Meaidi, Marianna and Irshad, Irfatha and de Sousa, Luís Alves and Fonager, Jannik and Rasmussen, Morten and Gubbels, Sophie Madeleine and Rasmussen, Lasse Dam. SARS-CoV-2 Reinfections in Denmark Confirmed by Whole Genome Sequencing. *Available at SSRN: https://ssrn.com/abstract=4054457 or http://dx.doi.org/10.2139/ssrn.4054457* (2022).

15. Bowe, B., Xie, Y., Xu, E. & Al-Aly, Z. Kidney Outcomes in Long COVID. *Journal of the American Society of Nephrology*, ASN.2021060734 (2021).

16. Kind, A.J.H. & Buckingham, W.R. Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. *N Engl J Med* **378**, 2456-2458 (2018).

17. Xie, Y., Xu, E., Bowe, B. & Al-Aly, Z. Long-term Cardiovascular Outcomes of COVID-19. *Nature Medicine* (2022).

18. Xie, Y., Xu, E. & Al-Aly, Z. Risks of Mental Health Outcomes in People with Covid-19: cohort study. *BMJ* (2022).

19. Taquet, M.*, et al.* Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *PLOS Medicine* **18**, e1003773 (2021).

20. Davis, H.E.*, et al.* Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine* **38**, 101019 (2021).

21. Taquet, M., Geddes, J.R., Husain, M., Luciano, S. & Harrison, P.J. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* **8**, 416-427 (2021).

22. Xie, Y. & Al-Aly, Z. Risks and burdens of incident diabetes in long COVID: a cohort study. *The Lancet Diabetes & Endocrinology* **10**, 311-321 (2022).

23. Xie, Y., Xu, E. & Al-Aly, Z. Risks of mental health outcomes in people with covid-19: cohort study. *BMJ* **376**, e068993 (2022).

24. Yan Xie, Z.A.-A. Risks and burdens of incident diabetes in long COVID-19: a cohort study. *Lancet Diabetes Endocrinol* (2022).

25. Spudich, S. & Nath, A. Nervous system consequences of COVID-19. *Science* **375**, 267-269 (2022).

26. Cai, M., Bowe, B., Xie, Y. & Al-Aly, Z. Temporal trends of COVID-19 mortality and hospitalisation rates: an observational cohort study from the US Department of Veterans Affairs. *BMJ Open* **11**, e047369 (2021).

27. Nalbandian, A.*, et al.* Post-acute COVID-19 syndrome. *Nature Medicine* **27**, 601-615 (2021).

28. Daugherty, S.E.*, et al.* Risk of clinical sequelae after the acute phase of SARS-CoV-2 infection: retrospective cohort study. *BMJ* **373**, n1098 (2021).

29. Sharma, A., Oda, G. & Holodniy, M. COVID-19 Vaccine Breakthrough Infections in Veterans Health Administration. in *medRxiv* 2021.2009.2023.21263864 (2021).

30. Schneeweiss, S.*, et al.* High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* **20**, 512-522 (2009).

31. 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 (2019).

32. 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 (2020).

33. HCUP CCSR. Healthcare cost and utilization project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD. Vol. 2021.

34. 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 (2010).

35. 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. *J Am Geriatr Soc* **62**, 447-453 (2014).

36. Lipsitch, M., Tchetgen Tchetgen, E. & Cohen, T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* **21**, 383-388 (2010).

37. Shi, X., Miao, W. & Tchetgen, E.T. A Selective Review of Negative Control Methods in Epidemiology. *Current Epidemiology Reports* **7**, 190-202 (2020).

Figures

Figure 1

Risk and burden of sequelae in people with SARS-CoV-2 reinfection vs one infection. Risk and 6-month excess burden of all-cause mortality, hospitalization, at least one sequela, and sequelae by organ system are plotted. Incident outcomes were assessed from reinfection to end of follow-up. Results are in comparison of SARS-CoV-2 reinfection (n=38,926) to first SARS-CoV-2 infection (n=257,427). Adjusted hazard ratios (dots) and 95% confidence intervals (error bars) are presented, as are estimated excess burden (bars) and 95% confidence intervals (error bars). Burdens are presented per 1000 persons at 6 months of follow-up from time of reinfection.

Figure 2

Risk and burden of sequelae in people with SARS-CoV-2 reinfection vs one infection by vaccination status prior to second infection. Risk of all-cause mortality, hospitalization, at least one sequela, and sequelae by organ system are plotted. Incident outcomes were assessed from reinfection to end of follow-up. Results are in comparison of SARS-CoV-2 reinfection (n=38,926) to first SARS-CoV-2 infection (n=257,427). At time of comparison, there were 69.49%, 9.09%, and 21.42% with no, one, and two or more vaccinations, respectively, among those with reinfection. At time of comparison, there were 59.86%, 9.18%, and 30.96% with no, one, and two or more vaccinations, respectively, among the first reinfection group. Adjusted hazard ratios (dots) and 95% confidence intervals (error bars) are presented.

Figure 3

Risk and burden of all-cause mortality, hospitalization, and at least one sequela in the acute and post-acute phase of SARS-CoV-2 reinfection vs one infection. Risk and 6-month burden of all-cause mortality, hospitalization, and at least one sequela of SARS-CoV-2 reinfection vs one infection in 30-day intervals covering the acute and post-acute phase of reinfection. Incident outcomes were assessed from reinfection to end of follow-up. Results are in comparison of SARS-CoV-2 reinfection (n=38,926) to first SARS-CoV-2 infection (n=257,427) by time since reinfection. Adjusted hazard ratios (dots) and 95% confidence intervals (error bars) are presented for each 30-day period since time of reinfection, as are estimated excess burden (bars) and 95% confidence intervals (error bars). Burdens are presented per 1000 persons at 30 days of follow-up from time of reinfection.

Figure 4

Risk and burden of sequelae by organ system in the acute and post-acute phase of SARS-CoV-2 reinfection vs one infection. Risk and 6-month excess burden of sequelae by organ system of SARS-CoV-2 reinfection vs one infection in 30-day intervals covering the acute and post-acute phase of reinfection. Incident outcomes were assessed from reinfection to end of follow-up. Results are in comparison of SARS-CoV-2 reinfection (n=38,926) to first SARS-CoV-2 infection (n=257,427) by time since reinfection. Adjusted hazard ratios (dots) and 95% confidence intervals (error bars) are presented for each 30-day period since time of reinfection, as are estimated excess burden (bars) and 95% confidence intervals (error bars). Burdens are presented per 1000 persons at 30 days of follow-up from time of reinfection.

Figure 5

Cumulative risk and burden of sequelae in people with one, two, and three or more SARS-CoV-2 infections compared to a non-infected control. Risk and 6-month excess burden of all-cause mortality,

hospitalization, at least one sequela, and sequelae by organ system are plotted. Incident outcomes were assessed from 30 days after first positive SARS-CoV-2 test to end of follow-up. Results are in comparison of first SARS-CoV-2 infection (n=257,427), two SARS-CoV-2 infections (n=36,417), and three or more SARS-CoV-2 infection (n=2,509) to a non-infected control (n=5,396,855). Adjusted hazard ratios (dots) and 95% confidence intervals (error bars) are presented, as are estimated excess burden (bars) and 95% confidence intervals (error bars). Burdens are presented per 1000 persons at 6 months of follow-up.

Supplementary Files

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

- SupplementaryMaterial.docx
- SupplementaryTables.xlsx