medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

EFFECT OF PAXLOVID TREATMENT DURING ACUTE COVID-19 ON LONG COVID ONSET: AN EHR-BASED TARGET TRIAL EMULATION FROM THE N3C AND RECOVER CONSORTIA

Alexander Preiss*[a], Abhishek Bhatia*[b], Leyna V. Aragon[c], John M. Baratta[b], Monika Baskaran[b], Frank Blancero[d], M. Daniel Brannock[a], Robert F. Chew[a], Iván Díaz[e], Megan Fitzgerald[f], Elizabeth P. Kelly[b], Andrea Zhou[g], Thomas W. Carton[h], Christopher G. Chute[i], Melissa Haendel[b], Richard Moffitt[j], and Emily Pfaff[b], on behalf of the N3C Consortium and the RECOVER EHR Cohort

a. RTI International, Durham, NC, USA
b. University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
c. University of New Mexico, Albuquerque, NM, USA
d. RECOVER Community Representative
e. New York University Grossman School of Medicine, New York, NY, USA
f. Patient Led Research Collaborative
g. University of Virginia, Charlottesville, VA, USA
h. Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
i. Johns Hopkins University School of Medicine, Public Health, and Nursing, Baltimore, MD, USA
j. Emory University School of Medicine, Atlanta, GA, USA

*Authors contributed equally to this work

Corresponding Author: Alexander Preiss, apreiss@rti.org

N3C and RECOVER Consortia Representative: Emily Pfaff, epfaff@email.unc.edu

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Preventing and treating post-acute sequelae of SARS-CoV-2 infection (PASC), commonly known as Long COVID, has become a public health priority. In this study, we examined whether treatment with Paxlovid in the acute phase of COVID-19 helps prevent the onset of PASC. We used electronic health records from the National Covid Cohort Collaborative (N3C) to define a cohort of 426,352 patients who had COVID-19 since April 1, 2022, and were eligible for Paxlovid treatment due to risk for progression to severe COVID-19. We used the target trial emulation (TTE) framework to estimate the effect of Paxlovid treatment on PASC incidence. We estimated overall PASC incidence using a computable phenotype. We also measured the onset of novel cognitive, fatigue, and respiratory symptoms in the post-acute period. Paxlovid treatment did not have a significant effect on overall PASC incidence (relative risk [RR] = 0.98, 95% confidence interval [CI] 0.95-1.01). However, it had a protective effect on cognitive (RR = 0.90, 95% CI 0.84-0.96) and fatigue (RR = 0.95, 95% CI 0.91-0.98) symptom clusters, which suggests that the etiology of these symptoms may be more closely related to viral load than that of respiratory symptoms.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Post-acute sequelae of SARS-CoV-2 infection (PASC), commonly known as Long COVID, has become a public health priority. PASC affects people from all walks of life, and it is difficult to predict whether an individual will get PASC at the time of acute infection. Many people with PASC continue to feel the impacts of the disease years after infection. Mechanisms causing PASC remain largely unknown, and we have yet to identify a treatment that is consistently effective across the array of PASC manifestations. Therefore, developing effective PASC prevention strategies will be crucial to alleviating the long-term public health impact of COVID-19. There is an urgent need for research on this topic, including identifying novel interventions and assessing whether and how known interventions could help prevent PASC.

Nirmatrelvir with ritonavir (Paxlovid) was given an emergency use authorization (EUA) in the United States in December 2021 for the treatment of patients with mild-to-moderate COVID-19 who are at high risk for progression to severe COVID-19. Paxlovid has proven effective at preventing severe COVID-19, hospitalization, and death, with supporting evidence from clinical trials and real-world evidence, although a recent study found that Paxlovid was less effective at preventing hospitalization from SARS-CoV-2 Omicron subvariants compared to prior variants.[1–7]

In 2022, several teams published case reports where Paxlovid was used to treat PASC. Across three early reports, treatment was effective in five of six treated patients.[8–10] A larger 2023 report found mixed effects in 13 patients, suggesting that Paxlovid treatment "may have meaningful benefits for some people with Long COVID but not others".[11] In sum, this evidence motivated the RECOVER-VITAL trial to evaluate Paxlovid as a potential treatment for PASC.[12]

In addition to treating PASC, researchers have begun to explore whether Paxlovid treatment in the acute phase of COVID-19 infection could help prevent the onset of PASC. One plausible pathway could be reducing infection severity. Several studies have found that more severe acute infection or hospitalization is associated with a higher risk of PASC.[13-16]

Few studies have explored Paxlovid as a PASC preventative, and results are mixed. The largest study to date (281,793 individuals) used data from the US Department of Veterans Affairs (VA).[17] The VA study found that Paxlovid treatment during the acute phase of COVID-19 reduced the risk of a composite outcome of 13 post-acute sequelae, with a hazard ratio of 0.74.[17]

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

However, two smaller studies found that Paxlovid treatment was not associated with a reduced risk of PASC: a survey of 4,684 individuals from the Covid Citizen Science cohort and a survey of 500 individuals from Montefiore Medical Center.[18,19] Although these studies are much smaller than the VA study, they are more representative of the general population, and survey methods may capture symptoms that electronic health record (EHR) data do not. Also, because there is still no consensus definition of PASC, these studies use different outcome measures. In sum, the relationship between Paxlovid treatment and PASC onset remains uncertain.

At the time of writing, the PANORAMIC trial in the United Kingdom and the CanTreatCOVID trial in Canada are both recruiting for arms which will receive Paxlovid during acute COVID-19.[20,21] The PANORAMIC trial will focus on acute outcomes, but the CanTreatCOVID trial will include follow-up at 90 days and 36 weeks. CanTreatCOVID will provide valuable insight to the relationship between Paxlovid treatment and PASC onset, but as of March 5, 2024, the trial has recruited only 393 participants, with a planned completion date of January 2025.[22]

Through the National Institute of Health's National COVID Cohort Collaborative (N3C), and as part of the Researching COVID to Enhance Recovery (RECOVER) Initiative's EHR data team, we have the opportunity to study Paxlovid as a PASC preventative using a large, nationally sampled cohort and an up-to-date study period consisting mostly of Omicron BA and later subvariant infections.[23,24] This study adds to the evidence base while we await results from CanTreatCOVID and, hopefully, additional future trials. All analyses described here were performed within the secure N3C Data Enclave, which integrates EHR data for 21 million patients from over 230 data partners across the United States. N3C's methods for data acquisition, ingestion, and harmonization have been reported elsewhere.[23,25,26]

We used the target trial emulation (TTE) framework to estimate the effect of Paxlovid treatment in the acute phase of COVID-19 infection on the cumulative incidence of PASC among a cohort of patients eligible for Paxlovid treatment (i.e., with one or more risk factors for developing severe COVID-19).[27] We followed the two-step process for emulating target trials with observational data suggested by Hernán et al.[28] First, we articulated the causal question of interest in the form of a hypothetical trial protocol. Second, we emulated each component of this protocol using observational EHR data.

We measured overall PASC incidence using a machine learning-based computable phenotype model, which gathers data for each patient in overlapping 100-day periods that progress through

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

time, and issues a probability of PASC for each 100-day period.[29] The model was trained to classify whether patients have a U09.9 ("Post COVID-19 Condition") ICD-10 diagnosis code in each period, based on the patients' diagnoses during each period. Diagnoses surrounding known COVID-19 infections are blacked out.

To measure PASC at a more granular level, we also measured the novel onset of PASC symptoms in the cognitive, fatigue, and respiratory clusters proposed by the Global Burden of Disease (GBD) Study ("GBD symptom clusters" henceforth).[30] These clusters were the most frequently reported symptoms in a meta-analysis of Long COVID studies.[30] Their full definitions are cognitive problems (forgetfulness or difficulty concentrating, commonly referred to as brain fog); persistent fatigue with bodily pain (myalgia) or mood swings; and ongoing respiratory problems (shortness of breath and persistent cough as the main symptoms).[30]

We also conducted two sub-analyses: the first using a "VA-like cohort" designed to mirror the study period and demographics used in Xie et al. (2023)[17] and the second including COVID-19 vaccination status as an additional covariate, conducted in a subset of sites with high-quality vaccination data. Finally, we conducted several sensitivity analyses to test sensitivity to estimation methods, computable phenotype prediction threshold, COVID-19 index definition, and time period of outcome observation.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

Patient Characteristics

All results are reported in adherence with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[31] After inclusion and exclusion criteria, a total of 426,352 patients had a valid COVID-19 index date within the study period of April 1, 2022 to February 28, 2023, of whom 123,186 (28.89%) were treated with Paxlovid, and 24,469 (5.74%) had PASC (U09.9 diagnosis or computable phenotype prediction over 0.9 from 29 to 180 days after index). During the study period, 107 (0.09%) patients treated with Paxlovid and 622 (0.21%) untreated patients died. A total of 5,822 (1.37%) patients had a post-acute symptom in the cognitive symptom cluster, 15,239 (3.57%) patients had a post-acute symptom in the fatigue symptom cluster, and 24,833 (5.83%) had a post-acute symptom in the respiratory symptom cluster. Among patients with a PASC diagnosis or computable phenotype prediction, 8.57% had a post-acute symptom in the cognitive symptom cluster, 20.92% had a post-acute symptom in the fatigue symptom cluster, and 35.76% had a post-acute symptom in the respiratory symptom cluster. A co-occurrence matrix, showing the percentage of patients with each outcome who also had other outcomes, is shown in Extended Display Figure 1. After applying the eligibility criteria to the patient population and study sites, a total of 28 of 36 study sites were retained. The CONSORT flow diagram is shown in Figure 1. The characteristics of all patients during the study period are presented in Table 1, stratified by treatment group. Inverse probability of treatment weighting achieved balance across all covariates, as shown in Extended Display Figure 2. The target trial protocol and emulation approach are presented in Table 2.

Effect of Paxlovid Treatment on PASC Incidence

Overall, we found that Paxlovid treatment during acute COVID-19 did not have a significant effect on overall PASC onset as defined by our computable phenotype, but it did have a significant protective effect against cognitive and fatigue symptoms. Table 3 shows inverse probability of treatment-weighted Aalen-Johansen estimates of cumulative incidence for main analyses. Extended Display Table 1 shows estimated treatment effects across all analyses, including subanalyses and sensitivity analyses. Figure 2 shows corresponding risk ratios for all analyses.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

For overall PASC onset, measured by our PASC computable phenotype, adjusted cumulative incidence estimates were 6.92% (95% CI 6.74-7.09) for treated patients and 7.03% (95% CI 6.93-7.14) for untreated patients. The adjusted relative risk of PASC was 0.98 (95% CI 0.95-1.01). For the GBD symptom clusters, adjusted relative risk was 0.90 (95% CI 0.84-0.96) for the cognitive symptom cluster, 0.95 (95% CI 0.91-0.98) for the fatigue symptom cluster, and 1.00 (95% CI 0.97-1.03) for the respiratory symptom cluster. Figure 3 shows cumulative incidence functions.

Subanalyses

In the subanalysis using a VA-like cohort, we altered the cohort from the primary analysis to mirror the demographics of VA patients, which skew male and older compared to the general population. We also altered the study period to the one used in Xie et al, 2023, thereby including patients from the Omicron wave in early 2022, when Paxlovid was less widely available.[17] These changes remove potential sources of difference between our primary analysis and Xie et al, 2023, making the studies more directly comparable. The VA-like cohort included 64,233 male patients 65 years or older with a COVID-19 index between January 3, 2022, and December 31, 2022.[17] Of this cohort, 16,876 (26.27%) were treated with Paxlovid. Adjusted relative risk for PASC overall was 0.92 (95% CI 0.84-1.01). For the cognitive, fatigue, and respiratory GBD symptom clusters, adjusted relative risk was 0.78 (95% CI 0.66-0.90), 0.89 (95% CI 0.79-0.99), and 0.92 (95% CI 0.84-1.00), respectively. Cumulative incidence functions are shown in Extended Display Figure 4.

In the vaccination-aware subanalysis, we replicated our primary analysis while adjusting for COVID-19 vaccination status. We considered vaccination to be a plausible confounder of Paxlovid treatment and documented PASC, either through acute infection severity or propensity to seek care. However, vaccination status in N3C (like most EHRs) is subject to missingness. In this subanalysis, we used a subcohort of patients from sites with reliable vaccination data, which we identified, as in prior work, by comparing each site's data to public vaccination rates for its catchment area.[3,32] We categorized patients by their vaccination status prior to their COVID-19 index date, defined as having completed a full course of vaccination at least 14 days prior to index. Partially vaccinated patients and patients who became fully vaccinated fewer than 14 days prior to index were excluded from the analysis. The vaccination-aware cohort included 164,940 patients from 8 sites that met vaccination data quality criteria. Of this cohort, 59,257 (35.93%)

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

were treated with Paxlovid, and 8,824 (5.35%) had PASC according to our primary outcome measure. Adjusted relative risk was 0.96 (95% CI 0.92-1.01). The cumulative incidence function is shown in Extended Display Figure 5.

Our findings were not sensitive to the use of different computable phenotype prediction thresholds, or to the use of different PASC timing windows.

However, other sensitivity analyses produced different results. Treating only positive lab tests as index events, Paxlovid appeared to have an anti-protective effect on PASC (see Extended Display Figure 6). There is no plausible mechanism for this to be the case, and it is likely due to bias in the subset of COVID-19 patients who had documented lab tests in an era when home testing was common.

In our analysis including treatment with Paxlovid as a COVID-19 index event (i.e., including patients who received Paxlovid but did not have a COVID-19 diagnosis or positive lab result), we also found a significant, protective treatment effect (RR 0.97, 95% CI 0.94-0.99, see Extended Display Figure 7). In the absence of a COVID-19 diagnostic code (U07.1) or positive laboratory confirmed SARS-CoV-2 test to mark a COVID-19 index date, the additional subset of the patient population treated with Paxlovid here may still represent true COVID-19 cases, treated for suspected (but not laboratory-confirmed COVID-19), or with a recent personal history of COVID-19 (Z86.16).[33] Within the resulting patient cohort, the treatment effect of Paxlovid on PASC was statistically significant, but the effect size remained practically insignificant.

Using a doubly robust estimator and the hazard ratio (HR) of Paxlovid treatment as the estimand, we found a small but significant treatment effect (HR 0.95, 95% CI 0.92-0.98). This suggests that some residual confounding may remain after IPTW, however, the treatment effect remains practically insignificant.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

In this target trial emulation using the N3C database and a cohort of patients eligible for Paxlovid treatment (i.e., with one or more risk factors for severe COVID-19), we found that Paxlovid treatment during the acute phase of COVID-19 did not have a significant effect on overall PASC incidence as defined by our computable phenotype. However, we found that Paxlovid had a protective effect against the onset of novel cognitive and fatigue symptoms in the post-acute period and no effect on respiratory symptoms. The relationships among these effect sizes is explained by overall incidence rates. Computable phenotype PASC predictions and respiratory symptoms were far more common than cognitive and fatigue symptoms in this cohort. Although Paxlovid had a significant effect on cognitive and fatigue symptoms, few patients with PASC had these symptoms, and the effects were not large enough to contribute to a significant overall effect.

Differing effects by symptom cluster also suggest that Paxlovid may have more impact on the underlying causes of certain symptoms. In the literature, multiple PASC etiologies have been proposed. The chief hypotheses are that, relative to healthy convalescents, those with PASC may be experiencing (1) an aberrant autoimmune response triggered by the virus, (2) organ, tissue, or vascular dysfunction related to inflammatory cascades following infection, and/or (3) persistent viremia due to increased viral load or viral reservoirs. We do not yet know which symptoms are caused by which mechanisms. Paxlovid treatment decreases viral load, and thus could plausibly have more impact on symptoms arising from the third factor.[34] Our findings allow us to generate the hypothesis that cognitive symptoms (against which Paxlovid is most protective) may be caused by mechanisms that Paxlovid would affect (e.g., viral load).

The VA-like subanalysis, limited to a cohort of males at least 65 years old, found a much smaller treatment effect than Xie et al. (2023).[17] Despite our efforts to align outcome measures, cohort characteristics, and methodology, significant differences remain between our subanalysis and the VA study. Chief among them are remaining differences in our cohort and a true VA cohort. Veterans are more likely than demographically similar non-veterans to have been exposed to traumatic brain injury, post-traumatic stress disorder, biohazards, and other risk factors.[35–39] Through consistent access to the VA, the EHR for veterans may also be more complete.[40] Veterans may also differ from demographically similar non-veterans in their access to care.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

These factors may account for the large difference in PASC incidence and unadjusted and adjusted relative risk between our subanalysis and Xie et al. (2023).[17]

Although our study did not find that Paxlovid has a practically or statistically significant effect in preventing PASC overall, it shows promise at reducing the risk of certain post-acute symptoms. Although cognitive and fatigue symptoms were less common in this cohort, countless people suffer from these forms of PASC. The RECOVER-VITAL and CanTreatCOVID trials will provide strong evidence on whether Paxlovid is safe and effective in treating and preventing PASC, and will provide further insight to the differing effects by symptom observed in this study.[41] We hope future trials of Paxlovid treatment during acute COVID-19 will include post-acute endpoints. The target trial emulation framework employed here allows us to draw initial conclusions while we lack results from a randomized controlled trial. Notably, our results contrast with recent media coverage of Paxlovid's effect on PASC, with NBC News reporting in October 2023 that "A consensus has emerged among experts who study and treat long Covid: Paxlovid seems to reduce the risk of lingering symptoms among those eligible to take it."[42] Our findings bring this interpretation into question. Ultimately, although Paxlovid may help with certain symptoms, effective treatment and prevention of PASC remains elusive.

This study has several strengths that underscore the value of large-scale EHR repositories. We used a large, nationally sampled cohort from 28 sites across the United States, increasing generalizability and decreasing the potential for misclassification present in administrative or claims data.[43] The volume of data in the N3C database allowed for the aggressive inclusion/exclusion criteria necessary for TTE while preserving statistical power.[44,45] Our use of the TTE framework allowed us to account for confounding and estimate the causal effect of Paxlovid treatment using observational data.[46–49]

Our use of a PASC computable phenotype is also a strength. Although several institutions have proposed definitions of PASC, they disagree on the symptoms and timing that constitute the condition.[50–53] Varying definitions of PASC can lead to widely varying incidence estimates. Furthermore, measuring PASC as the novel onset of a specific set of symptoms can lead to false positives (symptoms with etiologies other than COVID-19) and false negatives (related symptoms not included in the definition). A machine learning-based computable phenotype may learn to avoid these errors. Furthermore, it does not require the selection of a principled set of symptoms, instead learning from all symptoms associated with PASC diagnoses. However,

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

computable phenotypes also have weaknesses: they are less interpretable and more complex than symptom-based definitions. PASC is also a heterogeneous outcome, so the use of symptom cluster outcomes is an important complement to the computable phenotype.

The study period makes our findings more relevant than prior studies of this topic, which have included cases from the initial Omicron wave, when Paxlovid was less available and disease dynamics were markedly different. Finally, our subanalyses shed further light on potential demographic and cohort effects.

This study also has limitations. Because EHR data do not include information on adherence, we can only measure whether a patient was prescribed Paxlovid. However, this is adequate for estimating the intention-to-treat effect. Also, our inclusion criteria of Paxlovid treatment within five days of COVID-19 index differs from the indication of treatment within five days of symptom onset. However, we note that within our cohort, 96% of treated patients were treated within one day of COVID-19 index.

This study's eligibility criteria include eligibility for on-label Paxlovid treatment (i.e., at risk for developing severe COVID-19 due to the presence of one or more risk factors). Therefore, results can only be generalized to a high-risk population. Ideally, a clinical trial on Paxlovid as a PASC preventative would also assess treatment among lower-risk populations. We chose not to emulate such a trial because it would complicate the study design and make exchangeability harder to establish due to confounding by indication. We note that the CanTreatCOVID trial also focuses only on high-risk patients. The effect of Paxlovid treatment on PASC onset among lower-risk patients is an important area for future research.

Several variables in this study are subject to measurement error. Many COVID-19 infections during this period were not documented due to the prevalence of home testing. Paxlovid prescriptions from providers outside N3C data partner systems may not be documented. The PASC computable phenotype may also misclassify patients.[29] For this reason, the confidence intervals around computable phenotype-based incidence estimates are likely too narrow. Vaccination status is poorly documented in most EHRs, which precluded its use as a covariate. However, our vaccination-aware sensitivity analysis, conducted in a subset of sites with high-quality vaccination data, found similar results. The fact that our findings were not sensitive to the inclusion of vaccination status as a covariate suggests that vaccination did not cause substantial unmeasured confounding in the primary analysis.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Finally, this study is subject to limitations common to EHR-based studies. EHRs are susceptible to missing data, and our estimates may be biased if missingness was related to unobserved confounding.[54–56] This study is also subject to the assumptions of all causal inference studies, in particular, that there is no unmeasured confounding. One potential unmeasured confounder is acute COVID-19 severity prior to diagnosis. Sicker patients may be more likely to seek Paxlovid and develop PASC. The EHR contains no reliable measure of this construct, but we control for pre-diagnosis comorbidities, which have been shown to correlate so strongly with COVID-19 severity that they can be considered proxies, thus mitigating the potential unmeasured confounding from this source.[57–59]

In summary, there is overwhelming evidence that Paxlovid is effective in preventing hospitalization and death due to acute COVID-19, and as such is a critical treatment option to improve COVID-19 outcomes. We used a target trial emulation framework to determine whether Paxlovid might also be effective in preventing PASC. Among patients with COVID-19 in our study period who were eligible for Paxlovid treatment, the cumulative incidence of PASC within a 180-day follow-up period was not significantly lower in patients treated with Paxlovid. However, the cumulative incidence of post-acute cognitive and fatigue symptoms was significantly lower. Based on these findings, we see Paxlovid as unlikely to become a definitive solution for PASC prevention. Cognitive and fatigue symptoms were relatively rare in this cohort, and its effect on these symptoms was not enough to move the needle on overall PASC incidence. Nevertheless, any protective effect is worth further exploration. Future research will dig deeper into potential heterogeneous treatment effects across PASC subphenotypes.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ACKNOWLEDGEMENTS

This study is part of the NIH Researching COVID to Enhance Recovery (RECOVER) Initiative, which seeks to understand, treat, and prevent the post-acute sequelae of SARS-CoV-2 infection (PASC). For more information on RECOVER, visit https://recovercovid.org/. This research was funded by the National Institutes of Health (NIH) Agreement OTA OT2HL161847 as part of the Researching COVID to Enhance Recovery (RECOVER) research program.

The analyses described in this manuscript were conducted with data or tools accessed through the NCATS N3C Data Enclave https://covid.cd2h.org and N3C Attribution & Publication Policy v 1.2-2020-08-25b supported by NCATS U24 TR002306, Axle Informatics Subcontract: NCATS-P00438-B, and by the RECOVER Initiative (OT2HL161847–01). The N3C Publication committee confirmed that this manuscript msid:1733.497 is in accordance with N3C data use and attribution policies; however, this content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the RECOVER or N3C programs. This research was possible because of the patients whose information is included within the data and the organizations (https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories) and scientists who have contributed to the on-going development of this community resource [https://doi.org/10.1093/jamia/ocaa196]. We would also like to thank the National Community Engagement Group (NCEG), all patient, caregiver and community Representatives, and all the participants enrolled in the RECOVER Initiative.

We also acknowledge the following institutions whose data is released or pending:

Available: Advocate Health Care Network — UL1TR002389: The Institute for Translational Medicine (ITM) • Aurora Health Care Inc — UL1TR002373: Wisconsin Network For Health Research • Boston University Medical Campus — UL1TR001430: Boston University Clinical and Translational Science Institute • Brown University — U54GM115677: Advance Clinical Translational Research (Advance-CTR) • Carilion Clinic — UL1TR003015: iTHRIV Integrated Translational health Research Institute of Virginia • Case Western Reserve University — UL1TR002548: The Clinical & Translational Science Collaborative of Cleveland (CTSC) • Charleston Area Medical Center — U54GM104942: West Virginia Clinical and Translational Science Institute (WVCTSI) • Children's Hospital Colorado — UL1TR002535: Colorado Clinical and Translational Sciences Institute • Columbia University Irving Medical Center — UL1TR001873: Irving Institute for Clinical and Translational Research • Dartmouth College — None (Voluntary) Duke University — UL1TR002553: Duke Clinical and Translational Science Institute • George Washington Children's Research Institute — UL1TR001876: Clinical and Translational Science Institute at Children's National (CTSA-CN) • George Washington University — UL1TR001876: Clinical and Translational Science Institute at Children's National (CTSA-CN) • Harvard Medical School — UL1TR002541: Harvard Catalyst • Indiana University School of Medicine — UL1TR002529: Indiana Clinical and Translational Science Institute •

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Johns Hopkins University — UL1TR003098: Johns Hopkins Institute for Clinical and Translational Research • Louisiana Public Health Institute — None (Voluntary) • Loyola Medicine — Loyola University Medical Center • Loyola University Medical Center — UL1TR002389: The Institute for Translational Medicine (ITM) • Maine Medical Center — U54GM115516: Northern New England Clinical & Translational Research (NNE-CTR) Network • Mary Hitchcock Memorial Hospital & Dartmouth Hitchcock Clinic — None (Voluntary) • Massachusetts General Brigham — UL1TR002541: Harvard Catalyst • Mayo Clinic Rochester — UL1TR002377: Mayo Clinic Center for Clinical and Translational Science (CCaTS) • Medical University of South Carolina — UL1TR001450: South Carolina Clinical & Translational Research Institute (SCTR) • MITRE Corporation — None (Voluntary) • Montefiore Medical Center — UL1TR002556: Institute for Clinical and Translational Research at Einstein and Montefiore • Nemours — U54GM104941: Delaware CTR ACCEL Program • NorthShore University HealthSystem — UL1TR002389: The Institute for Translational Medicine (ITM) • Northwestern University at Chicago — UL1TR001422: Northwestern University Clinical and Translational Science Institute (NUCATS) • OCHIN — INV-018455: Bill and Melinda Gates Foundation grant to Sage Bionetworks • Oregon Health & Science University — UL1TR002369: Oregon Clinical and Translational Research Institute • Penn State Health Milton S. Hershey Medical Center — UL1TR002014: Penn State Clinical and Translational Science Institute • Rush University Medical Center — UL1TR002389: The Institute for Translational Medicine (ITM) • Rutgers, The State University of New Jersey — UL1TR003017: New Jersey Alliance for Clinical and Translational Science • Stony Brook University — U24TR002306 • The Alliance at the University of Puerto Rico, Medical Sciences Campus — U54GM133807: Hispanic Alliance for Clinical and Translational Research (The Alliance) • The Ohio State University — UL1TR002733: Center for Clinical and Translational Science • The State University of New York at Buffalo — UL1TR001412: Clinical and Translational Science Institute • The University of Chicago — UL1TR002389: The Institute for Translational Medicine (ITM) • The University of Iowa — UL1TR002537: Institute for Clinical and Translational Science • The University of Miami Leonard M. Miller School of Medicine — UL1TR002736: University of Miami Clinical and Translational Science Institute • The University of Michigan at Ann Arbor — UL1TR002240: Michigan Institute for Clinical and Health Research • The University of Texas Health Science Center at Houston — UL1TR003167: Center for Clinical and Translational Sciences (CCTS) • The University of Texas Medical Branch at Galveston — UL1TR001439: The Institute for Translational Sciences • The University of Utah — UL1TR002538: Uhealth Center for Clinical and Translational Science • Tufts Medical Center — UL1TR002544: Tufts Clinical and Translational Science Institute • Tulane University — UL1TR003096: Center for Clinical and Translational Science • The Queens Medical Center — None (Voluntary) • University Medical Center New Orleans — U54GM104940: Louisiana Clinical and Translational Science (LA CaTS) Center • University of

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Alabama at Birmingham — UL1TR003096: Center for Clinical and Translational Science • University of Arkansas for Medical Sciences — UL1TR003107: UAMS Translational Research Institute • University of Cincinnati — UL1TR001425: Center for Clinical and Translational Science and Training • University of Colorado Denver, Anschutz Medical Campus — UL1TR002535: Colorado Clinical and Translational Sciences Institute • University of Illinois at Chicago — UL1TR002003: UIC Center for Clinical and Translational Science • University of Kansas Medical Center — UL1TR002366: Frontiers: University of Kansas Clinical and Translational Science Institute • University of Kentucky — UL1TR001998: UK Center for Clinical and Translational Science • University of Massachusetts Medical School Worcester — UL1TR001453: The UMass Center for Clinical and Translational Science (UMCCTS) • University Medical Center of Southern Nevada — None (voluntary) • University of Minnesota — UL1TR002494: Clinical and Translational Science Institute • University of Mississippi Medical Center — U54GM115428: Mississippi Center for Clinical and Translational Research (CCTR) • University of Nebraska Medical Center — U54GM115458: Great Plains IDeA-Clinical & Translational Research • University of North Carolina at Chapel Hill — UL1TR002489: North Carolina Translational and Clinical Science Institute • University of Oklahoma Health Sciences Center — U54GM104938: Oklahoma Clinical and Translational Science Institute (OCTSI) • University of Pittsburgh — UL1TR001857: The Clinical and Translational Science Institute (CTSI) • University of Pennsylvania — UL1TR001878: Institute for Translational Medicine and Therapeutics • University of Rochester — UL1TR002001: UR Clinical & Translational Science Institute • University of Southern California — UL1TR001855: The Southern California Clinical and Translational Science Institute (SC CTSI) • University of Vermont — U54GM115516: Northern New England Clinical & Translational Research (NNE-CTR) Network • University of Virginia — UL1TR003015: iTHRIV Integrated Translational health Research Institute of Virginia • University of Washington — UL1TR002319: Institute of Translational Health Sciences • University of Wisconsin-Madison — UL1TR002373: UW Institute for Clinical and Translational Research • Vanderbilt University Medical Center — UL1TR002243: Vanderbilt Institute for Clinical and Translational Research • Virginia Commonwealth University — UL1TR002649: C. Kenneth and Dianne Wright Center for Clinical and Translational Research • Wake Forest University Health Sciences — UL1TR001420: Wake Forest Clinical and Translational Science Institute • Washington University in St. Louis — UL1TR002345: Institute of Clinical and Translational Sciences • Weill Medical College of Cornell University — UL1TR002384: Weill Cornell Medicine Clinical and Translational Science Center • West Virginia University — U54GM104942: West Virginia Clinical and Translational Science Institute (WVCTSI) Submitted: Icahn School of Medicine at Mount Sinai — UL1TR001433: ConduITS Institute for Translational Sciences • The University of Texas Health Science Center at Tyler — UL1TR003167: Center for Clinical and Translational Sciences (CCTS) • University of California, Davis — UL1TR001860: UCDavis

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Health Clinical and Translational Science Center • University of California, Irvine — UL1TR001414: The UC Irvine Institute for Clinical and Translational Science (ICTS) • University of California, Los Angeles — UL1TR001881: UCLA Clinical Translational Science Institute • University of California, San Diego — UL1TR001442: Altman Clinical and Translational Research Institute • University of California, San Francisco — UL1TR001872: UCSF Clinical and Translational Science Institute Pending: Arkansas Children's Hospital — UL1TR003107: UAMS Translational Research Institute • Baylor College of Medicine — None (Voluntary) • Children's Hospital of Philadelphia — UL1TR001878: Institute for Translational Medicine and Therapeutics • Cincinnati Children's Hospital Medical Center — UL1TR001425: Center for Clinical and Translational Science and Training • Emory University — UL1TR002378: Georgia Clinical and Translational Science Alliance • HonorHealth — None (Voluntary) • Loyola University Chicago — UL1TR002389: The Institute for Translational Medicine (ITM) • Medical College of Wisconsin — UL1TR001436: Clinical and Translational Science Institute of Southeast Wisconsin • MedStar Health Research Institute — None (Voluntary) • Georgetown University — UL1TR001409: The Georgetown-Howard Universities Center for Clinical and Translational Science (GHUCCTS) • MetroHealth — None (Voluntary) • Montana State University — U54GM115371: American Indian/Alaska Native CTR • NYU Langone Medical Center — UL1TR001445: Langone Health's Clinical and Translational Science Institute • Ochsner Medical Center — U54GM104940: Louisiana Clinical and Translational Science (LA CaTS) Center • Regenstrief Institute — UL1TR002529: Indiana Clinical and Translational Science Institute • Sanford Research — None (Voluntary) • Stanford University — UL1TR003142: Spectrum: The Stanford Center for Clinical and Translational Research and Education • The Rockefeller University — UL1TR001866: Center for Clinical and Translational Science • The Scripps Research Institute — UL1TR002550: Scripps Research Translational Institute • University of Florida — UL1TR001427: UF Clinical and Translational Science Institute • University of New Mexico Health Sciences Center — UL1TR001449: University of New Mexico Clinical and Translational Science Center • University of Texas Health Science Center at San Antonio — UL1TR002645: Institute for Integration of Medicine and Science • Yale New Haven Hospital — UL1TR001863: Yale Center for Clinical Investigation

AUTHOR CONTRIBUTIONS

Authorship was determined according to ICMJE recommendations.

Alexander Preiss and Abhishek Bhatia contributed to study design, methodology, analysis, writing, and manuscript review.

Leyna Aragon contributed to study design and writing.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

John M Baratta contributed to study design and manuscript review.

Monika Baskaran contributed to analysis.

Frank Blancero contributed to study design and writing.

M Daniel Brannock contributed to study design, methodology, writing, and manuscript review.

Robert F Chew and Ivan Diaz contributed to study design, methodology, and manuscript review.

Elizabeth P Kelly contributed to project management.

Andrea G Zhou contributed to analysis.

Thomas W Carton, Chistopher G Chute, and Melissa Haendel contributed to study design and manuscript review.

Richard Moffitt contributed to study design, methodology, and manuscript review.

Emily Pfaff contributed to study design, methodology, writing, and manuscript review.

COMPETING INTERESTS STATEMENT

This group of authors has no relevant disclosures to report.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Dryden-Peterson, S. *et al.* Nirmatrelvir Plus Ritonavir for Early COVID-19 in a Large U.S. Health System : A Population-Based Cohort Study. *Ann. Intern. Med.* **176**, 77–84 (2023).
2. Hammond, J. *et al.* Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *N. Engl. J. Med.* **386**, 1397–1408 (2022).
3. Bhatia, A. *et al.* Effect of Nirmatrelvir/Ritonavir (Paxlovid) on Hospitalization among Adults with COVID-19: an EHR-based Target Trial Emulation from N3C. *medRxiv* (2023) doi:10.1101/2023.05.03.23289084.
4. Najjar-Debbiny, R. *et al.* Effectiveness of Paxlovid in Reducing Severe Coronavirus Disease 2019 and Mortality in High-Risk Patients. *Clin. Infect. Dis.* **76**, e342–e349 (2023).
5. Malden DE, Hong V, Lewin BJ, et al. *Hospitalization and Emergency Department Encounters for COVID-19 After Paxlovid Treatment — California*. vol. 71 830–833 http://dx.doi.org/10.15585/mmwr.mm7125 (2022).
6. Yip, T. C.-F. *et al.* Impact of the Use of Oral Antiviral Agents on the Risk of Hospitalization in Community Coronavirus Disease 2019 Patients (COVID-19). *Clin. Infect. Dis.* **76**, e26–e33 (2023).
7. Lin, D.-Y. *et al.* Nirmatrelvir or Molnupiravir Use and Severe Outcomes From Omicron Infections. *JAMA Netw Open* **6**, e2335077 (2023).
8. Peluso, M. J. *et al.* Effect of Oral Nirmatrelvir on Long COVID Symptoms: 4 Cases and Rationale for Systematic Studies. *Pathog Immun* **7**, 95–103 (2022).
9. Geng, L. N., Bonilla, H. F., Shafer, R. W., Miglis, M. G. & Yang, P. C. The use of nirmatrelvir-ritonavir in a case of breakthrough long COVID. *Explor. Res. Hypothesis Med.* **000**, 000–000 (2023).
10. Visvabharathy, L., Orban, Z. S. & Koralnik, I. J. Case report: Treatment of long COVID with a SARS-CoV-2 antiviral and IL-6 blockade in a patient with rheumatoid arthritis and SARS-CoV-2 antigen persistence. *Front. Med.* **9**, 1003103 (2022).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

11. Cohen, A. K. *et al.* Impact of extended-course oral nirmatrelvir/ritonavir (Paxlovid) in established Long COVID: Case series and research considerations. *Res Sq* (2023) doi:10.21203/rs.3.rs-3359429/v1.

12. McCarthy, M. W. Paxlovid as a potential treatment for long COVID. *Expert Opin. Pharmacother.* **24**, 1839–1843 (2023).

13. 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).

14. Perez Giraldo, G. S. *et al.* Neurologic Manifestations of Long COVID Differ Based on Acute COVID-19 Severity. *Ann. Neurol.* **94**, 146–159 (2023).

15. Hill, E. L. *et al.* Risk factors associated with post-acute sequelae of SARS-CoV-2: an N3C and NIH RECOVER study. *BMC Public Health* **23**, 2103 (2023).

16. Fernández-de-Las-Peñas, C. *et al.* Prevalence of post-COVID-19 symptoms in hospitalized and non-hospitalized COVID-19 survivors: A systematic review and meta-analysis. *Eur. J. Intern. Med.* **92**, 55–70 (2021).

17. Xie, Y., Choi, T. & Al-Aly, Z. Association of Treatment With Nirmatrelvir and the Risk of Post-COVID-19 Condition. *JAMA Intern. Med.* **183**, 554–564 (2023).

18. Durstenfeld, M. S. *et al.* Association of nirmatrelvir for acute SARS-CoV-2 infection with subsequent Long COVID symptoms in an observational cohort study. *J. Med. Virol.* **96**, e29333 (2024).

19. Congdon, S. *et al.* Nirmatrelvir/ritonavir and risk of long COVID symptoms: A retrospective cohort study. (2023) doi:10.21203/rs.3.rs-3231786/v1.

20. Gbinigie, O. *et al.* Platform adaptive trial of novel antivirals for early treatment of COVID-19 In the community (PANORAMIC): protocol for a randomised, controlled, open-label, adaptive platform trial of community novel antiviral treatment of COVID-19 in people at increased risk of more severe disease. *BMJ Open* **13**, e069176 (2023).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

21. Canadian Adaptive Platform Trial of Treatments for COVID-19 in Community Settings (CanTreatCOVID). *ClinicalTrials.gov* https://clinicaltrials.gov/study/NCT05614349 (2023).

22. CanTreatCOVID. *CanTreatCOVID - Finding therapies to treat COVID* https://cantreatcovid.org/ (2020).

23. Haendel, M. A. *et al.* The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment. *J. Am. Med. Inform. Assoc.* **28**, 427–443 (2021).

24. RECOVER: Researching COVID to Enhance Recovery. *RECOVER: Researching COVID to Enhance Recovery* https://recovercovid.org/.

25. Pfaff, E. R. *et al.* Synergies between centralized and federated approaches to data quality: a report from the national COVID cohort collaborative. *J. Am. Med. Inform. Assoc.* **29**, 609–618 (2022).

26. *Phenotype_Data_Acquisition: The Repository for Code and Documentation Produced by the N3C Phenotype and Data Acquisition Workstream*. (Github).

27. Hernán, M. A. & Robins, J. M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am. J. Epidemiol.* **183**, 758–764 (2016).

28. Hernán, M. A., Wang, W. & Leaf, D. E. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* **328**, 2446–2447 (2022).

29. Crosskey, M. *et al.* Reengineering a machine learning phenotype to adapt to the changing COVID-19 landscape: A study from the N3C and RECOVER consortia. *bioRxiv* (2023) doi:10.1101/2023.12.08.23299718.

30. Global Burden of Disease Long COVID Collaborators *et al.* Estimated Global Proportions of Individuals With Persistent Fatigue, Cognitive, and Respiratory Symptom Clusters Following Symptomatic COVID-19 in 2020 and 2021. *JAMA* **328**, 1604–1615 (2022).

31. British Medical Journal Publishing Group. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* **335**, (2007).

32. Brannock, M. D. *et al.* Long COVID risk and pre-COVID vaccination in an EHR-based cohort study

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

from the RECOVER program. *Nat. Commun.* **14**, 2914 (2023).

33. Centers for Disease Control (CDC). *New ICD-10-CM Code for the 2019 Novel Coronavirus (COVID-19)*. https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-19-508.pdf. (2020).
34. Iwasaki, A. & Putrino, D. Why we need a deeper understanding of the pathophysiology of long COVID. *Lancet Infect. Dis.* **23**, 393–395 (2023).
35. Anschuetz, N. DU Study Finds History of TBI Likely Worsens Long COVID Symptoms. *University of Denver* https://www.du.edu/news/du-study-finds-history-tbi-likely-worsens-long-covid-symptoms (2023).
36. Schneider, A. L. C. *et al.* Traumatic Brain Injury and Long-Term Risk of Stroke Among US Military Veterans. *Stroke* **54**, 2059–2068 (2023).
37. Granados, G. *et al.* Asbestos Exposure and Severity of COVID-19. *Int. J. Environ. Res. Public Health* **19**, (2022).
38. PTSD: National Center for PTSD. https://www.ptsd.va.gov/professional/treat/essentials/epidemiology.asp.
39. Munz, A. *Why Should Veterans Diagnosed with Mesothelioma Seek Legal Help?* (Mesothelioma Center - Vital Services for Cancer Patients & Families, 2022).
40. Morgan, R. O., Teal, C. R., Reddy, S. G., Ford, M. E. & Ashton, C. M. Measurement in Veterans Affairs Health Services Research: veterans as a special population. *Health Serv. Res.* **40**, 1573–1583 (2005).
41. RECOVER-VITAL: Platform Protocol to Measure the Effects of Antiviral Therapies on Long COVID Symptoms - Full Text View - ClinicalTrials.gov. https://classic.clinicaltrials.gov/ct2/show/NCT05595369.
42. Bendix, A., Mogg, K. & Murphy, J. Paxlovid may reduce the chance of long Covid. Why don't doctors prescribe it more? *NBC News* https://www.nbcnews.com/health/health-news/paxlovid-may-

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

prevent-long-covid-but-not-prescribed-rcna119737 (2023).

43. van Walraven, C., Bennett, C. & Forster, A. J. Administrative database research infrequently used validated diagnostic or procedural codes. *J. Clin. Epidemiol.* **64**, 1054–1059 (2011).

44. Stuart, E. A. Matching methods for causal inference: A review and a look forward. *Stat. Sci.* **25**, 1–21 (2010).

45. Imai, K., King, G. & Stuart, E. A. Misunderstandings between experimentalists and observationalists about causal inference. *J. R. Stat. Soc. Ser. A Stat. Soc.* **171**, 481–502 (2008).

46. Gupta, S. *et al.* Association Between Early Treatment With Tocilizumab and Mortality Among Critically Ill Patients With COVID-19. *JAMA Intern. Med.* **181**, 41–51 (2021).

47. Petito, L. C. *et al.* Estimates of Overall Survival in Patients With Cancer Receiving Different Treatment Regimens: Emulating Hypothetical Target Trials in the Surveillance, Epidemiology, and End Results (SEER)-Medicare Linked Database. *JAMA Netw Open* **3**, e200452 (2020).

48. Dickerman, B. A., García-Albéniz, X., Logan, R. W., Denaxas, S. & Hernán, M. A. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat. Med.* **25**, 1601–1606 (2019).

49. Admon, A. J. *et al.* Emulating a Novel Clinical Trial Using Existing Observational Data. Predicting Results of the PreVent Study. *Ann. Am. Thorac. Soc.* **16**, 998–1007 (2019).

50. Soriano, J. B. *et al.* A clinical case definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect. Dis.* **22**, e102–e107 (2022).

51. Department of Health and Human Services, Office of the Assistant Secretary for Health. *National Research Action Plan on Long COVID*. (2022).

52. Munblit, D. *et al.* Long COVID: aiming for a consensus. *Lancet Respir Med* **10**, 632–634 (2022).

53. EnSpark Consulting. *What We Heard: Engagement Report on the Working Definition for Long COVID*. (2023).

54. Wells, B. J., Chagin, K. M., Nowacki, A. S. & Kattan, M. W. Strategies for handling missing data in electronic health record derived data. *EGEMS (Wash DC)* **1**, 1035 (2013).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

55. Lin, K. J. *et al.* Out-of-system Care and Recording of Patient Characteristics Critical for Comparative Effectiveness Research. *Epidemiology* **29**, 356–363 (2018).

56. Wang, L. E., Shaw, P. A., Mathelier, H. M., Kimmel, S. E. & French, B. Evaluating Risk-Prediction Models Using Data from Electronic Health Records. *Ann. Appl. Stat.* **10**, 286–304 (2016).

57. Sanyaolu, A. *et al.* Comorbidity and its Impact on Patients with COVID-19. *SN Compr Clin Med* **2**, 1069–1076 (2020).

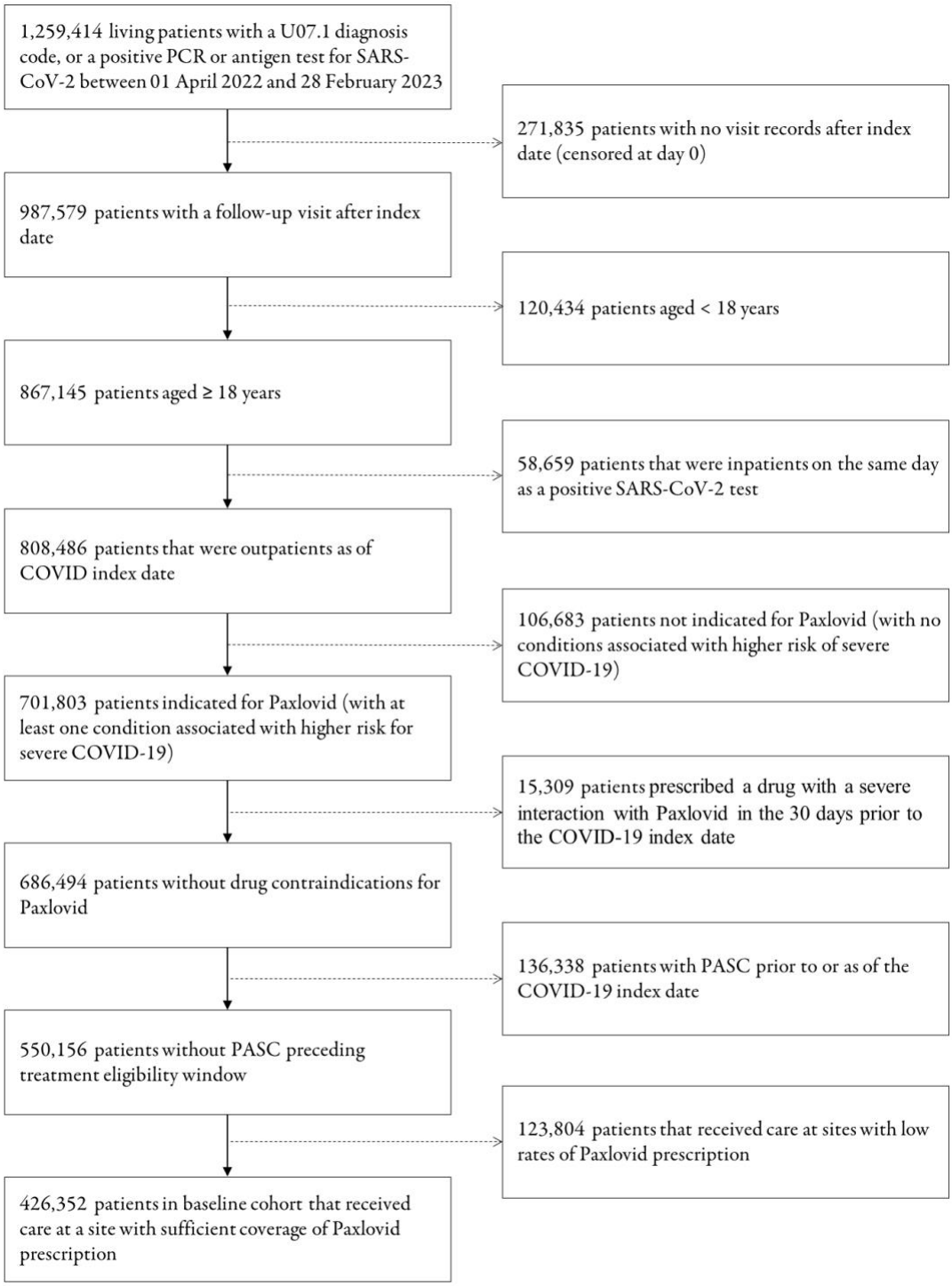
58. Fang, X. *et al.* Epidemiological, comorbidity factors with severity and prognosis of COVID-19: a systematic review and meta-analysis. *Aging* **12**, 12493–12503 (2020).

59. Honardoost, M., Janani, L., Aghili, R., Emami, Z. & Khamseh, M. E. The Association between Presence of Comorbidities and COVID-19 Severity: A Systematic Review and Meta-Analysis. *Cerebrovasc. Dis.* **50**, 132–140 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FIGURE LEGENDS

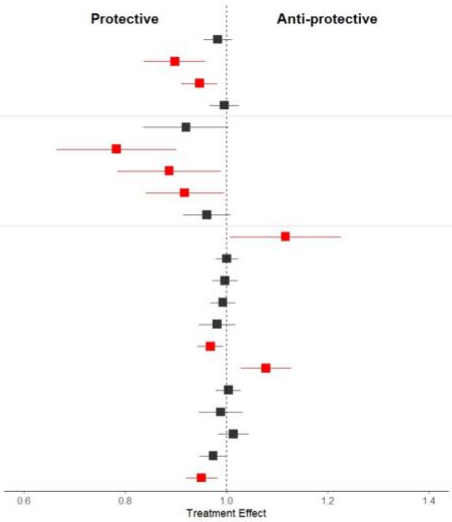
Figure 1: CONSORT Diagram: Study Cohort and Flow of Emulated Trial



medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2: Estimated Treatment Effects (Rate Ratios) of Paxlovid on PASC, across all analyses

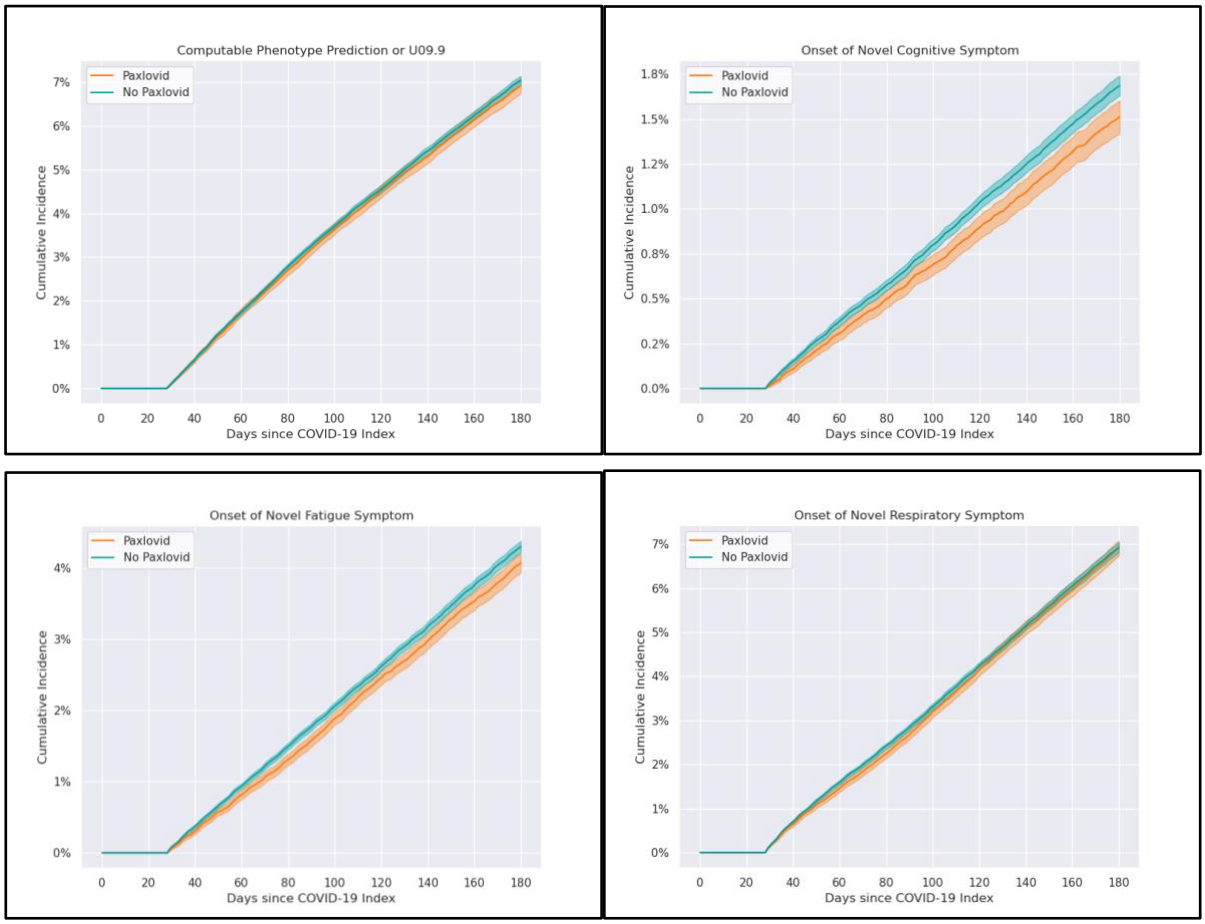
	Analysis	Treatment Effect (95% CI)	p-value
Main Analyses	Computable phenotype PASC prediction or U09.9 diagnosis[a]	0.983 (0.954-1.012)	0.245
	Cognitive Symptom Cluster[a]	0.898 (0.837-0.959)	<0.001
	Fatigue Symptom Cluster[a]	0.947 (0.911-0.982)	0.003
	Respiratory Symptom Cluster[a]	0.996 (0.966-1.025)	0.786
Sub-analyses	VA – like Cohort CP prediction or U09.9 diagnosis[a]	0.920 (0.836-1.005)	0.054
	VA – like Cohort Cognitive Symptom Cluster[a]	0.783 (0.664-0.901)	<0.001
	VA – like Cohort Fatigue Symptom Cluster[a]	0.887 (0.785-0.989)	0.021
	VA – like Cohort Respiratory Symptom Cluster[a]	0.918 (0.840-0.995)	0.029
	CP prediction or U09.9 diagnosis – Vaccination – Aware[a]	0.962 (0.915-1.009)	0.102
Sensitivity Analyses	U09.9 Code Diagnosis[a]	1.117 (1.007-1.226)	0.048
	PASC Computable Phenotype Threshold – 0.75[a]	1.001 (0.979-1.023)	0.920
	PASC Computable Phenotype Threshold – 0.8[a]	0.997 (0.972-1.023)	0.829
	PASC Computable Phenotype Threshold – 0.85[a]	0.993 (0.967-1.018)	0.582
	PASC Computable Phenotype Threshold – 0.95[a]	0.982 (0.945-1.018)	0.312
	Paxlovid Treatment as Index Event[a]	0.969 (0.943-0.994)	0.016
	Positive Lab – only Index Events[a]	1.079 (1.028-1.129)	0.003
	CP prediction or U09.9 diagnosis – 29 to 365 days[a]	1.004 (0.979-1.029)	0.751
	CP prediction or U09.9 diagnosis – 90 to 180 days[a]	0.989 (0.946-1.032)	0.615
	CP prediction or U09.9 diagnosis – 90 to 365 days[a]	1.014 (0.983-1.044)	0.378
	CP prediction or U09.9 diagnosis – 0 to 180 days[a]	0.974 (0.947-1.001)	0.055
	N3C Doubly Robust Adjustment[b]	0.951 (0.920-0.984)	0.003



Note: [a]Analysis reports Relative Risk; [b]Analysis reports adjusted Hazard Ratio

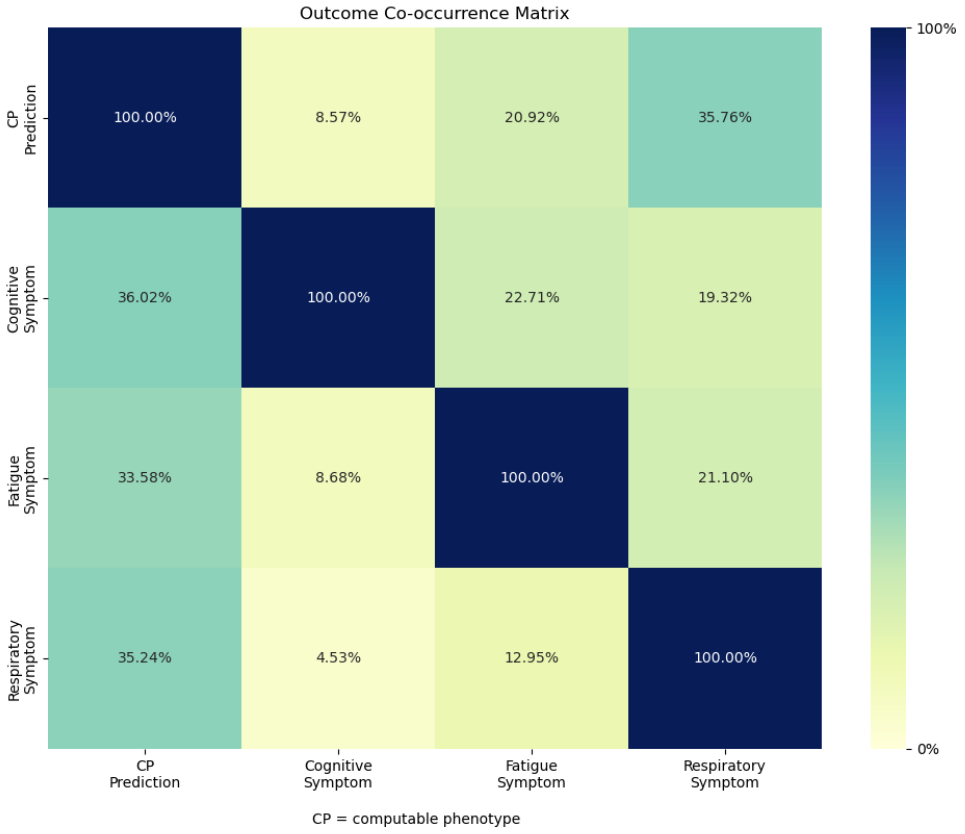
Figure 3: Cumulative incidence of PASC in Paxlovid treated vs. Non-Paxlovid-Treated patients by outcome measure; between 29-180 days

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

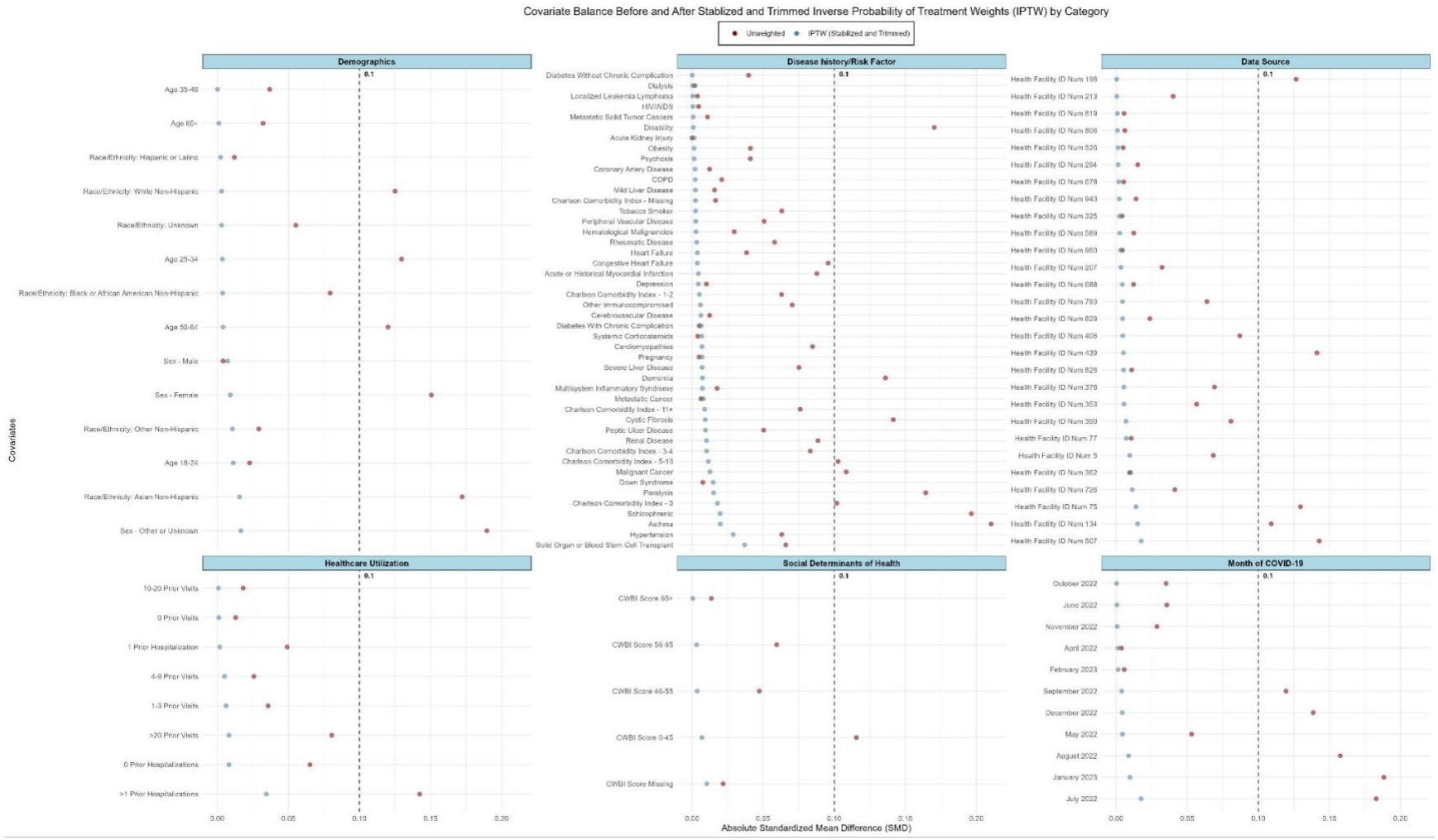


medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Extended Display Figure 1: Outcome co-occurrence matrix. Each cell represents the percentage of patients with the row outcome who also had the column outcome.

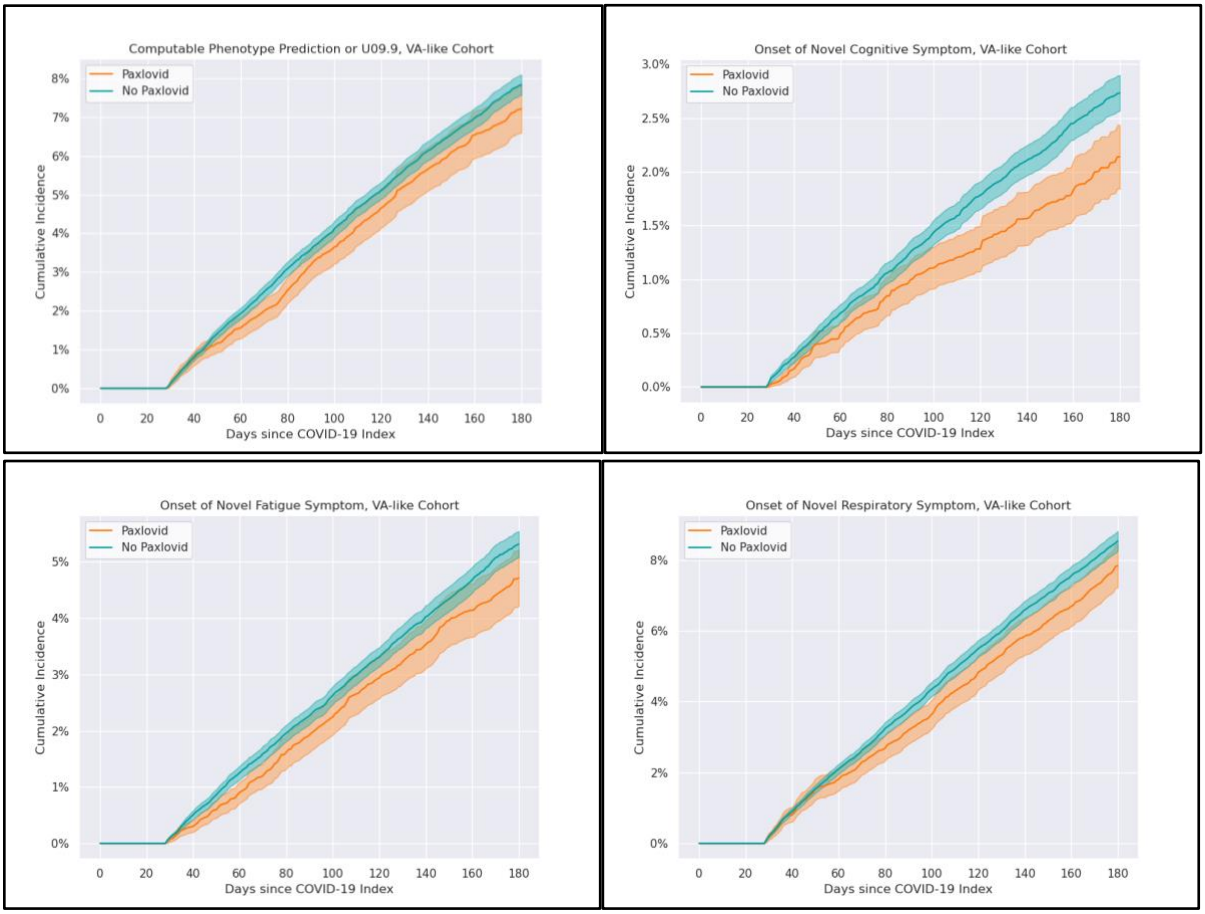


Extended Display Figure 2: Covariate balance before and after stabilized and trimmed inverse probability of treatment weighting (IPTW)



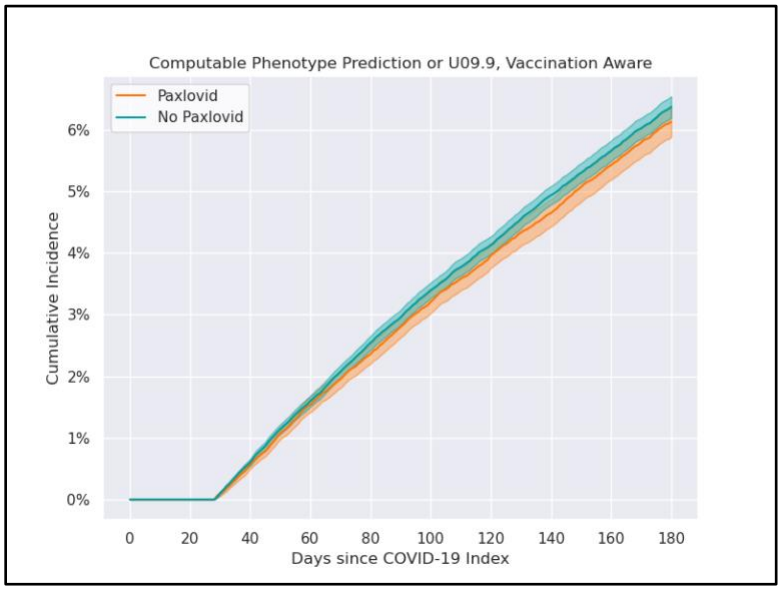
medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Extended Display Figure 4: Cumulative incidence of PASC in Paxlovid treated vs. Non-Paxlovid-Treated patients by outcome measure; between 29-180 days; VA-like subanalysis

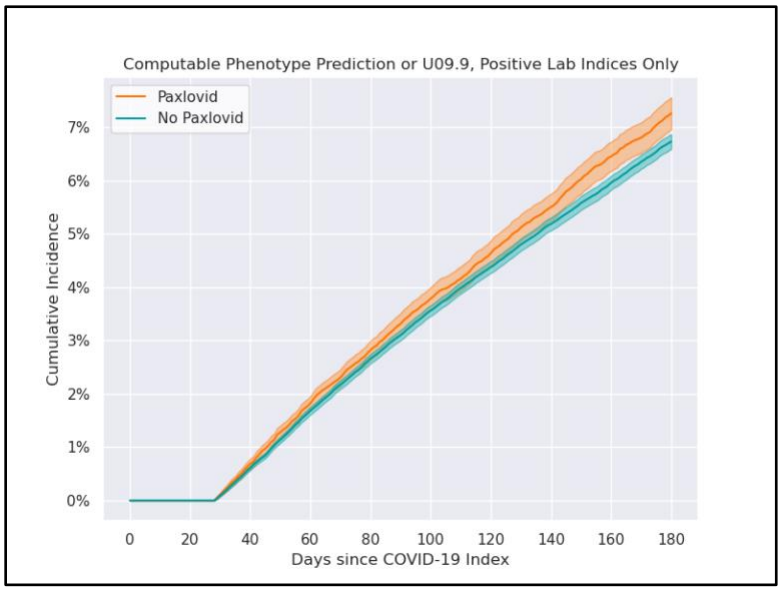


Extended Display Figure 5: Cumulative incidence of PASC in Paxlovid treated vs. Non-Paxlovid-Treated patients by predicted outcome from CP model with threshold of 0.9 or U09.9, additionally adjusted for vaccination status and among data partners meeting vaccination data quality criteria

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

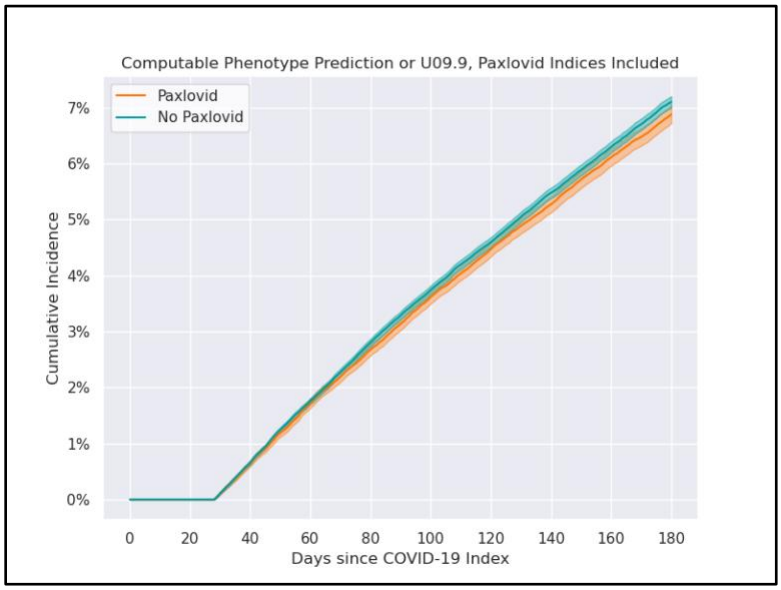


Extended Display Figure 6: Cumulative incidence of PASC in Paxlovid treated vs. Non-Paxlovid-Treated patients by predicted outcome from CP model with threshold of 0.9 or U09.9, positive lab required for COVID-19 index (i.e., U07.1 diagnoses without accompanying lab tests not included as index events)



Extended Display Figure 7: Cumulative incidence of PASC in Paxlovid treated vs. Non-Paxlovid-Treated patients by predicted outcome from CP model with threshold of 0.9 or U09.9, Paxlovid treatments without accompanying lab tests or U07.1 diagnoses included as index events

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

TABLES

Table 1: Descriptive population characteristics within the N3C Cohort

Characteristic	Treatment Group	
	No Paxlovid (N=303,166)	Paxlovid (N=123,186)
PASC		
Computable phenotype prediction or diagnosis[1]	17,238 (5.7%)	7,231 (5.9%)
Cognitive symptom cluster	4,221 (1.4%)	1,601 (1.3%)
Fatigue symptom cluster	10,944 (3.6%)	4,295 (3.5%)
Respiratory symptom cluster	17,595 (5.8%)	7,238 (5.9%)
Sex		
Female	189,621 (62.5%)	74,590 (60.6%)
Male	113,503 (37.4%)	48,576 (39.4%)
Missing	42 (0.0%)	20 (0.0%)
Age (in years)		
18-24	19,254 (6.4%)	2,908 (2.4%)
25-34	41,955 (13.8%)	9,109 (7.4%)
35-49	62,358 (20.6%)	22,270 (18.1%)
50-64	84,856 (28.0%)	39,255 (31.9%)
65+	94,743 (31.3%)	49,644 (40.3%)
Race and Ethnicity		
Asian Non-Hispanic	13,621 (4.5%)	5,653 (4.6%)
Black or African American Non-Hispanic	37,593 (12.4%)	11,892 (9.7%)
Hispanic or Latino Any Race	32,543 (10.7%)	9,796 (8.0%)
White Non-Hispanic	198,073 (65.3%)	88,250 (71.6%)
Other Non-Hispanic	5,237 (1.7%)	1,389 (1.1%)
Unknown	16,099 (5.3%)	6,206 (5.0%)
Charlson Comorbidity Index		
0	165,575 (54.6%)	66,812 (54.2%)
1-2	80,124 (26.4%)	38,600 (31.3%)
3-4	22,125 (7.3%)	9,148 (7.4%)
5-10	12,055 (4.0%)	4,017 (3.3%)
11+	994 (0.3%)	289 (0.2%)
Missing	22,293 (7.4%)	4,320 (3.5%)
Number of Visits in Prior Year		
0	26,602 (8.8%)	6,380 (5.2%)
1-3	52,029 (17.2%)	14,627 (11.9%)
4-9	71,675 (23.6%)	29,336 (23.8%)
10-20	75,420 (24.9%)	37,520 (30.5%)
> 20	77,440 (25.5%)	35,323 (28.7%)
Number of Hospitalizations in Prior Year		
0	287,800 (94.9%)	118,942 (96.6%)
1	12,369 (4.1%)	3,556 (2.9%)
> 1	2,997 (1.0%)	688 (0.6%)
Community Wellbeing Index[2]		
0-45	2,195 (0.7%)	670 (0.5%)
46-55	109,675 (36.2%)	37,068 (30.1%)
56-65	137,969 (45.5%)	58,324 (47.3%)
65+	22,136 (7.3%)	13,223 (10.7%)
Missing	31,191 (10.3%)	13,901 (11.3%)
Censoring Events		
Death	622 (0.2%)	107 (0.1%)
PASC diagnosis or prediction, day 0 to 28	5,145 (1.7%)	2,151 (1.8%)
Lost to follow-up (no further visits in EHR)	86,566 (28.65%)	31,378 (25.5%)
Paxlovid prescription >5 days after index	1,226 (0.4%)	N.A.
Month of COVID-19 diagnosis		
April 2022	18,340 (6.0%)	3,801 (3.1%)
May 2022	42,297 (14.0%)	13,075 (10.6%)

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

June 2022	43,248 (14.3%)	14,950 (12.1%)
July 2022	46,539 (15.4%)	18,691 (15.2%)
August 2022	38,485 (12.7%)	15,133 (12.3%)
September 2022	24,245 (8.0%)	10,028 (8.1%)
October 2022	19,118 (6.3%)	7,407 (6.0%)
November 2022	17,917 (5.9%)	8,526 (6.9%)
December 2022	23,910 (7.9%)	14,455 (11.7%)
January 2023	17,592 (5.8%)	9,734 (7.9%)
February 2023	11,475 (3.8%)	7,386 (6.0%)

Notes: [1]Any PASC (CP or U09.9) between 28-days following a positive SARS-CoV-2 test result to 180 days post-index; [2]CWBI is a measure of five interrelated community-level domains: Healthcare access (ratios of healthcare providers to population), Resource access (libraries and religious institutions, employment, and grocery stores), Food access (access to grocery stores and produce), Housing & transportation (home values, ratio of home value to income, and public transit use), and Economic security (rates of employment, labor force participation, health insurance coverage rate, and household income above the poverty level).[60]

Table 2: Protocol of a Target Trial Emulation to Estimate the Effect of Paxlovid Treatment during Acute COVID-19 on Cumulative PASC Incidence

Protocol Component	Description under Target Trial Conditions	Method of Target Trial Emulation
Eligibility criteria	Persons aged 18 and older, with no history of PASC, who are not currently hospitalized and who have an acute COVID-19 infection and are eligible for Paxlovid treatment due to presence of one or more risk factors for severe COVID-19 as per CDC guidelines.[61]	Persons aged 18 and older, with no history of PASC, who are not currently hospitalized, who have one or more risk factors for severe COVID-19 as per CDC guidelines[61] documented in their EHR, and who have a COVID-19 index (either a documented COVID-19 diagnosis or positive SARS-CoV-2 lab test) during the study period.
Treatment strategies	Paxlovid prescribed within a five-day grace period of the date patient presented with acute COVID-19.	Paxlovid prescribed within 5 days of COVID-19 index, indicated by a Paxlovid or nirmatrelvir drug exposure record in their EHR within 5 days of the COVID-19 diagnosis or positive SARS-CoV-2 lab test that constitutes their COVID-19 index date for the study.
Assignment procedures	Participants will be randomly assigned to treatment or control at the date they present with acute COVID-19 and will be aware of their treatment assignment.	Patients will be assigned weights based on treatment propensity scores to ensure exchangeability of treatment and control groups and emulate random assignment conditional on measured variables.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Follow-up period	Each patient will be followed for 180 days after treatment. Patients in the control group who obtain Paxlovid from an outside source will be censored. Patients who are lost to follow-up will be censored.	Patients will be censored at 180 days after COVID-19 index or the time of their last recorded visit, whichever is earlier. Patients in the control group will additionally be censored if they receive Paxlovid more than 5 days after COVID-19 index. We assume these cases constitute misdiagnoses or
Outcome	Clinical diagnosis of PASC within follow-up period	Clinical diagnosis of PASC, computable phenotype probability of PASC > 0.9, or onset of a novel PASC symptom between 29 and 180 days after COVID-19 index. Patients with a PASC diagnosis or computable phenotype prediction within 28 days of COVID-19 index will be censored.
Causal contrasts	Intention-to-treat effect	Intention-to-treat effect
Analysis plan	Measure relative risk of PASC diagnosis across treatment arms.	Estimate cumulative incidence of PASC in each treatment arm using Aalen-Johansen estimators weighted for treatment propensity; estimate relative risk based on point estimates and variances of cumulative incidence estimates.

Table 3: Estimated Cumulative Incidence of PASC among Paxlovid-treated patients compared to non-Paxlovid treated patients across all analyses

Analysis	Cumulative Incidence (95% CI)	
	Paxlovid	**No Paxlovid**
Main Results		
Computable phenotype PASC prediction or U09.9 diagnosis	0.069 (0.067, 0.071)	0.070 (0.069, 0.071)
Cognitive Symptom Cluster	0.015 (0.014, 0.016)	0.017 (0.016, 0.017)
Fatigue Symptom Cluster	0.041 (0.039, 0.042)	0.043 (0.042, 0.044)
Respiratory Symptom Cluster	0.069 (0.067, 0.071)	0.069 (0.068, 0.070)
Validation Analysis		
VA-like Cohort, Computable phenotype PASC prediction or U09.9 diagnosis	0.072 (0.066, 0.078)	0.078 (0.076, 0.081)
VA-like Cohort, Cognitive Symptom Cluster	0.021 (0.018, 0.024)	0.027 (0.026, 0.029)
VA-like Cohort, Fatigue Symptom Cluster	0.047 (0.042, 0.052)	0.053 (0.051, 0.055)

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

VA-like Cohort, Respiratory Symptom Cluster	0.078 (0.072, 0.084)	0.085 (0.083, 0.088)
Vaccination-Aware, CP prediction or U09.9 diagnosis	0.061 (0.059, 0.064	0.064 (0.062, 0.065

Extended Display Table 1: Cumulative incidence estimates across all analyses

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Analysis	Cumulative Incidence (95% CI)	
	Paxlovid	**No Paxlovid**
Main Results		
Computable phenotype PASC prediction or U09.9 diagnosis	0.069 (0.067, 0.071)	0.070 (0.069, 0.071)
Cognitive Symptom Cluster	0.015 (0.014, 0.016)	0.017 (0.016, 0.017)
Fatigue Symptom Cluster	0.041 (0.039, 0.042)	0.043 (0.042, 0.044)
Respiratory Symptom Cluster	0.069 (0.067, 0.071)	0.069 (0.068, 0.070)
Subanalyses		
VA-like Cohort, CP prediction or U09.9 diagnosis	0.072 (0.066, 0.078)	0.078 (0.076, 0.081)
VA-like Cohort, Cognitive Symptom Cluster	0.021 (0.018, 0.024)	0.027 (0.026, 0.029)
VA-like Cohort, Fatigue Symptom Cluster	0.047 (0.042, 0.052)	0.053 (0.051, 0.055)
VA-like Cohort, Respiratory Symptom Cluster	0.078 (0.072, 0.084)	0.085 (0.083, 0.088)
Vaccination-Aware, CP prediction or U09.9 diagnosis	0.061 (0.059, 0.064)	0.064 (0.062, 0.065)
Sensitivity Analyses		
U09.9 Code Diagnosis	0.008 (0.007, 0.008)	0.007 (0.007, 0.007)
PASC Computable Phenotype Threshold - 0.75	0.137 (0.135, 0.140)	0.137 (0.136, 0.139)
PASC Computable Phenotype Threshold - 0.80	0.114 (0.112, 0.117)	0.115 (0.113, 0.116)
PASC Computable Phenotype Threshold - 0.85	0.092 (0.090, 0.094)	0.093 (0.092, 0.094)
PASC Computable Phenotype Threshold - 0.95	0.047 (0.046, 0.049)	0.048 (0.047, 0.049)
Paxlovid Treatment as Index Event	0.069 (0.067, 0.070)	0.071 (0.070, 0.072)
Positive Lab-only Index Events	0.073 (0.070, 0.076)	0.067 (0.066, 0.069)
CP prediction or U09.9 diagnosis (29-365 days)	0.127 (0.124, 0.129)	0.126 (0.125, 0.128)
CP prediction or U09.9 diagnosis (90-180 days)	0.039 (0.038, 0.041)	0.040 (0.039, 0.041)
CP prediction or U09.9 diagnosis (90-365 days)	0.099 (0.096, 0.101)	0.098 (0.096, 0.099)
CP prediction or U09.9 diagnosis (0-180 days)	0.086 (0.084, 0.088)	0.088 (0.087, 0.089)
Doubly Robust Adjustment	Hazard Ratio: 0.951 (0.920, 0.984)	

Extended Display Table 2: ICD-10 codes used to define Global Burden of Disease symptom clusters[30]

ICD-10 Code	ICD-10 Code Description	Symptom Cluster

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

R404	Transient alteration of awareness	Cognitive
R410	Disorientation unspecified	Cognitive
R411	Anterograde amnesia	Cognitive
R412	Retrograde amnesia	Cognitive
R413	Other amnesia	Cognitive
R4182	Altered mental status unspecified	Cognitive
R41840	Attention and concentration deficit	Cognitive
R41841	Cognitive communication deficit	Cognitive
R4189	Other symptoms and signs involving cognitive functions and awareness	Cognitive
R419	Unspecified symptoms and signs involving cognitive functions and awareness	Cognitive
R531	Weakness	Fatigue
R5381	Other malaise	Fatigue
R5382	Chronic fatigue unspecified	Fatigue
R5383	Other fatigue	Fatigue
J9610	Chronic respiratory failure unspecified whether with hypoxia or hypercapnia	Respiratory
J9611	Chronic respiratory failure with hypoxia	Respiratory
J9612	Chronic respiratory failure with hypercapnia	Respiratory
J9620	Acute and chronic respiratory failure unspecified whether with hypoxia or hypercapnia	Respiratory
J9621	Acute and chronic respiratory failure with hypoxia	Respiratory
J9622	Acute and chronic respiratory failure with hypercapnia	Respiratory
J9690	Respiratory failure unspecified unspecified whether with hypoxia or hypercapnia	Respiratory
J9691	Respiratory failure unspecified with hypoxia	Respiratory

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

J9692	Respiratory failure unspecified with hypercapnia	Respiratory
J988	Other specified respiratory disorders	Respiratory
J989	Respiratory disorder unspecified	Respiratory
J99	Respiratory disorders in diseases classified elsewhere	Respiratory
R05	Cough	Respiratory
R0600	Dyspnea unspecified	Respiratory
R0602	Shortness of breath	Respiratory
R0603	Acute respiratory distress	Respiratory
R0609	Other forms of dyspnea	Respiratory
R071	Chest pain on breathing	Respiratory

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

METHODS

Ethics Approval

The N3C data transfer to the National Center for Advancing Translational Sciences is performed under a Johns Hopkins University reliance protocol (IRB00249128). The RECOVER Publications and Presentations Oversight Committee and the N3C Publications Committee approved this manuscript (manuscript IDs R_1OK0bL9LIsmS7Hu and 1020.56, respectively). Results downloads were approved per N3C Attribution and Publication Principles.[1]

Eligibility Criteria

The study period spanned April 1, 2022, to August 14, 2023, with an index cutoff date of February 28, 2023 (180 days before the end of the study period). We excluded the period between December 21, 2021 (date of Paxlovid EUA) and March 31, 2022 due to the variability in case counts and prescription patterns during the first wave of the Omicron variant.[2] We used data from RECOVER release v141 (August 2023) in the N3C Enclave.

Our inclusion criteria emulated the target trial's eligibility criteria: 1) having a documented COVID-19 index date within the study period (with *index date* defined as the earliest date of either a COVID-19 diagnosis [ICD-10 code U07.1] or a positive SARS-CoV-2 test result), 2) being ≥ 18 years of age at the COVID-19 index date (due to potential differences in clinical characteristics and prescription practices between pediatric and adult patients[3,4]), and 3) having ≥ 1 risk factor for severe COVID-19 per CDC guidelines (age ≥ 50 years or diagnosis of a comorbidity associated with higher risk of severe COVID-19[5]). For patients with > 1 COVID-19 index date in the study period, we selected a single index date per the following criteria: 1) if Paxlovid was prescribed within 5 days of one index date, use that index date, 2) if Paxlovid was prescribed within 5 days of > 1 index date, use the first, and 3) if Paxlovid was not prescribed within 5 days of any index date, use the first index date.

We also applied a set of exclusion criteria, to exclude: 1) patients who were hospitalized on the COVID-19 index date, 2) patients with PASC (see Treatment and Outcome) prior to or on the COVID-19 index date, 3) patients who were prescribed a drug with a severe interaction with

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Paxlovid in the 30 days prior to the COVID-19 index.[6] Furthermore, to ensure that data were captured from sites with high fidelity and adequate coverage, we only included data from 28 sites with at least 5% of eligible patients, and a minimum of 500 patients, treated with Paxlovid during the study period.

Treatment and Outcome

Eligible patients were categorized by their treatment exposure. The treatment group was defined as having been prescribed Paxlovid within 5 days of their COVID-19 index date. The control group was defined as the complement, with one exception. Patients who were prescribed Paxlovid within 5 days of COVID-19 index, but were hospitalized prior to treatment, were included in the control group and censored at the date of Paxlovid prescription (see Statistical Analysis for more on censoring). We took this approach because inpatient Paxlovid treatment (presumably after COVID-19 is already severe) is a different treatment modality, and we intended to study on-label outpatient treatment. We selected a treatment window of 5 days from COVID-19 index to adhere as closely as possible to treatment guidelines (within 5 days of symptom onset) with the available data. We identified 10 Observational Medical Outcomes Partnership [OMOP] concepts that correspond to Paxlovid in N3C and used these concepts to measure treatment.[7]

We considered two measures of the PASC outcome. To measure PASC overall, we used a computable phenotype: a machine learning model trained to predict PASC diagnoses (ICD-10 code U09.9). An earlier version of this computable phenotype was used in prior work.[8] For this study, we used an updated version better suited for the later phase of the pandemic.[8,9] The model gathers data for each patient in overlapping 100-day periods that progress through time, and issues a probability of PASC for each 100-day period. The model was trained to classify whether patients have a U09.9 ("Post COVID-19 Condition") ICD-10 diagnosis code in each period, based on the patients' diagnoses during each period. We followed patients for 180 days following their COVID-19 index date. PASC date was defined as the start date of the 100-day period which had the maximum computable phenotype prediction above a threshold of 0.9, or, if

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

present, the date of U09.9 diagnosis, whichever was earlier. Patients over 100 years old at COVID-19 index did not receive model scores and were excluded from analysis of this outcome

To measure PASC at a more granular level, we examined the PASC symptom clusters--cognitive, fatigue, and respiratory--proposed by the Global Burden of Disease (GBD) Study.[10] These clusters were the most frequently reported symptoms in a meta-analysis of Long COVID studies. Their full definitions are cognitive problems (forgetfulness or difficulty concentrating, commonly referred to as brain fog); persistent fatigue with bodily pain (myalgia) or mood swings; and ongoing respiratory problems (primarily shortness of breath and persistent cough). For the GBD symptom cluster outcomes, PASC date was defined as the first diagnosis date of any novel symptom in the cluster at least 28 days after COVID-19 index (we defined novel symptoms as symptoms that did not occur in the three years prior to COVID-19 index).The list of ICD-10 codes to define each GBD symptom cluster cluster was based on the GBD study and is presented in Extended Display Table 2.

A positive prediction from the computable phenotype model does not necessarily imply that a patient must have a positive outcome for one or more symptom clusters. The model considers many more diagnosis codes than those included in the symptom clusters (see the "SNOMED Roll Up" section in the supplement of Crosskey et al., 2023), and a positive prediction may be based on other diagnosis codes.[11] Also, the computable phenotype model does not include a novelty restriction. For example, if a patient had a dyspnea diagnosis in the three years prior to index, a post-acute dyspnea diagnosis would not count for the respiratory symptom cluster, but it would be considered by the computable phenotype model.

Statistical Analysis

Our estimand was the cumulative incidence of PASC from 29 to 180 days after COVID-19 index. We applied a potential outcomes framework to compare the rate of PASC among patients who received treatment to those who did not. We use inverse probability of treatment (IPT) weighting to emulate random assignment through exchangeability between treatment arms. Our treatment model included the following pre-treatment covariates: sex, age (binned), race and ethnicity, prior history of individual comorbid conditions captured in the Charlson Comorbidity

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

index, value of the composite Charlson Comorbidity Index (CCI; binned), prior history of conditions associated with risk of severe COVID-19 (as defined by the CDC Paxlovid eligibility criteria[5]), Community Well-Being Index (CWBI; binned), number of visits in the year prior to index (binned), number of hospitalizations in the year prior to index (binned), month of COVID-19 onset, and site of care provision. Our rationale for using these covariates is as follows. Many studies have shown disparity in COVID-19 treatment and outcome by race, ethnicity, and social determinants of health.[12–15] Sex, age, and comorbidities are known to affect care seeking and the outcome of COVID-19. Past healthcare utilization could affect the likelihood of treatment seeking and PASC documentation. Finally, the index month was included because Paxlovid treatment rates, viral variants, and infection rates changed during the study period. CCI was coded as missing when no condition records were present in N3C prior to index. CWBI was coded as missing when patient ZIP code was not reported. We used this treatment model to generate stabilized IPT weights trimmed at the 99.5th percentile. We assessed covariate balance using absolute standardized differences. To estimate the cumulative incidence of PASC, we used IPT-weighted Aalen-Johansen estimators. We used bootstrapping with 200 iterations to estimate the 95% confidence interval at a two-sided alpha of 0.05.

We censored patients at the following events: 1) death, 2) last documented visit in the study period, 3) PASC outcome within 28 days of COVID-19 index, and 4) 180 days after index (end of study period). We also censored patients in the control group if they received Paxlovid. This could occur if they received Paxlovid within 5 days of index, but after hospitalization (see Treatment and Outcome). It could also occur if they received Paxlovid later in the study period, but not within 5 days of a COVID-19 index (see Eligibility Criteria). By treating death as a censoring event rather than a competing risk, we estimate the direct effect of Paxlovid treatment on PASC incidence, rather than the total effect.[16] The total effect would include any effect of Paxlovid treatment on PASC incidence that is mediated by death, which is less interpretable.

In addition, we conducted two subanalyses and five sensitivity analyses.

Subanalysis

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

In the first subanalysis, we attempted to mirror the cohort used in Xie et al (2023). We refer to this as the VA-like cohort subanalysis. In this analysis , we used the same study start and end dates as Xie et al. (January 3, 2022, and December 31, 2022). To mirror VA demographics, we filtered the cohort to males >= 65 years old at COVID-19 index. To reflect the high continuity of care of the VA system, we filtered our cohort to patients with at least two visits in the year prior to COVID-19 index.

In the second supplementary analysis, we included COVID-19 vaccination status as a covariate, and replicated our primary analysis. We considered vaccination to be a plausible confounder of Paxlovid treatment and documented PASC, either through acute infection severity or propensity to seek care. We followed a similar procedure as in our earlier work estimating the effect of Paxlovid treatment on hospitalization.[17] Because vaccination status in N3C is subject to misclassification, we used a subcohort of patients from sites with reliable vaccination data. We categorized patients by their vaccination status prior to their COVID-19 index date, defined as having completed a full course of vaccination at least 14 days prior to index. Partially vaccinated patients and patients who became fully vaccinated fewer than 14 days prior to index were excluded from the analysis.

Sensitivity Analyses

We conducted five sensitivity analyses.

First, we used a doubly-robust estimation method in case the treatment model was misspecified. Targeted maximum likelihood estimation was not feasible with our cohort and computing environment, so we were unable to estimate cumulative incidence using a doubly-robust method. Instead, we estimated the hazard ratio (HR) of Paxlovid treatment as a secondary estimand. We used inverse probability of treatment-weighted Cox proportional hazards models adjusted for the same baseline covariates as the treatment model. The same bootstrap procedure was used to estimate confidence intervals.

Second, we tested various computable phenotype prediction thresholds. In addition to the 0.9 threshold used in the primary analysis, we tested prediction thresholds at 0.75, 0.8, 0.85, and 0.95.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Third, we included Paxlovid treatment as a COVID-19 index event. This added 33,571 additional patients who were treated with Paxlovid during the study period, but did not have a U07.1 diagnosis or a positive lab test in the five days prior to treatment.

Fourth, we also tested sensitivity to COVID-19 index definition by including only positive lab tests as index events (i.e., we did not include U07.1 diagnoses without accompanying lab results).

Fifth, we tested sensitivity to outcome definition in three ways: by requiring outcomes to occur 90 days after COVID-19 index (rather than 29 days), by observing patients for up to 365 days (rather than 180 days), and by treating PASC predictions or diagnoses from 0 to 28 days after index as outcome events rather than censoring events.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DATA AVAILABILITY STATEMENT

The N3C data transfer to NCATS is performed under a Johns Hopkins University reliance protocol (IRB00249128) or individual site agreements with the NIH. The N3C Data Enclave is managed under the authority of the NIH; more information can be found at ncats.nih.gov/n3c/resources. Enclave data is protected, and can be accessed for COVID-19-related research with an institutional review board-approved protocol and data use request. The Data Use Request ID for this study is RP-5677B5. Enclave and data access instructions can be found at https://covid.cd2h.org/for-researchers.

CODE AVAILABILITY STATEMENT

All code used to produce the analyses in this manuscript is available within the N3C Data Enclave to users with valid login credentials. Enclave access instructions can be found at https://covid.cd2h.org/for-researchers. Enclave access requires a signed data use agreement between NIH and a research institution (or directly with a researcher in the case of an individual citizen scientist who is not affiliated with an institution) and a signed data use request by the individual researcher. Researchers must also complete human subjects and security training. The registration checklist is available at https://covid.cd2h.org/account-instructions/checklist.jsp. Guidance locating the pertinent code workbooks within the Enclave is available upon request. In recognition that Enclave access may not be feasible for everyone, code is also available upon request.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

METHODS REFERENCES

1. Creators N3C Publication Committee N3C Governance Committee. *N3C and Tenant Pilot Attribution and Publication Principles*. doi:10.5281/zenodo.10698093.
2. Chatterjee, S., Bhattacharya, M., Nag, S., Dhama, K. & Chakraborty, C. A Detailed Overview of SARS-CoV-2 Omicron: Its Sub-Variants, Mutations and Pathophysiology, Clinical Characteristics, Immunological Landscape, Immune Escape, and Therapies. *Viruses* **15**, (2023).
3. Howard-Jones, A. R. *et al.* COVID-19 in children. II: Pathogenesis, disease spectrum and management. *J. Paediatr. Child Health* **58**, 46–53 (2022).
4. Bose-Brill, S. *et al.* Pediatric Nirmatrelvir/Ritonavir Prescribing Patterns During the COVID-19 Pandemic. *medRxiv* (2022) doi:10.1101/2022.12.23.22283868.
5. CDC. Interim Clinical Considerations for COVID-19 Treatment in Outpatients. *Centers for Disease Control and Prevention* https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/outpatient-treatment-overview.html (2023).
6. Paxlovid Drug-Drug Interactions. *COVID-19 Treatment Guidelines* https://www.covid19treatmentguidelines.nih.gov/therapies/antivirals-including-antibody-products/ritonavir-boosted-nirmatrelvir--paxlovid-/paxlovid-drug-drug-interactions/.
7. Hripcsak, G. *et al.* Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud. Health Technol. Inform.* **216**, 574–578 (2015).
8. Brannock, M. D. *et al.* Long COVID risk and pre-COVID vaccination in an EHR-based cohort study from the RECOVER program. *Nat. Commun.* **14**, 2914 (2023).
9. Pfaff, E. R. *et al.* Identifying who has long COVID in the USA: a machine learning approach using N3C data. *Lancet Digit Health* **4**, e532–e541 (2022).
10. Global Burden of Disease Long COVID Collaborators *et al.* Estimated Global Proportions of Individuals With Persistent Fatigue, Cognitive, and Respiratory Symptom Clusters Following

medRxiv preprint doi: https://doi.org/10.1101/2024.01.20.24301525; this version posted April 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Symptomatic COVID-19 in 2020 and 2021. *JAMA* **328**, 1604–1615 (2022).

11. Crosskey, M. *et al.* Reengineering a machine learning phenotype to adapt to the changing COVID-19 landscape: A study from the N3C and RECOVER consortia. *bioRxiv* (2023) doi:10.1101/2023.12.08.23299718.
12. Klein, E. J., Hardesty, A., Vieira, K. & Farmakiotis, D. Use of anti-spike monoclonal antibodies in kidney transplant recipients with COVID-19: Efficacy, ethnic and racial disparities. *Am. J. Transplant* **22**, 640–645 (2022).
13. Wiltz, J. L. *et al.* Racial and Ethnic Disparities in Receipt of Medications for Treatment of COVID-19 - United States, March 2020-August 2021. *MMWR Morb. Mortal. Wkly. Rep.* **71**, 96–102 (2022).
14. Wu, E.-L. *et al.* Disparities in COVID-19 Monoclonal Antibody Delivery: a Retrospective Cohort Study. *J. Gen. Intern. Med.* **37**, 2505–2513 (2022).
15. Boehmer, T. K. *et al.* Racial and Ethnic Disparities in Outpatient Treatment of COVID-19 - United States, January-July 2022. *MMWR Morb. Mortal. Wkly. Rep.* **71**, 1359–1365 (2022).
16. Young, J. G., Stensrud, M. J., Tchetgen Tchetgen, E. J. & Hernán, M. A. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat. Med.* **39**, 1199–1236 (2020).
17. Bhatia, A. *et al.* Effect of Nirmatrelvir/Ritonavir (Paxlovid) on Hospitalization among Adults with COVID-19: an EHR-based Target Trial Emulation from N3C. *medRxiv* (2023) doi:10.1101/2023.05.03.23289084.