

397 Demographic and co-morbidity characteristics of patients tested for SARS-CoV-2  
398 from March 2020 to January 2022 in a national clinical research network: results  
399 from PCORnet®

400

401 Short Title: Characteristics of patients tested for SARS-CoV-2 in a national research network

402 Jason P. Block, MD, MPH<sup>1\*</sup>; Keith A. Marsolo, PhD<sup>2</sup>; Kshema Nagavedu, MPH<sup>1</sup>; L. Charles  
403 Bailey, MD, PhD<sup>3</sup>; Tegan K. Boehmer, PhD<sup>4,5</sup>; Julia Fearrington<sup>1</sup>; Aaron M. Harris, MD,  
404 MPH<sup>4,5</sup>; Nedra Garrett, MS<sup>4</sup>; Alyson B. Goodman, MD, MPH<sup>4,5</sup>; Adi V. Gundlapalli, MD, PhD,  
405 MS<sup>4</sup>; Rainu Kaushal, MD, MPH<sup>6</sup>; Abel Kho, MD<sup>7</sup>; Kathleen M. McTigue, MD, MPH, MS<sup>8</sup>;  
406 Vinit P. Nair, BPharm, MS<sup>9</sup>; Jon Puro, MPA/HA<sup>10</sup>; Elizabeth Shenkman, PhD<sup>11</sup>; Mark G.  
407 Weiner, MD<sup>3</sup>; Neely Williams, EdD<sup>12</sup>; Thomas W. Carton, PhD, MS<sup>13</sup> on behalf of PCORnet®  
408 Network Partners†

409 <sup>1</sup>Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical  
410 School, Boston, Massachusetts, United States of America

411 <sup>2</sup>Department of Population Health Sciences, Duke University, Durham, North Carolina, United  
412 States of America

413 <sup>3</sup>Applied Clinical Research Center, Department of Pediatrics, Children's Hospital of  
414 Philadelphia, Philadelphia, Pennsylvania, United States of America

415 <sup>4</sup>US Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America

416 <sup>5</sup>US Public Health Service, Rockville, Maryland, United States of America

417 <sup>6</sup>Department of Population Health Sciences, Weill Cornell Medicine, New York, New York,  
418 United States of America

419 <sup>7</sup>Center for Health Information Partnerships, Northwestern University, Chicago, Illinois, United  
420 States of America

421 <sup>8</sup>Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of  
422 America

423 <sup>9</sup>PRACnet, Sharon, MA, United States of America

424 <sup>10</sup>OCHIN, Inc., Portland, OR, United States of America

425 <sup>11</sup>Health Outcomes & Biomedical Informatics, College of Medicine, University of Florida,  
426 Gainesville, Florida, United States of America

427 <sup>12</sup>Community Partners Network Inc, Nashville, Tennessee, United States of America

428 <sup>13</sup>Louisiana Public Health Institute, New Orleans, LA, United States of America

429

430

431 †Collaborative authors are listed in an Appendix

432

433 **\*Corresponding author:** [jblock1@partners.org](mailto:jblock1@partners.org) (JPB)

434

435

436 Abbreviations: EHR – electronic health record, CDM – Common Data Model, ETL – extract-  
437 transfer-load procedure to move data from a clinical data warehouse to the PCORnet Common  
438 Data Model; Arrhyth – arrhythmia, CAD – coronary artery disease, CHF – congestive heart  
439 failure, CKD – chronic kidney disease, Chron pulm – chronic pulmonary disease, Coag –  
440 Coagulopathy, DM2 – type 2 diabetes mellitus, HTN – hypertension, Mental – mental health  
441 disorder; Amb – ambulatory care setting, ED –emergency department, Inpt – inpatient, Vent –  
442 inpatient with procedure codes indicating use of mechanical ventilation; conv plasma –  
443 convalescent plasma, dex – dexamethasone, HCQ – hydroxychloroquine, remdes – remdesivir,  
444 steroids – corticosteroids other than dexamethasone, tocili - tocilizumab

445

446 Manuscript Word Count: 4096

447 Abstract Word Count: 413

448 **Abstract**

449

450 **Background:** Prior studies have documented differences in the age, racial, and ethnic  
451 characteristics among patients with SARS-CoV-2 infection. However, little is known about how  
452 these characteristics changed over time during the pandemic and whether racial, ethnic, and age  
453 disparities evident early in the pandemic were persistent over time. This study reports on trends  
454 in SARS-CoV-2 infections among U.S. adults from March 1, 2020 to January, 31 2022, using  
455 data from electronic health records.

456 **Methods and Findings:** We captured repeated cross-sectional information from 43 large  
457 healthcare systems in 52 U.S. States and territories, participating in PCORnet<sup>®</sup>, the National  
458 Patient-Centered Clinical Research Network. Using distributed queries executed at each  
459 participating institution, we acquired information for all patients  $\geq 20$  years of age who were  
460 tested for SARS-CoV-2 (both positive and negative results), including care setting, age, sex,  
461 race, and ethnicity by month as well as comorbidities (assessed with diagnostic codes).

462 During this time period, 1,325,563 patients had positive (13% inpatient) and 6,705,868 patients  
463 had negative (25% inpatient) viral tests for SARS-CoV-2. Disparities in testing positive were  
464 present across racial and ethnic groups, especially in the inpatient setting. Compared to White  
465 patients, Black or African American and other race patients had relative risks for testing positive  
466 of 1.5 or greater in the inpatient setting for 12 of the 23-month study period. Compared to non-  
467 Hispanic patients, Hispanic patients had relative risks for testing positive in the inpatient setting  
468 of 1.5 or greater for 16 of 23. Ethnic and racial differences were present in emergency  
469 department and ambulatory settings but were less common across time than in inpatient settings.  
470 Trends in infections by age group demonstrated higher test positivity for older patients in the  
471 inpatient setting only for most months, except for June and July of 2020, April to August 2021,  
472 and January 2022. Comorbidities were common, with much higher rates among those  
473 hospitalized; hypertension (38% of patients SARS-CoV-2 positive vs. 29% for those negative)  
474 and type 2 diabetes mellitus (22% vs. 13%) were the most common.

475 **Conclusion and Relevance:** Racial and ethnic disparities changed over time among persons  
476 infected with SARS-CoV-2. These trends highlight potential underlying mechanisms, such as  
477 poor access to care and differential vaccination rates, that may have contributed to greater  
478 disparities, especially early in the pandemic. Monitoring data on characteristics of patients  
479 testing positive in real time could allow public health officials and policymakers to tailor  
480 interventions to ensure that patients and communities most in need are receiving adequate  
481 testing, mitigation strategies, and treatment.

482 Multiple studies have documented the substantial differences in characteristics of patients with  
483 SARS-CoV-2 early in the pandemic, with individuals of minority race or ethnicity and older age  
484 bearing a much higher burden of severe COVID-19 illness (1–4). The US Centers for Disease  
485 Control and Prevention (CDC) has provided detailed information about cases and deaths by  
486 demographic characteristics, over time, gathered from data provided by state and local health  
487 departments and reference laboratories (5). These data have shown that rates of deaths per  
488 100,000 spiked among non-Hispanic Black individuals early in the pandemic, followed by a  
489 decline compared to other racial and ethnic groups in the Summer and Fall 2020 and another  
490 increase starting in early 2021. Similarly, death rates among Hispanic individuals have remained  
491 higher than non-Hispanic individuals throughout the pandemic. However, to date, there has been  
492 limited information about patient characteristics and disparities among COVID-19 patients  
493 across care settings over time, compared to non-COVID-19 patients, such as those testing  
494 negative for SARS-CoV-2. These comparisons can help determine whether characteristics of  
495 patients and temporal trends were unique to patients with COVID-19 or were consistent with  
496 trends among patients with other illnesses. Further, this information can foster understanding of  
497 pandemic dynamics, both for ongoing public health response and to help determine how best to  
498 prepare for future public health crises.

499  
500 Analyzing trends over time across care settings requires availability of longitudinal information  
501 on large populations throughout the entire course of the pandemic with detailed information  
502 about individuals, which is typically not available from public health authorities and public  
503 health departments. Research networks fill this void by using granular data from many healthcare  
504 systems, providing real-world information on diverse populations, followed for long periods.

505 Funded by the Patient-Centered Outcomes Research Network<sup>®</sup> (PCORI<sup>®</sup>), PCORnet<sup>®</sup>, the  
506 National Patient-Centered Clinical Research Network, is a network-of-networks that facilitates  
507 research using electronic health record (EHR), administrative health plan claims, and other  
508 healthcare data (6,7). PCORnet currently has participation from more than 60 data contributing  
509 sites (i.e., individual health systems) embedded within nine Clinical Research Networks (8).  
510 Using data from 23 months of the pandemic (March 2020 through January 2022) in 43  
511 institutions participating in PCORnet, we describe demographic trends over time of one of the  
512 largest US national sample of patients testing positive or negative for SARS-CoV-2.

513

## 514 **Methods**

### 515 *PCORnet<sup>®</sup> Common Data Model*

516 PCORnet<sup>®</sup> operates as a distributed research network, where data are held locally behind  
517 institutional firewalls, with queries sent to institutions from a central PCORnet<sup>®</sup> Coordinating  
518 Center. The participating sites utilize the PCORnet<sup>®</sup> Common Data Model, in which each data  
519 contributing site maps its clinical and/or administrative claims data into a standardized  
520 representation (9). These data are normally updated quarterly by sites, coinciding with data  
521 quality curation cycles in which data are examined for conformance with the Common Data  
522 Model specifications as well as data plausibility and data completeness.

523

524 Prompted by the COVID-19 pandemic, PCORnet<sup>®</sup> sites agreed to rapidly implement  
525 supplemental data extraction/curation procedures to allow for rapid capture of information on  
526 patients with COVID-19. Starting in early April 2020, sites began updating COVID-related data  
527 at least monthly (as often as daily) instead of quarterly. Rather than doing this for all patients

528 (which requires considerable processing time and effort), PCORnet<sup>®</sup> sites included a subset of all  
529 patients, including only those who had at least one International Classification of Diseases, 10<sup>th</sup>  
530 revision, Clinical Modification (ICD-10-CM) diagnostic code for a potential viral illness,  
531 including COVID-19, or had a lab result for a SARS-CoV-2 test recorded beginning January 1,  
532 2020 (**Supplemental Figure 1**) (10). Through January 31, 2022, this database included  
533 16,317,253 individual patients 20+ years of age.

534

### 535 *Queries and Patient Characterization*

536 From this larger population of patients, we used existing modular programs to characterize all  
537 patients tested for SARS-CoV-2 (**Supplemental Methods Note**). These programs return only  
538 aggregate data from sites. Submission, and use of these data are covered by a PCORnet-wide  
539 Data Sharing Agreement and individual site Institutional Review Board approvals for their  
540 PCORnet Common Data Model data. Using designated Logical Observation Identifiers Names  
541 and Codes (LOINC) (11), we captured information on patients with results of “positive,”  
542 “presumptive positive,” or “detected” on nuclear acid amplification (NAAT)/polymerase chain  
543 reaction (PCR) or antigen viral tests for SARS-CoV-2 (referred to as “testing positive” or  
544 “positive viral test” hereafter) as well as those with results of “negative (“testing negative” or  
545 “negative viral test”). Patients included in the negative viral test strata had only records of  
546 negative tests and no positive tests. We stratified patients based on the setting of their viral test:  
547 1) all, 2) ambulatory, 3) emergency department, and 4) inpatient. The latter three categories were  
548 mutually exclusive. We assigned a care setting by determining the setting for encounters before  
549 and after the test was performed (1 day before and 16 days following). To make care settings  
550 exclusive, we removed patients from a lower care setting stratum if they had encounters in a

551 higher care setting (for ambulatory strata, we assessed emergency and inpatient encounters)  
552 before and after their positive test (21 days before and 16 days following the date of the test)  
553 **(Supplemental Note)**. Patients identified as receiving care in the emergency department  
554 included patients who were discharged directly from this setting and not admitted to the hospital  
555 within 16 days; those admitted directly from the emergency department or soon thereafter were  
556 characterized with the inpatient care setting cohort. One-third of patients did not have a care  
557 setting associated with their positive test, perhaps because many patients were tested in drive-  
558 thru testing settings that might not have registered an encounter; some of these tests might also  
559 have been ordered during telehealth or virtual visits that were not specifically assessed in our  
560 queries. We captured all information with respect to the date for an index viral laboratory test,  
561 which was the first positive test for patients testing positive and the first negative test for patients  
562 testing negative. All codes for case definitions, comorbidities, treatments, and labs were posted  
563 on GitHub.

564

### 565 *Demographics, Comorbidities, and Disease Severity*

566 Using information available at the time of the test, we assessed demographic characteristics of  
567 patients (i.e., age group, sex, race, and ethnicity) using value sets defined in the PCORnet  
568 Common Data Model. Sex categories allowed for Male, Female, and Other. Several race  
569 categories (i.e., Multiple race, Other race, Native Hawaiian/Other Pacific Islander, and Alaskan  
570 Native/American Indian) were less common and typically <1% for all strata; these were grouped  
571 together under “Other Race.” Ethnicity was defined as Hispanic or not Hispanic, with a separate  
572 missing category. We assessed monthly trends by age group, race, and ethnicity by care setting.  
573 We assessed disparities by calculating monthly patient positivity as the total number of patients

574 testing positive separately for each racial, ethnic and age group, divided by the total number of  
575 patients tested in that month by care setting. We calculated relative risks (RRs) by month as the  
576 percentage of patients testing positive by racial, ethnic, and age group divided by the percentage  
577 in a reference group (White; non-Hispanic; age 20-39 years, respectively for each assessment of  
578 patient characteristics); these calculations were done separately for race, ethnicity, and age. We  
579 noted that disparities were present when RRs were  $\geq 1.5$ .

580  
581 We included a detailed assessment of comorbidities, which were selected based on information  
582 presented in publications regarding the association of comorbidities with COVID-19 outcomes  
583 (12–15). Patients were considered to have a comorbidity if they had at least two ICD-9-CM or  
584 ICD-10-CM codes for the disease/condition recorded any time over a three-year period prior to  
585 the index date of their viral test. We also captured use of critical care procedure/billing Current  
586 Procedural Terminology (CPT) codes (99291, 99292) during this period and mechanical  
587 ventilation usage through presence of ICD-10-CM, Healthcare Common Procedure Coding  
588 System (HCPCS) and CPT codes any time from the index date of the SARS-CoV-2 test through  
589 21 days after the index date. Lastly, we categorized weight status using the most recent height  
590 and weight available within the year prior to the index date. This study was reviewed by the  
591 Institutional Review Board of Harvard Pilgrim Health Care, Inc., which deemed it exempt from  
592 review because it was a public health surveillance project, conducted under the direction of a  
593 public health authority. This activity was reviewed by the CDC and was conducted consistent  
594 with applicable federal law and CDC policy.

595

596



597 **Results**

598 *Characteristics of patients testing positive for SARS-CoV-2*

599 The query captured aggregate data on 1,325,563 patients who tested positive for SARS-CoV-2  
600 and 6,705,868 patients who tested negative from March 1, 2020 through January 31, 2022.

601 Across the 43 sites, percent contributions to the overall sample ranged from <1% to 15% with 15  
602 sites each contributing  $\geq 3\%$  of the total; these 15 sites included 73% of all patients testing  
603 positive for SARS-CoV-2. Of patients testing positive, 38% were treated in the ambulatory  
604 setting, 15% were seen and discharged from the emergency department, and 13% were  
605 hospitalized (**Table 1**). Among hospitalized patients, 10% and 19% were documented as  
606 receiving mechanical ventilation and critical care treatment, respectively. Test positivity among  
607 patients (not tests) was 17% overall and 15%, 20%, and 9% in the ambulatory, emergency  
608 department, and inpatient setting, respectively.

609  
610 Black or African American patients were 16% of all patients testing positive but were 12%, 28%,  
611 and 23% of patients testing positive in the ambulatory, emergency department, and inpatient  
612 settings, respectively (**Table 1**). Other Race patients were 10% of all patients testing positive but  
613 were 8%, 16%, and 15% in the ambulatory, emergency department, and inpatient settings. The  
614 RR for testing positive for Black or African American patients vs. White patients was 1.0, 1.2,  
615 and 1.5 across the care settings, and RR for Other Race patients was 1.1, 1.3, and 1.8 (**Table 1**).  
616 Compared to non-Hispanic patients, the RR for testing positive in the ambulatory, emergency  
617 department, and inpatient settings for Hispanic patients was 1.5, 1.5, and 1.8 (**Table 1**).

618

619 Overall, more patients testing positive were in the 20 to 39 years age group (41%) than other age  
620 groups, primarily driven by patients testing positive in the ambulatory setting (**Table 1**). Patients  
621 20-39 years of age comprised the largest proportion of patients testing positive in the ambulatory  
622 and emergency department settings, but 55-74 year old patients comprised the largest proportion  
623 in the inpatient setting (Table 1). Compared to patients who were 20-39 years, the RR for testing  
624 positive in the inpatient setting was 1.7 for patients 40 to <55 years, 1.8 for patients 55 to <65  
625 years, 1.7 for patients 65 to <75 years, 1.9 for patients 75-<85 years, and 2.2 for patients 85+  
626 years. Women comprised a larger proportion of patients testing positive overall (56%), with men  
627 more likely to test positive only in the inpatient setting (RR 1.4).

628 **Fig 1. Monthly relative risks for testing positive for SARS-CoV-2 by race and ethnicity and**  
629 **care setting, March 1, 2020 to January 31, 2022.**

630 These figures show the relative risk for testing positive by month for Black or African American,  
631 Asian, Other Race or Missing Race compared to White race as a reference in the (A) ambulatory  
632 or (B) inpatient setting and those who were Hispanic, or had Missing Hispanic ethnicity  
633 compared to Not Hispanic as a reference in the (C) ambulatory or (D) inpatient setting. Care  
634 settings are mutually exclusive. Proportions of patients testing positive by care setting are  
635 available in Supplemental Figures 2 and 4; relative risks for testing positive in the emergency  
636 department are available in Supplemental Figures 3 and 5. Abbreviations: Hisp-Hispanic; Not  
637 Hisp-Not Hispanic; Miss-Missing.

638

639 *Trends by race, ethnicity, and age*

640 Disparities by race were most apparent during the early phase of the pandemic, especially in the  
641 inpatient setting. Among patients hospitalized with SARS-CoV-2 in March and April 2020,  
642 Black or African American, Other race, and White patients comprised nearly an equal proportion  
643 of patients (**Supplemental Figure 2**). Black or African American patients still comprised more  
644 than 20% of patients hospitalized for 17 of the 23 months examined. When we calculated relative  
645 risks by month, we demonstrated that hospitalized Black or African American patients were  
646 twice as likely to test positive for SARS-CoV-2 compared to White hospitalized patients from

647 April through August 2020 and 1.5 times as likely to tested positive for 12 of the 23 months,  
648 with the lowest RRs from September 2020 to January 2021 and again from June to November  
649 2021 (**Figure 1**). Patients with “Other Race” had similar patterns when compared to patients of  
650 White race, with relative risks of 1.5 or greater for 12 of the 23 months. Patients with Missing or  
651 Asian race followed a similar pattern to Black and Other Race groups early in the pandemic and  
652 from December 2020 through April 2021. These disparities were similar in the emergency  
653 department, with lower relative risks than the inpatient setting (**Supplemental Figure 3**).  
654 Disparities also were noted in the ambulatory setting but only in the early months of the  
655 pandemic (**Figure 1**).

656

657 Trends were similar when examined by ethnicity, with a disproportionate number of patients  
658 with Hispanic ethnicity testing positive across all care settings; these disparities persisted through  
659 mid-2021 in the emergency department in inpatient settings (**Figure 1; Supplemental Figures**  
660 **4- 5**). Compared to non-Hispanic patients, Hispanic patients had relative risks for testing positive  
661 in the inpatient setting of 1.5 or greater for 16 months.

662 **Fig 2. Monthly relative risks for testing positive for SARS-CoV-2 by age group and care**  
663 **setting, March 1, 2020 to January 31, 2022.**

664 These figures show the relative risks for testing positive by month and age group in the (A)  
665 ambulatory and (B) inpatient setting; reference was 20-39 years of age. Proportions of patients  
666 testing positive by care setting are available in Supplemental Figure 6; relative risks for testing  
667 positive in the emergency department are available in Supplemental Figure 7. Care settings are  
668 mutually exclusive.

669

670 Trends in infections by age demonstrated consistently higher test positivity for 20-39-year-olds  
671 in the ambulatory setting (**Figure 2, Supplemental Figures 6-7**). In the emergency department,  
672 younger patients were more likely to test positive than older patients in most months except for  
673 March to April 2020 (**Supplemental Figure 7**). Patients in older age groups (55 years and

674 higher) were more likely to test positive in the inpatient setting than patients who were 20-39  
675 years, with RRs  $\geq 1.5$  consistently except in June-July 2020, April-July 2021, and December  
676 2021-January 2022 (**Figure 2, Supplemental Figure 6**); patients who were 40 to 54 years had  
677 RRs  $\geq 1.5$  for nearly all months of the pandemic compared to patients 20-39 years, except  
678 December 2021-January 2022.

679

### 680 *Comorbidities*

681 Among patients testing positive for SARS-CoV-2, hypertension was the most common  
682 comorbidity (19% for all patients) followed by type 2 diabetes (DM2) (9%), chronic pulmonary  
683 disease (7%), and mental health disorders (8%); 7% of patients had hypertension and either Type  
684 1 (T1DM) or T2DM (**Figure 3; Supplemental Table 3**). These diagnoses were more common  
685 among patients with SARS-CoV-2 treated in the inpatient setting, with hypertension present in  
686 38% and DM2 in 22% (19% had hypertension and either DM1 or DM2) of inpatients. The  
687 prevalence of obesity also was high at 46% among all patients testing positive but was similar  
688 across care settings (45% to 48%). Patients testing negative while inpatient had a lower  
689 prevalence of hypertension (29% vs. 38%), diabetes (13% vs. 22%), and obesity (41% vs. 47%)  
690 compared with those testing positive (**Supplemental Table 4**); prevalence was similar for  
691 coronary artery disease (10% vs. 13%), mental health disorders (11% vs. 12%), or chronic  
692 pulmonary disease (11% vs. 14%).

### 693 **Fig 3. Comorbidities for patients with positive viral tests for SARS-CoV-2 across care** 694 **settings, March 1, 2020 to January 31, 2022.**

695 These figures show prevalence of various comorbidities for patients with positive viral tests for  
696 SARS-CoV-2 across care settings. Each condition includes a broad capture of ICD-9-CM and  
697 ICD-10-CM codes, with codes lists available on Github. For example, mental health codes  
698 include those for a large number associated with psychotic and mood disorders. Care settings  
699 were mutually exclusive. Abbreviations: AV-ambulatory visit; ED-emergency department; IP-  
700 inpatient or hospitalized; HTN-hypertension; DM2-diabetes mellitus type 2, DM-diabetes

701 mellitus Types 1 and 2, Arrhythmia; Chron Pulm-Chronic Pulmonary Disease; CAD-  
702 coronary artery disease; CKD-chronic kidney disease; MH Dis-mental health disease; CHF-  
703 congestive heart failure.  
704

## 705 **Discussion**

706 Using real world electronic health care data, we examined a large population of patients with  
707 SARS-CoV-2 infection, using information on 1.3 million in 43 institutions testing positive for  
708 SARS-CoV-2 infection and 6.7 million who tested negative without any positive tests through  
709 January 2022. Unlike most other studies, we compared patients across care settings, with rich  
710 demographic and comorbidity data by ambulatory, emergency department, and inpatient settings;  
711 we also compared patients who tested positive to those who tested negative. Trends by race,  
712 ethnicity and age revealed important information about health inequities during the COVID-19  
713 pandemic. While variable over time, individuals who were Black or African American or  
714 Hispanic were more likely to test positive compared to White and non-Hispanic patients in the  
715 inpatient setting for most months.

716  
717 Beyond basic demographic trends in cases and death overall, as reported by public health  
718 authorities, few studies have reported demographic trends in infections. Studies that have  
719 reported on these trends have covered short-term trends early in the pandemic; this study covers  
720 23 months of the pandemic. Using data from a national reference laboratory (ARUP  
721 Laboratories, with nearly half of the sample of 19,320 patients from Utah), investigators at the  
722 US Centers for Disease Control and Prevention (CDC) found a changing age distribution of  
723 patients infected with SARS-CoV-2, with a decline in median age in June and July 2020 of 35.8  
724 years compared to 40.5 in March and April 2020 (16). From May to August 2020, state  
725 electronic laboratory reporting data from 37 states showed similar trends by age, with 20-to-29-

726 year-olds having the highest incidence of SARS-CoV-2 among all age groups from June through  
727 August (4). This increase in incidence among young people continued through early September,  
728 as 18- to 22-year-olds started to return to colleges (3). In our study, we found peaks in test  
729 positivity of young people during the Summer of 2020 and 2021.

730

731 In our study, racial and ethnic disparities in infections, especially among those hospitalized, were  
732 evident throughout the pandemic but were most evident early in the pandemic. Disparities by  
733 race, defined by test positivity, in the inpatient setting were highest early in the pandemic and  
734 during the Winters of 2020/21 and 2021/22. Ethnic disparities were persistent for most of the  
735 pandemic in the inpatient setting, with a higher proportion of Hispanic individuals testing  
736 positive when hospitalized compared to non-Hispanic patients; this trend ended in Fall 2021.

737 Other studies have found similar patterns for shorter periods. In COVID-19 mortality data from  
738 the National Vitals Statistics System, disparities were evident from May to August 2020, with  
739 some differences over time (1). Overall, 19% of individuals who died from COVID-19 were  
740 Black or African American, with a reduction in proportion from 20% in May 2020 to 17% in  
741 August; 24% of all people who died were Hispanic, with an increase from 16% to 26% during  
742 this same time period. From January 26 to October 3, 2020, there were nearly 300,000 excess  
743 deaths in the United States; this rate summarizes all-cause mortality compared to expected rates  
744 from prior years, in this case 2015 to 2019 (2). Substantial disparities were present. The excess  
745 death rate was 11.9% for Whites but 29% to 37% for racial minorities and 53% for Hispanic  
746 persons; this was highest in March and April 2020 with a second peak among Hispanics in July  
747 and August 2020.

748

749 Among all conditions assessed, we found higher prevalence of hypertension, type 2 diabetes  
750 mellitus, and obesity among adults treated in the inpatient setting with SARS-CoV-2 compared  
751 to those testing negative. These differences could reflect the demographic differences, including  
752 age, found between those testing negative and positive or an increased risk of severe disease  
753 among patients with these conditions. Diabetes has been found to be an independent risk factor  
754 for severe COVID-19, including mortality, likely secondary to underlying physiologic  
755 mechanisms, such as immune effects from diabetes and a higher affinity for cellular binding of  
756 the virus in patients with diabetes (17–20). The high prevalence of diabetes especially in our  
757 study is further evidence of a greater burden of severe SARS-CoV-2 in this population.

758  
759 COVID-19 surveillance in PCORnet provides several unique opportunities. Obtaining real-  
760 world, comprehensive information on a diverse, national population of patients with COVID-19  
761 is critical to help guide policymakers as well as clinical and public health leaders regarding  
762 planning for future COVID-19 response. The network infrastructure and large samples allows for  
763 creation of research-ready cohorts and in-depth assessment of patients, based on demographic or  
764 clinical characteristics. Analyses can be done in a distributed fashion without sharing patient-  
765 level data, including predictive modeling, and aggregation across diverse sites can increase  
766 generalizability. Large sample sizes with granular clinical information will be increasingly  
767 important to capture information on patient subgroups, defined by demographics or less common  
768 presentations or outcomes of COVID-19, such as rare complications of SARS-CoV-2. Prior  
769 PCORnet surveillance studies have examined disparities in use of monoclonal antibodies (21),  
770 the occurrence of post-acute sequelae of COVID-19 (22), and the incidence of myocarditis and  
771 pericarditis after SARS-CoV-2 infection and COVID-19 vaccination (23). The investment in

772 changing the processes for data management – moving to frequent updates for patients with  
773 respiratory illnesses from a quarterly update – prepared the network to make ongoing  
774 contributions to the investigation of COVID-19. Other national initiatives, such as The National  
775 COVID Cohort Collaborative (N3C) are gathering similar information from institutions, with  
776 patient-level data extracts, to allow for varied research on COVID-19 over time (24).

777

### 778 *Limitations*

779 Several limitations are inherent to working with structured healthcare data. First, electronic  
780 health records and administrative claims data may not have complete data on some patients,  
781 especially patients tested and treated entirely outside of a PCORnet health system or outside of  
782 the health plan insurance coverage. In such cases, the PCORnet Common Data Model may not  
783 have any information about a SARS-CoV-2 infection or inadequate information on chronic  
784 diseases and medication use. For example, some institutions had less data on symptomatic  
785 ambulatory patients, who may have been referred to outside facilities for testing; other  
786 institutions did their own testing for symptomatic ambulatory patients. Most importantly,  
787 missing information could have led to misclassification, such as when a patient only had  
788 negative tests in the healthcare system but tested positive at an outside institution or on a home  
789 test, results that would not be captured in the PCORnet Common Data Model. If missingness  
790 was differential by demographic characteristics, proportions reported could be inaccurate.

791 Second, even for patients who received all care within a PCORnet institution, missing data were  
792 still possible, as we found for variables such as race, weight status, and care setting. If  
793 missingness was non-random, which could be the case if data were less available on patients of  
794 certain race categories, our results could provide some incomplete conclusions. Third, the results



795 we report are descriptive and unadjusted for potential confounders. We discovered higher rates  
796 of several chronic conditions among patients with SARS-CoV-2 compared to those without;  
797 these differences might be explained by demographic or age differences. We required at least  
798 two ICD-10-CM codes for chronic conditions, which may have led to an undercount, especially  
799 for those patients who were not seen often in the healthcare system prior to their SARS-CoV-2  
800 test. Fourth, we report differences based on race/ethnicity information available in the EHR. We  
801 do not have further information in the Common Data Model, such as socioeconomic status or  
802 other information that underlie disparities, such as exposure to racism or other social  
803 determinants. We also do not have access to payer status, which could provide some information  
804 on disparities by type of insurance, such as governmental or private. Finally, PCORnet does not  
805 allow for representative estimates of overall COVID-19 cases across the U.S. Because PCORnet  
806 has a fixed number of institutions and is not designed to be representative of the U.S. population,  
807 the prevalence of cases across states in PCORnet is not equivalent to the actual state prevalence  
808 of patients who have been tested for SARS-CoV-2.

809

### 810 *Conclusion*

811 PCORnet has captured information on one of the largest samples to date of patients tested for  
812 SARS-CoV-2 infection. Racial and ethnic minority patients have borne a disproportionate  
813 burden of the pandemic, making up a larger proportion of those infected than other groups,  
814 especially for those hospitalized. Monitoring these data in real time could allow public health  
815 officials and policymakers to develop messaging and targeted programs for these communities to  
816 ensure that they are receiving adequate testing, mitigation strategies, and treatment. Patients

817 diagnosed with COVID-19 were found to have a high burden of underlying conditions,  
818 especially among those hospitalized.  
819

## 820 **Acknowledgement**

821 This study was funded through a cooperative agreement to the Task Force for Global Health,  
822 through the US Centers for Disease Control and Prevention: COVID-19 Electronic Healthcare  
823 Data Initiative. Each of the institutions received funding through this agreement.

824 The Clinical Research Networks and Coordinating Center reported in this publication are  
825 Network Partners in PCORnet, the National Patient-Centered Clinical Research Network.  
826 PCORnet has been developed with funding from the Patient-Centered Outcomes Research  
827 Institute (PCORI). The Network Partners' participation in PCORnet is funded through the  
828 following PCORI Awards: Coordinating Center (PCORI-CC2-Duke-2016); ADVANCE (RI-  
829 CRN-2020-001); CaPriCORN (RI-CRN-2020-002); GPC (RI-CRN-2020-003); HealthCore (HP-  
830 1510-32545); Insight (RI-CRN-2020-004); OneFlorida (RI-CRN-2020-005); PaTH (RI-CRN-  
831 2020-006); PEDSnet (RI-CRN-2020-007); REACHnet (RI-CRN-2020-008); and, STAR (RI-  
832 CRN-2020-009). Authors report no competing interests associated with this publication.

833 The funder provided support in the form of salaries or stipends for all authors. As the funding  
834 organization of the network infrastructure for PCORnet, PCORI participates in discussions and  
835 high-level decision-making about the network and its strategic priorities. With this participation,  
836 PCORI provided support and approval for the initiative to alter the PCORnet infrastructure and  
837 support queries related to COVID-19. However, PCORI but had no specific role in the study  
838 design, data collection and analysis, decision to publish, or preparation of the manuscript. The  
839 specific roles of these authors are articulated in the 'author contributions' section.

840 The findings and conclusions in this report are those of the authors and do not necessarily  
841 represent the official position of CDC or the US Public Health Service.

842

843

## 844 **Collaborative Authors: *PCORnet*<sup>®</sup> Network Partners**

845 Faraz S. Ahmad, MD, MS, Northwestern University Feinberg School of Medicine; Saul Blecker,  
846 NYU Grossman School of Medicine; H. Ryan Carnahan, PharmD, MS, The University of Iowa  
847 College of Public Health, Department of Epidemiology; Olveen Carrasquillo, MD, MPH,  
848 University of Miami Miller School of Medicine; Bernard P. Chang, MD, PhD, Columbia  
849 University Irvine Medical Center; Katherine Chung-Bridges, MD, MPH, Health Choice  
850 Network; Lindsay G. Cowell, MS, PHD, UT Southwestern Medical Center; Janis L. Curtis,  
851 MSPH, MA, Duke University School of Medicine; Christine Draper, Department of Population  
852 Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School; Soledad A.  
853 Fernandez, PhD, The Ohio State University; Christopher B. Forrest, MD, PhD, Children's  
854 Hospital of Philadelphia; Nidhi Ghildayal, PhD, Department of Population Medicine, Harvard  
855 Pilgrim Health Care Institute, Harvard Medical School; David A Hanauer, MD, MS, University  
856 of Michigan; Rachel Hess, MD, MS, University of Utah; Benjamin D. Horne, PhD, MStat, MPH,  
857 Intermountain Medical Center Heart Institute, Stanford University School of Medicine; Philip  
858 Giordano, MD, FACEP, Orlando Health, Inc.; William Hogan, MD, MS, University of Florida;  
859 Wenke Hwang, PhD, Penn State University College of Medicine; Harold P. Lehmann MD,

860 PhD, Johns Hopkins School of Medicine; Abel Kho, MD, MS, Northwestern University; Scott  
861 Mackey, DO, LCMC Health; Kenneth H. Mayer MD, Fenway Health; Narayana Mazumder, MS,  
862 Allina Health; James C. McClay, MD, University of Nebraska Medical Center; J. Greg Merritt,  
863 PhD, Patient is Partner, LLC; Abu Saleh Mohammad Mosa, Phd, MS, FAMIA, University of  
864 Missouri; Samyuktha Nandhakumar, MS, North Carolina Translational and Clinical Sciences  
865 Institute, UNC School of Medicine; Bridget Nolan, Department of Population Medicine, Harvard  
866 Pilgrim Health Care Institute; Jihad S. Obeid, MD, FAMIA, Medical University of South  
867 Carolina; Brian Ostasiewski, Wake Forest School of Medicine; Anuradha Paranjape, MD, MPH,  
868 Lewis Katz School of Medicine at Temple University; Lav Parshottambhai Patel, BS, MS,  
869 University of Kansas Medical Center; Patricia S. Robinson, PhD, APRN, AdventHealth  
870 Research Institute; Alexander Stoddard, MS, CTSI, Medical College of Wisconsin; William E.  
871 Trick, Cook County Health; Russell L. Rothman, MD, MPP, Vanderbilt University Medical  
872 Center; Neely Williams, MDiv, EdD  
873

874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919

## References

1. Gold JAW, Rossen LM, Ahmad FB, Sutton P, Li Z, Salvatore PP, et al. Race, Ethnicity, and Age Trends in Persons Who Died from COVID-19 - United States, May-August 2020. *MMWR Morbidity and mortality weekly report*. 2020;69(42):1517-21.
2. Rossen LM, Branum AM, Ahmad FB, Sutton P, Anderson RN. Excess Deaths Associated with COVID-19, by Age and Race and Ethnicity - United States, January 26-October 3, 2020. *MMWR Morbidity and mortality weekly report*. 2020;69(42):1522-7.
3. Salvatore PP, Sula E, Coyle JP, Caruso E, Smith AR, Levine RS, et al. Recent Increase in COVID-19 Cases Reported Among Adults Aged 18-22 Years - United States, May 31-September 5, 2020. *MMWR Morbidity and mortality weekly report*. 2020;69(39):1419-24.
4. Boehmer TK, DeVies J, Caruso E, van Santen KL, Tang S, Black CL, et al. Changing Age Distribution of the COVID-19 Pandemic - United States, May-August 2020. *MMWR Morbidity and mortality weekly report*. 2020;69(39):1404-9.
5. CDC COVID data tracker 2021 [Available from: <https://covid.cdc.gov/covid-data-tracker/#demographicsvertime>].
6. Collins FS, Hudson KL, Briggs JP, Lauer MS. PCORnet: turning a dream into reality. *Journal of the American Medical Informatics Association : JAMIA*. 2014;21(4):576-7.
7. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association : JAMIA*. 2014;21(4):578-82.
8. Forrest C, McTigue K, Hernandez A, Cohen L, Cruz H, Haynes K, et al. PCORnet 2020: current state, accomplishments, and future directions. Submitted for publication. 2020.
9. Greene DN, Jackson ML, Hillyard DR, Delgado JC, Schmidt RL. Decreasing median age of COVID-19 cases in the United States-Changing epidemiology or changing surveillance? *PloS one*. 2020;15(10):e0240783.
10. Qualls LG, Phillips TA, Hammill BG, Topping J, Louzao DM, Brown JS, et al. Evaluating Foundational Data Quality in the National Patient-Centered Clinical Research Network (PCORnet(R)). *EGEMS (Washington, DC)*. 2018;6(1):3.
11. International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) 2020 [Available from: <https://www.cdc.gov/nchs/icd/icd10cm.htm>].
12. Fu L, Wang B, Yuan T, Chen X, Ao Y, Fitzpatrick T, et al. Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: A systematic review and meta-analysis. *The Journal of infection*. 2020.
13. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England journal of medicine*. 2020;382(18):1708-20.
14. Liang WH, Guan WJ, Li CC, Li YM, Liang HR, Zhao Y, et al. Clinical characteristics and outcomes of hospitalised patients with COVID-19 treated in Hubei (epicenter) and outside Hubei (non-epicenter): A Nationwide Analysis of China. *The European respiratory journal*. 2020.
15. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *Jama*. 2020.
16. Muniyappa R, Gubbi S. COVID-19 pandemic, coronaviruses, and diabetes mellitus. *American journal of physiology Endocrinology and metabolism*. 2020;318(5):E736-e41.

- 920 17. Barron E, Bakhai C, Kar P, Weaver A, Bradley D, Ismail H, et al. Associations of type 1  
921 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *The lancet Diabetes & endocrinology*. 2020;8(10):813-22.
- 922 18. Holman N, Knighton P, Kar P, O'Keefe J, Curley M, Weaver A, et al. Risk factors for  
923 COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-  
924 based cohort study. *The lancet Diabetes & endocrinology*. 2020;8(10):823-33.
- 925 19. Navaratnam AV, Gray WK, Day J, Wendon J, Briggs TWR. Patient factors and temporal  
926 trends associated with COVID-19 in-hospital mortality in England: an observational study using  
927 administrative data. *The Lancet Respiratory medicine*. 2021;9(4):397-406.
- 928 20. Helms J, Kremer S, Merdji H, Clere-Jehl R, Schenck M, Kummerlen C, et al. Neurologic  
929 Features in Severe SARS-CoV-2 Infection. *The New England journal of medicine*. 2020.
- 930 21. Weiland NL, S. Zimmer, C. U.S. calls for pause on Johnson & Johnson vaccine after  
931 clotting cases. *New York Times*. 2021 13 April 2021.
- 932 22. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Guinney J, Kibbe WA, et al. The  
933 National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment.  
934 *Journal of the American Medical Informatics Association : JAMIA*. 2021;28(3):427-43.
- 935

936

937

938 Table 1: Characteristics of patients positive for SARS-CoV-2 on polymerase chain reaction or antigen laboratory tests, March 1, 2020  
 939 – January 30, 2022, 43 PCORnet Data Contributing Institutions

	All SARS-CoV-2 +			SARS-CoV-2 +, Ambulatory			SARS-CoV-2 +, Emergency Department			SARS-CoV-2 +, Inpatient		
	1,325,563			499,370			202,241			172,673		
	N	%	RR	N	%	RR	N	%	RR	N	%	RR
<b>Age (years)</b>												
20-<40	540,033	41%	Ref	221,535	44%	Ref	85,360	42%	Ref	31,125	18%	Ref
40-<55	346,074	26%	1.0	138,259	28%	0.9	53,762	27%	1.0	33,280	19%	1.7
55-<65	204,153	15%	0.8	74,873	15%	0.6	29,682	15%	0.9	34,186	20%	1.8
65-<75	135,503	10%	0.6	43,829	9%	0.5	18,926	9%	0.7	34,580	20%	1.7
75-<85	70,592	5%	0.6	17,017	3%	0.4	10,162	5%	0.6	25,748	15%	1.9
85+	29,208	2%	0.7	3,857	1%	0.4	4,349	2%	0.5	13,754	8%	2.2

940

<b>Sex</b>												
<i>Female</i>	736,587	56%	Ref	284,413	57%	Ref	112,905	56%	Ref	86,107	50%	Ref
<i>Male</i>	588,459	44%	1.1	214,851	43%	1.0	89,312	44%	1.0	86,545	50%	1.4
<i>Other<sup>6</sup></i>	511	0%	*	102	0%	*	24	0%	*	20	0%	*
<b>Race</b>												
<i>Asian</i>	37,423	3%	0.9	13,153	3%	0.9	4,935	2%	1.0	5,594	3%	1.2
<i>Black or African American</i>	214,540	16%	1.1	62,218	12%	1.0	56,796	28%	1.2	39,384	23%	1.5
<i>White</i>	831,180	63%	Ref	343,881	69%	Ref	100,859	50%	Ref	91,793	53%	Ref
<i>Other</i>	126,993	10%	1.1	38,029	8%	1.1	32,338	16%	1.3	26,716	15%	1.8
<i>Missing</i>	115,427	9%	1.0	42,089	8%	1.2	7,313	4%	1.1	9,186	5%	1.5

941



<b>Hispanic Ethnicity</b>												
<i>Yes</i>	217,931	16%	1.5	69,250	14%	1.5	45,620	23%	1.5	33,254	19%	1.8
<i>No</i>	937,127	71%	Ref	370,700	74%	Ref	139,387	69%	Ref	126,686	73%	Ref
<i>Other/Missing</i>	170,505	13%	0.9	59,420	12%	0.9	17,234	9%	1.2	12,733	7%	1.2
<b>Month</b>												
<i>March 2020</i>	19,823	1%		4,121	1%		4,007	2%		7,157	4%	
<i>April 2020</i>	39,561	3%		10,117	2%		6,257	3%		13,462	8%	
<i>May 2020</i>	25,207	2%		9,416	2%		4,989	2%		4,849	3%	
<i>June 2020</i>	36,650	3%		16,635	3%		6,045	3%		4,559	3%	
<i>July 2020</i>	60,330	5%		28,146	6%		8,698	4%		7,299	4%	
<i>August 2020</i>	37,832	3%		18,200	4%		4,552	2%		4,915	3%	
<i>September 2020</i>	34,122	3%		19,109	4%		3,247	2%		3,897	2%	

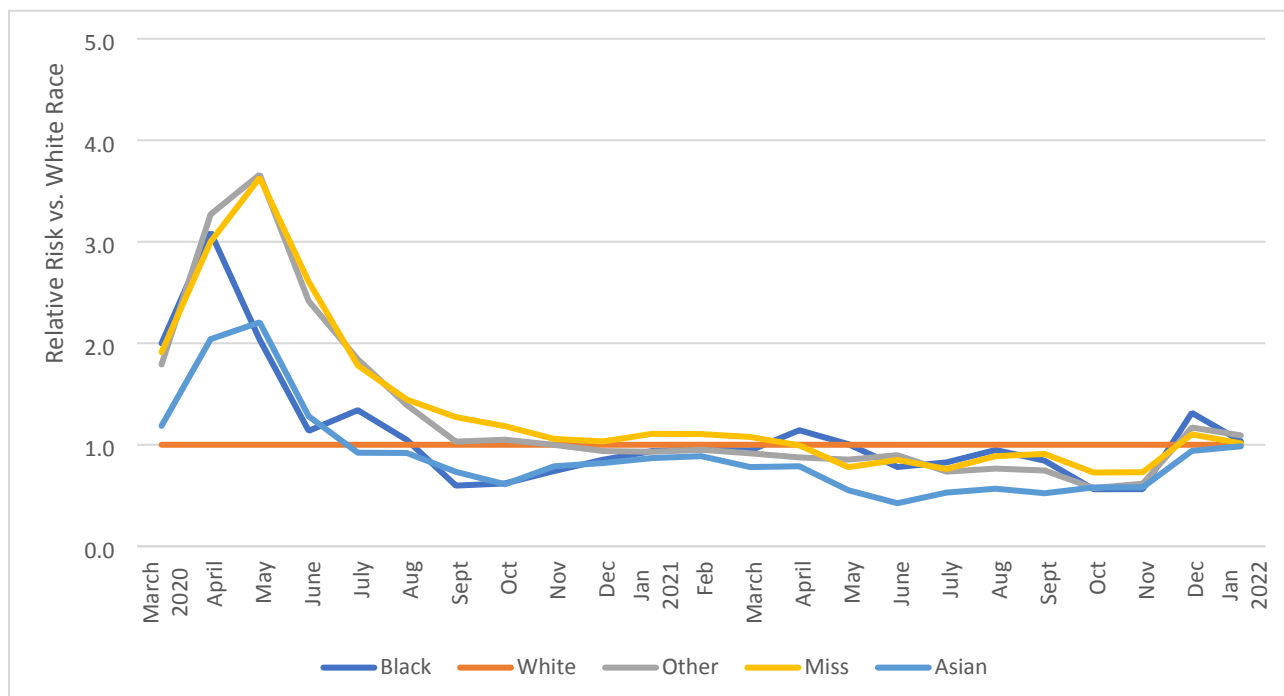
<i>October 2020</i>	60,474	5%		26,828	5%		6,386	3%		6,178	4%	
<i>November 2020</i>	125,513	9%		48,605	10%		13,789	7%		11,149	6%	
<i>December 2020</i>	131,017	10%		46,042	9%		14,255	7%		14,713	9%	
<i>January 2021</i>	97,096	7%		31,522	6%		12,466	6%		14,481	8%	
<i>February 2021</i>	39,816	3%		12,239	2%		5,749	3%		7,972	5%	
<i>March 2021</i>	32,802	2%		9,016	2%		5,905	3%		6,889	4%	
<i>April 2021</i>	32,097	2%		8,490	2%		6,735	3%		6,499	4%	
<i>May 2021</i>	15,190	1%		3,993	1%		3,631	2%		3,488	2%	
<i>June 2021</i>	8,414	1%		2,555	1%		1,815	1%		2,118	1%	
<i>July 2021</i>	22,919	2%		7,511	2%		6,776	3%		4,420	3%	
<i>August 2021</i>	62,668	5%		22,775	5%		14,955	7%		9,832	6%	
<i>September 2021</i>	52,779	4%		20,113	4%		10,079	5%		7,964	5%	

<i>October 2021</i>	33,004	2%		12,400	2%		5,478	3%		5,130	3%	
<i>November 2021</i>	35,651	3%		13,750	3%		6,220	3%		4,654	3%	
<i>December 2021</i>	109,629	8%		42,197	8%		21,346	11%		7,857	5%	
<i>January 2022</i>	212,969	16%		85,590	17%		28,861	14%		13,191	8%	

942

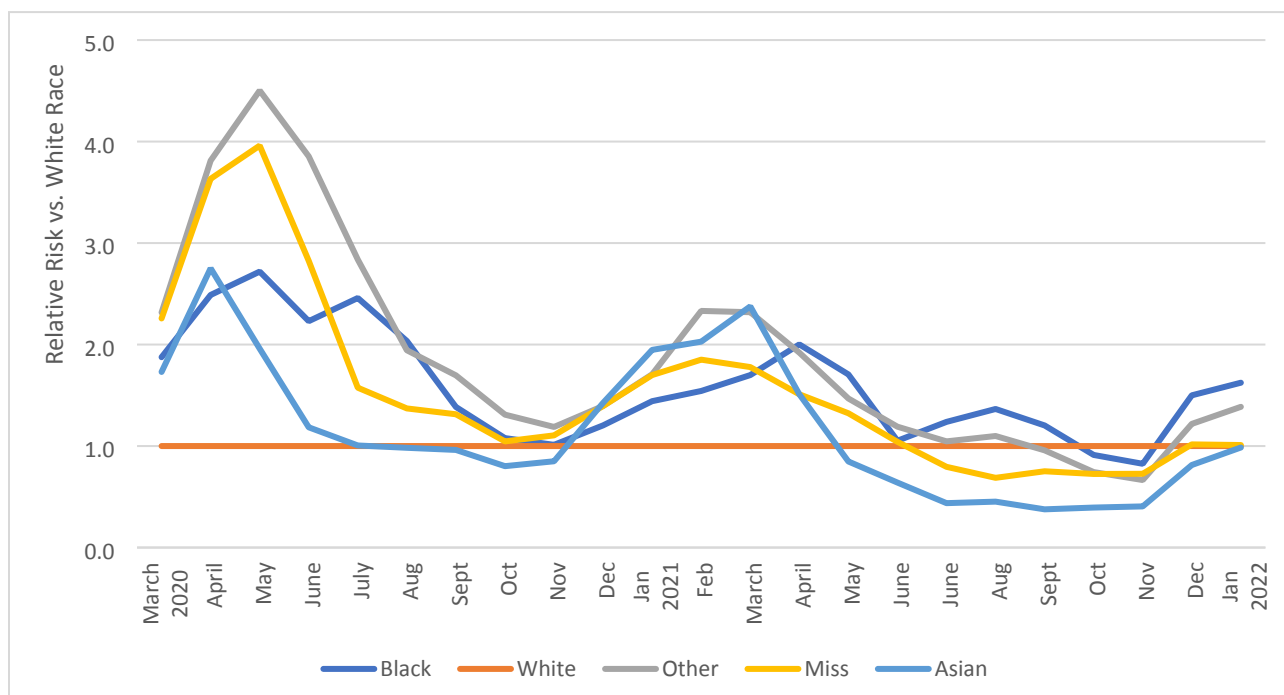
943 **Figure 1: Relative Risks by Race and Ethnicity of Patients with Positive SARS-CoV-2 Test**  
 944 **Result, by Care Setting and Month, March 2020—January 2022**

945 **A. Relative Risks by Race among SARS-CoV-2 Positive Patients in Ambulatory Setting**



946

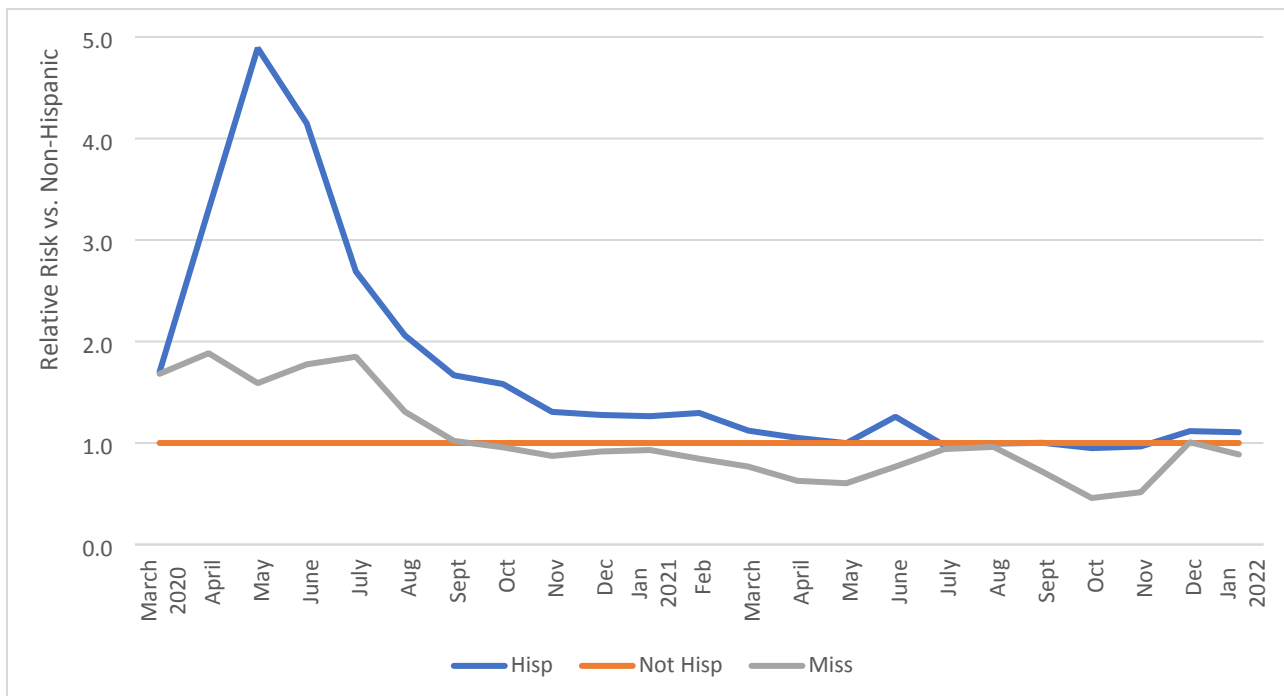
947 **B. Relative Risks by Race among SARS-CoV-2 Positive Patients in Inpatient Setting**



948

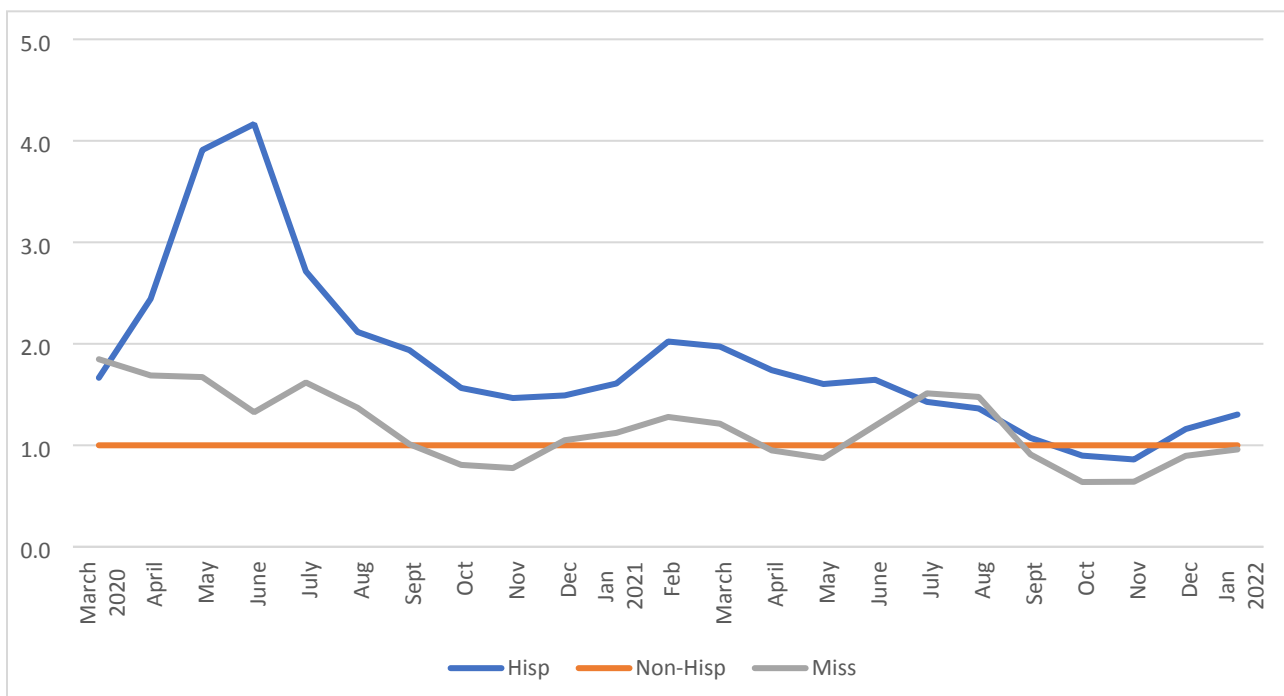
949

950 C. Ethnicity Distribution among SARS-CoV-2 Positive Patients in Ambulatory Setting



951

952 D. Ethnicity Distribution among SARS-CoV-2 Positive Patients in Inpatient Setting



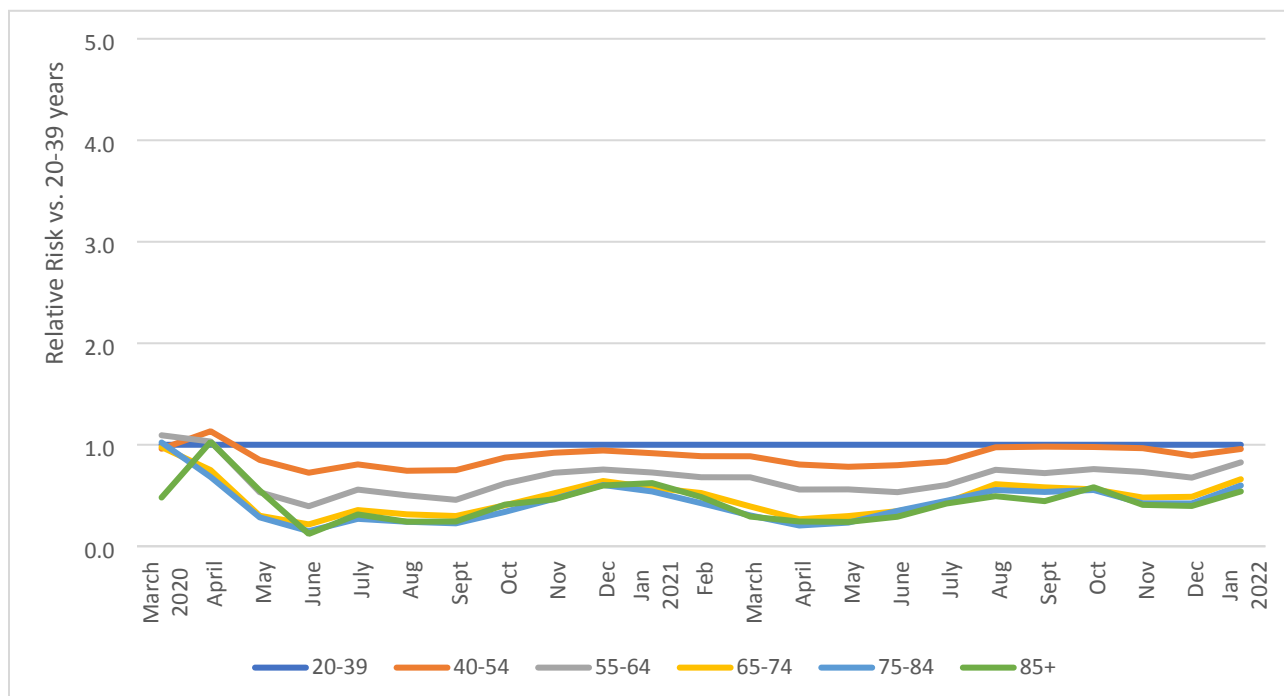
953

954

955

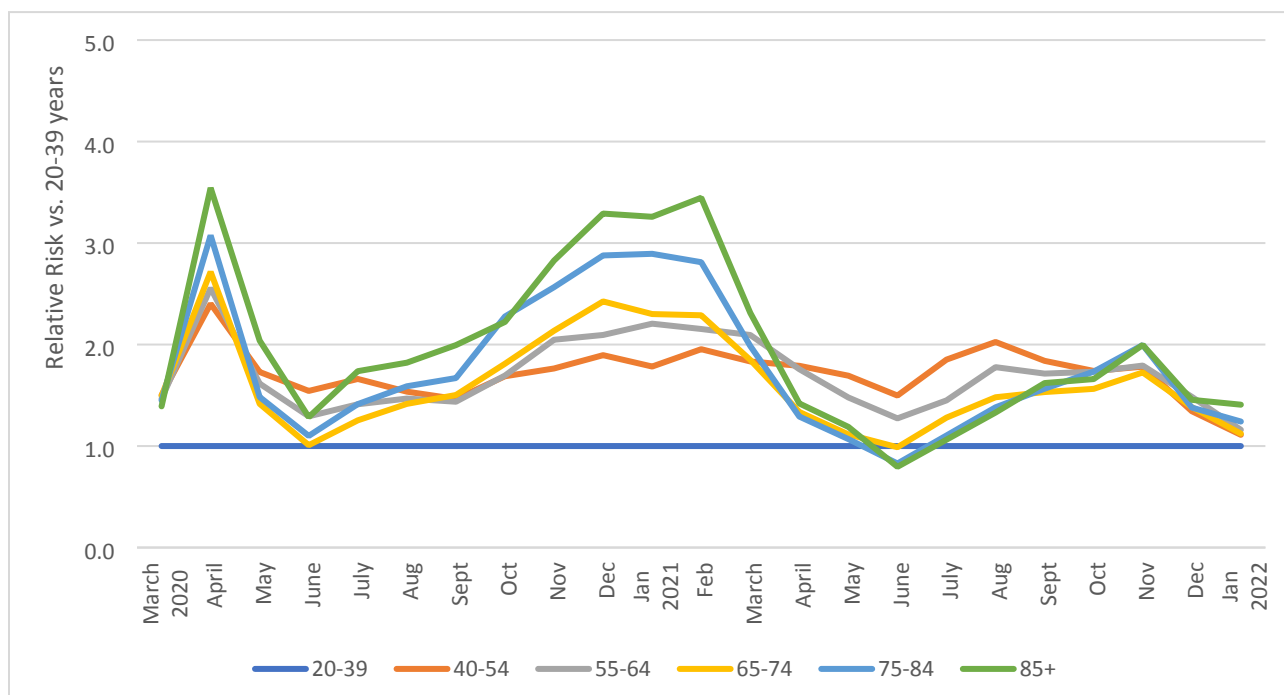
956 **Figure 2: Age Distribution of Patients with Positive SARS-CoV-2 Test Result, by Care**  
 957 **Setting and Month, March 2020—January 2022**

958 A. Age Distribution of SARS-CoV-2 Positive Patients in Ambulatory Setting



959

960 B. Age Distribution of SARS-CoV-2 Positive Patients in Inpatient Setting

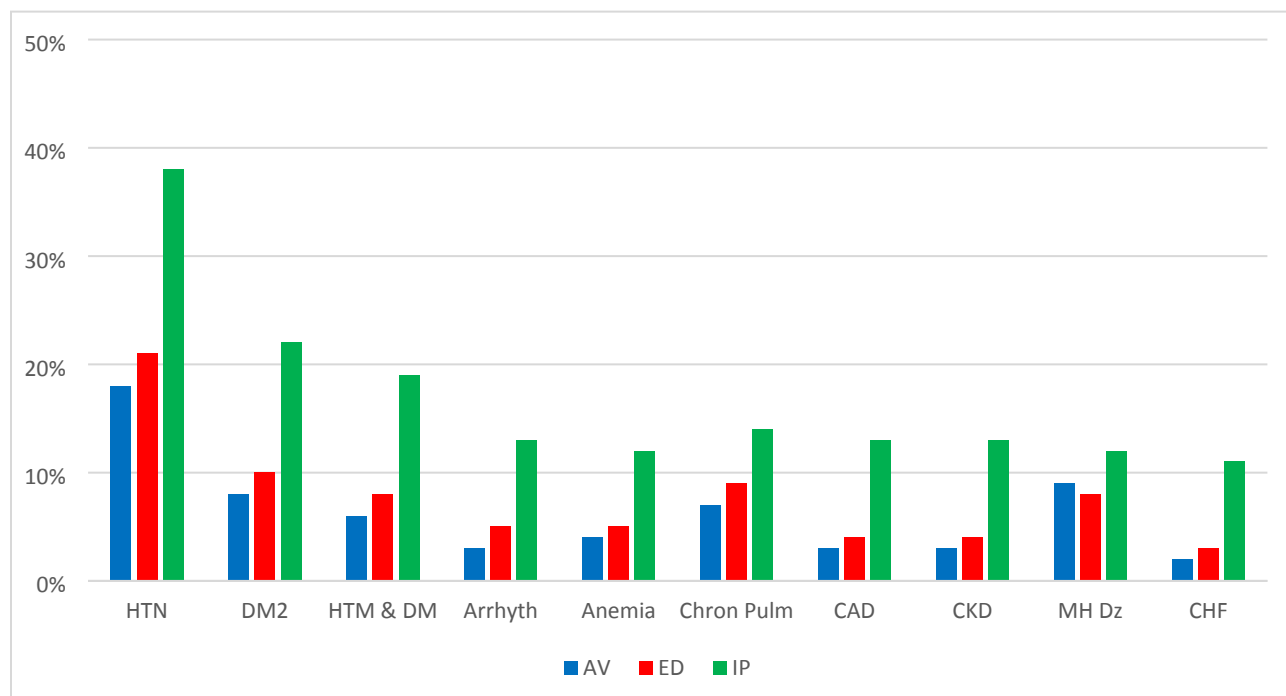


961

962

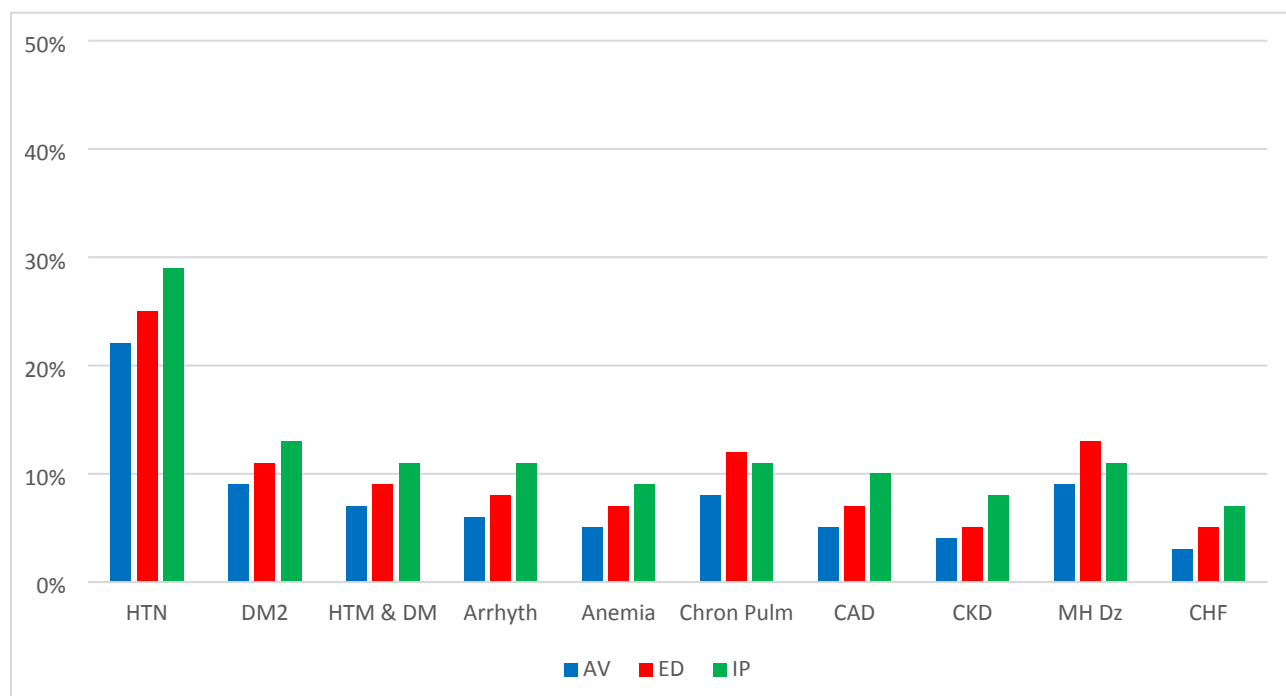
963 **Figure 3: Presence of Comorbidities among Patients Tested for SARS-CoV-2, by Care Setting and**  
 964 **Month, March 2020—January 2022**

965 A. Patients Testing Positive for SARS-CoV-2



966

967 B. Patients Testing Negative for SARS-CoV-2



968

969

## 970 Captions for Supporting Information

971 **S Methods Note:** Describes process of code list generation, COVID-19 code-based case  
972 definitions, and sensitivity analyses for the assessment of comorbidities

973

974 **S1 Table: Responding PCORnet Healthcare Institutions**

975

976 **S2 Table: Characteristics of Patients Testing Negative for SARS-CoV-2 with No Positive**  
977 **Tests, March 1, 2020 to January 31, 2022**

978 **S3 Table: Comorbidities among Patients SARS-CoV-2 +, by Care Setting**

979

980 **S4 Table: Comorbidities among Patients SARS-CoV-2 -, by Care Setting**

981

982 **S Figures**

983

984 **S Fig 1. PCORnet Common Data Models, including standard and COVID-19 Common**  
985 **Data Model.**

986 The PCORnet Common Data Model includes information collected through various clinical data  
987 systems, including electronic health records (EHR) and health insurance claims or pharmacy  
988 dispensing records. Once data are entered by the user, systems transfer data to a data warehouse  
989 that is unique to their own health system and system vendor (e.g., Epic, Cerner). The typical  
990 process for populating the PCORnet Common Data Model is through an extraction-transfer-  
991 loading (ETL) procedure that allows for transformation of raw information into a form that is  
992 compliant with the Common Data Model specifications. This transformation, which is done by  
993 each individual institution in keeping with PCORnet's federated data infrastructure, allows for  
994 consistency across health systems and centralized querying of data, and the PCORnet ETL and  
995 Common Data Model data refresh occurs quarterly. The PCORnet COVID-19 Common Data  
996 Model utilizes a similar process but rather than including information on all patients, it includes  
997 information only on patients who have had at least one respiratory illness diagnostic code or a  
998 lab test ordered for SARS-CoV-2 since January 1, 2020. This process occurs frequently (as often  
999 as daily) and includes data with lag times between 1 day to 7 days, depending on the site.

1000

1001 **S Fig 2. Monthly Proportion of Patients Testing Positive for SARS-CoV-2 by Race and**  
1002 **Care Setting**

1003 These figures show the proportion of patients by month who were White, Black or African  
1004 American, Asian, Other Race or Missing Race in the (A) Ambulatory, (B) Emergency  
1005 Department, (C) Inpatient, and (D) Non-Emergency or Inpatient Setting (includes all patients  
1006 except for those who were directly connected to an emergency department or inpatient  
1007 encounter). We included (D) to account for the approximately 1/3 of patients who had no care  
1008 setting connected to their positive test; patients may not have had a care setting if they had a test  
1009 done in a drive-thru setting without an encounter or if the test was ordered in a virtual care  
1010 setting.

1011

1012



1013 **S Fig 3. Monthly relative risks for testing positive for SARS-CoV-2 by Race in the**  
1014 **Emergency Department, March 1, 2020 to January 31, 2022.**

1015 These figures show the relative risks for testing positive by month and racial group in the  
1016 emergency department; reference was White Race. Care settings are mutually exclusive.

1017  
1018 **S Fig 4. Monthly Proportion of Patients Testing Positive for SARS-CoV-2 by Ethnicity and**  
1019 **Care Setting**

1020 These figures show the proportion of patients by month who were Hispanic, Not Hispanic, or  
1021 had Missing Hispanic ethnicity information in the (A) Ambulatory, (B) Emergency Department,  
1022 (C) Inpatient, and (D) Non-Emergency or Inpatient Setting (includes all patients except for those  
1023 who were directly connected to an emergency department or inpatient encounter). We included  
1024 (B) to account for the approximately 1/3 of patients who had no care setting connected to their  
1025 positive test; patients may not have had a care setting if they had a test done in a drive-thru  
1026 setting without an encounter or if the test was ordered in a virtual care setting.

1027  
1028 **S Fig 5. Monthly relative risks for testing positive for SARS-CoV-2 by Ethnicity in the**  
1029 **Emergency Department, March 1, 2020 to January 31, 2022.**

1030 These figures show the relative risks for testing positive by month and ethnicity group in the  
1031 emergency department; reference was White Race. Care settings are mutually exclusive.

1032  
1033 **S Fig 6. Monthly Proportion of Patients Testing Positive for SARS-CoV-2 by Age and Care**  
1034 **SettingSARS-CoV-2 +, Patients in Emergency Departments and Patients in the Non-**  
1035 **Emergency or Inpatient Settings, Age by Month**

1036 These figures show the proportion of patients by month and age group in the (A) Ambulatory,  
1037 (B) Emergency Department, (C) Inpatient, and (D) Non-Emergency or Inpatient Setting  
1038 (includes all patients except for those who were directly connected to an emergency department  
1039 or inpatient encounter). We included (B) to account for the approximately 1/3 of patients who  
1040 had no care setting connected to their positive test; patients may not have had a care setting if  
1041 they had a test done in a drive-thru setting without an encounter or if the test was ordered in a  
1042 virtual care setting.

1043  
1044 **S Fig 7. Monthly relative risks for testing positive for SARS-CoV-2 by Age in the**  
1045 **Emergency Department, March 1, 2020 to January 31, 2022.**

1046 These figures show the relative risks for testing positive by month and age group in the  
1047 emergency department; reference was 20-39-year-olds. Care settings are mutually exclusive.

1048  
1049  
1050  
1051  
1052  
1053