medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

Understanding pediatric long COVID using a tree-based scan statistic approach: An EHR-based cohort study from the RECOVER Program

Corresponding author:

Vitaly Lorman
Applied Clinical Research Center
Children's Hospital of Philadelphia
2716 South St
Philadelphia, PA 19146
Lormanv@chop.edu
(267) 428-2845

Authors:

Vitaly Lorman, PhD*[1], Hanieh Razzaghi, MPH*[1], Suchitra Rao, MBBS, MSCS[2], Ravi Jhaveri, MD[3], Abigail Case, MD[4], Asuncion Mejias, MD, PhD, MsCS[5], Nathan M. Pajor, MD, MS[6], Payal Patel, MD[7], Deepika Thacker, MD[8], Seuli Bose-Brill, MD[9], Jason Block, MD, MPH[10], Patrick C. Hanley, MD, MSHQS [11], Priya Prahalad, MD, PhD[12], Yong Chen, PhD[13], Christopher B. Forrest, MD, PhD[1], L. Charles Bailey, MD, PhD[1], Grace M. Lee, MD, MPH[14]

*Co-first authorship

Affiliations:

[1]Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA, United States

[2]Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital of Colorado, Aurora, CO, United States

[3]Division of Infectious Diseases, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

[4]Division of Physical Medicine & Rehabilitation, The Children's Hospital of Philadelphia, Philadelphia, PA, United States

[5]Division of Infectious Diseases, Department of Pediatrics, Nationwide Children's Hospital and The Ohio State University, Columbus, OH, United States

[6]Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, OH, United States

[7]Department of Neurology, University of Washington, Seattle, WA, United States

[8]Nemours Cardiac Center, Nemours Childrens Health, Wilmington, DE, United States

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2022.12.08.22283158; this version posted December 11, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

[9]Internal Medicine and Pediatrics Section, Division of General Internal Medicine, Department of Internal Medicine, Ohio State University College of Medicine and Ohio State University Wexner Medical Center, Columbus, OH, United States

[10]Division of Chronic Disease Research Across the Lifecourse, Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, Boston, MA, United States

[11]Division of Endocrinology, Nemours Children's Hospital, Wilmington, DE, United States

[12]Department of Pediatrics, Division of Endocrinology, Stanford University, Stanford, CA, United States

[13]Department of Biostatistics, Epidemiology and Informatics, the University of Pennsylvania, Philadelphia, PA, United States

[14]Division of Infectious Diseases, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States

Keywords: Long COVID, Post-Acute Sequelae of SARS-CoV-2 Infection, COVID-19

Word count: 3854

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

STRUCTURED ABSTRACT

Objectives

Post-acute sequalae of SARS-CoV-2 infection (PASC) is not well defined in pediatrics given its heterogeneity of presentation and severity in this population. The aim of this study is to use novel methods that rely on data mining approaches rather than clinical experience to detect signals associated with PASC.

Materials and Methods

We used a propensity-matched cohort design comparing children identified using the new PASC ICD10CM diagnosis code (U09.9) (N=1250) to children with (N=6250) and without (N=6250) SARS-CoV-2 infection. We used a tree-based scan statistic to identify potential condition clusters co-occurring more frequently in cases than controls.

Results

We found significant enrichment among children with PASC in cardiac, respiratory, neurologic, psychological, endocrine, gastrointestinal, and musculoskeletal systems, the most significant related to circulatory and respiratory such as dyspnea, difficulty breathing, and fatigue and malaise.

Discussion

Our study addresses methodological limitations of prior studies that rely on pre-specified clusters of potential PASC-associated diagnoses driven by clinician experience. Future studies are needed to identify patterns of diagnoses and their associations to derive clinical phenotypes.

Conclusion

We identified multiple conditions and body systems associated with pediatric PASC. Because we rely on a data-driven approach, several new or under-reported conditions and symptoms were detected that warrant further investigation.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

BACKGROUND AND SIGNIFICANCE

The National Institutions of Health (NIH) launched the new RECOVER initiative in 2021 (1) to leverage electronic health record (EHR) data to better identify and characterize patients with post-acute sequelae of SARS-CoV-2 infection (PASC), defined by the NIH as failure to recover from COVID-19, or those persistently symptomatic for >30 days.(2) In the adult literature, advances have been made to predict PASC among COVID-19 affected patients (3) and to describe etiology, risk factors, and outcomes (4-7). In contrast, there is currently a paucity of rigorous studies that accurately describe PASC in children (8). Research attempting to better elucidate PASC in children has been limited by small sample sizes (9-11), lack of a control group (12-14), or restrictions in source population (15) that limit generalizability to a broader cohort. (17)

Access to EHR data in a large population of children offers an opportunity to better understand the spectrum of PASC across a wide range of demographics and clinical trajectories. A recent exploratory analysis using EHR data to characterize pediatric PASC examined symptoms, diagnoses, and medications occurring more frequently in a large cohort of SARS-CoV-2 viral test-positive patients when compared with SARS-CoV-2-negative controls. Similar studies have also been carried out in adult populations(7). However, these studies are limited in two important ways. First, SARS-CoV-2 positive patients are heterogeneous and therefore analyses are at risk for spurious associations attributable to residual confounding rather than significant clinical findings. Second and relatedly, studies of PASC-associated features employ outcomes defined by clusters of clinically similar diagnosis codes. Clustering clinical codes in this way may bias findings that confirm clinical experience. The rapidly evolving nature of the pandemic and the lack of consensus of specific symptoms that define PASC in children necessitates a more data-driven approach for knowledge discovery.(18)

In this study, we explored syndromic and systemic features associated with a clinical diagnosis of PASC compared to children with and without SARS-CoV-2 infection. The diagnosis code for PASC was established in October 2021 – *U09.9, post COVID-19 condition, unspecified*. Prior to this, the non-specific code, *B94.8, Sequelae of other specified infectious and parasitic diseases*, was proposed as a temporary alternative (19). While these codes reflect clinician judgment in diagnosing patients who may suffer from PASC, they are likely to have a higher positive predictive value than identifying cases using COVID-19- alone. To identify clusters of PASC-associated diagnoses from tens of thousands of diagnosis codes, we used a tree-based scan statistic, a data mining tool which detects signals from hierarchical structures without relying on pre-specifying the clusters of codes of interest. This data-driven technique is especially advantageous in EHR research where mining a large corpus of data is feasible.

MATERIALS AND METHODS

Study Population and Design

The RECOVER PEDSnet EHR population includes EHR data from nine children's hospitals: Children's Hospital of Philadelphia, Cincinnati Children's Hospital Medical Center, Children's Hospital of Colorado, Ann & Robert H. Lurie Children's Hospital of Chicago, Nationwide

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Children's Hospital, Nemours Children's Health System (in Delaware and Florida), Seattle Children's Hospital, and Stanford Children's Health. The Children's Hospital of Philadelphia's institutional review board designated this study as not human subjects research and the need for consent was waived. The PEDSnet COVID-19 Database Version week 115 was used. 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/.

Our primary analyses focused on two comparisons of interest. Cases consisted of patients with evidence of PASC; the two comparator cohorts were patients with and without SARS-CoV-2 infection. Evidence for PASC comprised a U09.9 diagnosis code, an interface terminology (IMO) term in the EHR with the following strings: ('post' and 'acute' and 'covid') or ('complication' and 'covid'), or a B94.8 diagnosis code. Because the B94.8 code is more general, we did not include such diagnoses where the IMO term indicated something other than PASC. We excluded patients who had a MIS-C diagnosis at any point in time based on the presence of the M35.81, U10, and U10.9 ICD10CM diagnostic codes; therefore, we use 'PASC' to refer specifically to non-MIS-C variants. Patients were considered *SARS-COV-2 infected* if they had a diagnostic test that was positive for infection (Polymerase Chain Reaction (PCR), antigen, or serology). Serology tests included IgM, IgG anti- N antibodies, IgG anti-S or receptor binding domain (RBD) antibodies, and IgG and IgA undifferentiated antibodies based on criteria from Mejias et al. for positivity due to infection(20). Patients were also considered SARS-COV-2 infected if they had a diagnosis code for COVID-19 in the inpatient or emergency department (ED) setting. We excluded diagnosis codes for COVID-19 in the outpatient setting because of the possibility that the code was assigned as a rule-out diagnosis during a patient evaluation. Further, the observation period in the study aligned with the phase of the pandemic when viral testing was still widely performed in the healthcare setting. Patients were considered *SARS-COV-2 uninfected* if (1) all available diagnostic tests (PCR, Ag, serology) during the study period were negative and (2) the patient did not have any diagnosis codes indicating SARS-COV-2 infection, MIS-C, or PASC in any setting.

Cohort entry date for incident PASC cases was defined as: (1) the date of the first positive PCR or antigen test, or (2) 4 weeks before the date of the first positive serology test, or (3) 4 weeks before the first occurrence of a diagnosis of PASC if no preceding confirmatory test was available. We recognize that imposing a date of infection may misclassify the risk window between initial infection and PASC diagnosis if no testing data is available but chose to be overly inclusive of patients rather than exclude patients with the diagnosis based on unavailability of testing data. The cohort entry date for non-PASC SARS-CoV-2 infected patients was defined based on the date of the earliest SARS-CoV-2 diagnostic test (unless serology positive, then 4 weeks earlier) or earliest COVID-19 diagnosis/encounter. For SARS-CoV-2 uninfected patients, cohort entry dates were chosen as the date of a random negative test. All patients across case and control definitions were < 21 years of age at the date of cohort entry.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Additionally, we conducted a sensitivity analysis in which we compared SARS-CoV-2 positive to SARS-CoV-2 negative patients during the post-acute period—a description of the methods and results of this analysis are included in the supplementary appendix.

Matching

We extracted data for our study from the RECOVER pediatric database by identifying any patient who met study inclusion criteria between March 1, 2020, and June 22, 2022. We used a propensity score matched approach to balance covariates between cases and each of the comparison groups to minimize potential confounding (21). Propensity scores were estimated using logistic regression as the probability of a PASC diagnosis conditional on the following covariates: institution, age group (<1, 1-4, 5-11, 12-15, 16-20 years), sex, race/ethnicity (white non-Hispanic (NH), Black NH, Hispanic, Asian NH, multi-racial NH, other NH), and clinical setting of testing. Additionally, we used the Pediatric Medical Complexity Algorithm (PMCA) (22) to define for each body system an index indicating presence and complexity of chronic condition in that body system grouping and used these PMCA indices in our propensity score model. We then matched PASC patients to comparison patients (SARS-CoV-2 infected patients in the first cohort, SARS-COV-2 uninfected patients in the second cohort) using 5:1 nearest neighbor matching, additionally requiring exact matching on both cohort entry month and age group (23). We assessed the balance between the SARS-CoV-2 infected and uninfected patients via absolute standardized mean differences.

Tree-based scan statistic

The tree-based scan statistic is a data mining tool that is well established in vaccine safety(24) and disease surveillance and characterization(24, 25), making it salient for applications to understand pediatric PASC in its early stages of characterization. This approach simultaneously evaluates outcomes at different levels of granularity in a hierarchical structure, adjusting for multiple testing using a likelihood ratio statistic. In this study, we employed the unconditional analysis based on a Bernoulli probability model at each node of the tree, as described in the TreeScan User's Guide(26). This approach identifies branches of the tree at which outcomes belonged most disproportionately to cases as compared to controls.

As inputs, we used the ICD10CM vocabulary following the ICD10CM Tabular List of Diseases and Injuries(27) as developed by the National Center for Health Statistics (NCHS). For example, a path down the hierarchy might proceed as follows: C00-D49-> C00-C97 -> C44 -> C44.1 -> C44.10 -> C44.102 -> C44.1021. In all, the hierarchy has 7 levels, with the following counts of nodes per level starting from the top down: 22, 262, 2274, 9502, 14400, 22236, 48450.

We will alternately refer to the hierarchy as a tree, and the cluster consisting of a node together with all of its descendants will be referred to as a *branch* of the tree, or equivalently, as a *cut.* For each branch of the tree, we started with the observed count of cases who had an incident diagnosis in that branch during the outcome period. We also calculated the expected number of cases for each branch by multiplying the proportion of cases in the cohort (1/6 due to 5:1 matching) by the number of controls who had an incident diagnosis in the branch. The null hypothesis, for each cut, is that the number of observed cases equals the number of expected

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

cases. This method adjusts for multiple testing in the sense that under the null hypothesis that the number of observed cases equals the number of expected cases at each cut, there is a 95% probability that there is no cut in the tree with a p-value <0.05. For each cut demonstrating significance at the p<0.05 threshold, we calculated percent excess cases as the difference between number of observed cases and the number of expected cases divided by the number of expected cases. Analyses were carried out using R (4.1.0) and Treescan (2.0) software. Subclassification matching was implemented using the MatchIt R package and tree visualizations were built using the collapsibleTree R package.

Outcomes

For each patient and each diagnosis code, we calculated a binary indicator for whether the patient had an incident occurrence of that diagnosis code during the 28-179 days following cohort entry. We then used these to compute for each diagnosis code, and consequently each cluster of diagnosis codes defined by a branch of the tree, the number of cases and controls who had an incident diagnosis in that branch.

To define incident outcomes, we employed a washout period between 18 months and 7 days prior to cohort entry for conditions grouped at the 3rd level of the ICD hierarchy (I.e., the part of the ICD code before the decimal place) to ensure new incident diagnoses. In other words, conditions which occurred during the study period were not counted if a code in the same group occurred during the washout period. We also a priori excluded a handful of outcomes that were deemed unlikely to be attributable to COVID-19, specifically injury codes (S00-T98), external causes of morbidity and mortality (V01-Y98), and congenital malformations, deformations and chromosomal abnormalities (Q00-Q99).

RESULTS

Study population and matching

Between March 1, 2020, and June 22, 2022, there was a total of 13,750 patients identified for inclusion into the 3 cohorts, with PASC cases accounting for 1,250 patients (Table 1a). Older children and females were more likely to be included in the PASC cohort than younger children and males(overall: 55.2% vs 44.8%). The most common cohort entry months were in the fall of 2021. A majority of children in the PASC cohort (55.1%) did not have any chronic conditions. Following 5:1 nearest neighbor matching, both comparison cohorts had excellent balance (|SMD|<0.2 for unknown test location and one site, |SMD|<0.1 for all other variables). (Supplementary Appendix Figures 1a-b). After matching, the demographic distributions of both COVID-positive and COVID-negative cohorts reflect those of the PASC cohort.

TreeScan

When comparing patients with PASC to patients who were SARS-CoV-2 infected, we identified multiple statistical signals. At the highest level of the tree, significant cuts included R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified, M00-M99 Diseases of the musculoskeletal system and connective tissue, G00-G99 Diseases of the nervous system, J00-J99 Diseases of the respiratory system, F00-F99 Mental and behavioral disorders, E00-E90 Endocrine, nutritional, and metabolic diseases, I00-I99 Diseases of the

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

circulatory system, Z00-Z99 Factors influencing health status and contact with health services, L00-L99 Diseases of the skin and subcutaneous tissue, and K00-K93 Diseases of the digestive system.

Within the R00-R99 branch, the top three cuts were R00-R09 Symptoms and signs involving the circulatory and respiratory systems, R50-R69 General symptoms and signs and R40-R46 Symptoms and signs involving cognition, perception, emotional state and behavior. Within the R00-R09 branch, the three most significant cuts included R06.0 Dyspnea, R07 Pain in throat and chest, R09.9 Chest pain, unspecified. This part of the hierarchy is visualized in Figure 3a.

The PASC vs SARS-CoV-2 negative comparison showed significance for 283 cuts, and largely overlapped with the PASC vs SARS-CoV-2 negative comparison at the top of the tree, with the addition of H00-H59 Diseases of the eye and adnexa.

We refer the reader to Table 2 for a summary of systemic and syndromic findings collected from the two comparisons and Tables 3a and 3b for a full list of significant cuts, the likelihood ratio statistics, p-values, and percent excess cases.

DISCUSSION

We have employed a tree-based data mining approach to detect new-onset conditions presenting in children with a PASC (U09.9 or B94.8) diagnosis. Our data-driven approach identified and ranked features in a large corpus of EHR data. In the first of our two primary analyses, we compared PASC-diagnosed patients with non-PASC-diagnosed SARS-CoV-2 positive patients. The signals identified indicate long term sequelae of PASC that differentiate from having COVID-19.

Significant results from this comparison included both systemic and syndromic manifestations of PASC, with the former describing organ-level dysfunction and the latter focusing on signs and symptoms. Consistent with other studies and biologic mechanisms, we found significant enrichment among children with PASC in cardiac, respiratory, neurologic, psychological, endocrine, gastrointestinal, and musculoskeletal systems(28) .

We found the most significant point of difference between PASC and COVID-19 positive children was at the R00-R99 cluster of diagnostic codes (*Symptoms, signs, and abnormal clinical and laboratory findings, not otherwise classified*). These codes designate abnormal clinical findings not elsewhere described that encompass a range of organ systems. This is not surprising given the heterogeneity of previously reported manifestations of PASC and the difficulty in gaining clinical consensus to define PASC(29, 30).

Within the R00-R99 category, the codes with the greatest signal are symptoms related to circulatory and respiratory systems. These results align again with prior observations in smaller groups of patients or previous studies, and include dyspnea(12, 17), shortness of breath(17, 32), and other breathing abnormalities(33); chest pain(10, 34, 35); malaise and fatigue(14, 32,

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

35, 36); cough(14); and abnormalities of heartbeat including palpitations and tachycardia(37, 38).

Neurologic diagnoses such as headache and migraine as well as sleep disturbances are significant cuts and have also been previously reported (12, 33). Generalized pain was a key feature in our musculoskeletal system findings, which include non-localized myalgias, joint pain, and generalized soft tissue pain. The recurring appearance of multiple and diverse pain symptoms in this study warrants further investigation to better characterize causal biologic pathways.

Respiratory conditions were also significantly featured in our analysis. Both vocal cord disorders and dysautonomia were present as PASC-related features in our study. Interestingly, new onset asthma was present in our results as a specific diagnosis rather than a constellation of non-specific symptoms. Asthma is one of the few specific health conditions with a significant and strong signal in our study. Respiratory findings associated with PASC have been characterized in detail elsewhere (9).

In contrast to previous work, our study did not identify new-onset diabetes as a persistent long-term complication of COVID-19 (40). Endocrine findings in this study included nutritional deficiencies, volume depletion or fluid overload, and obesity. To our knowledge, these symptoms have not been described in detail elsewhere. It is unclear whether clinicians are recording obesity in the post-infection period as an indication that obesity complicated recovery of SARS-CoV-2 infection or that it points to a patient's exacerbation of health conditions unrelated to COVID-19.

Gastrointestinal (e.g., *Diseases of esophagus, stomach and duodenum*) and psychological findings (e.g., *Reaction to severe stress, and adjustment disorders*) followed a pattern of TreeScan cuts primarily at non-specific levels. Future work should focus on characterizing co-occurrences of these non-specific symptoms across body systems to detect clinically meaningful patterns.

We also compared children with a PASC diagnosis code to SARS-CoV-2 test-negative patients. The majority of cuts from this comparator cohort overlap with the findings where SARS-CoV-2 test-positive patients served as the control. These demonstrate consistency and increase the likelihood that these findings are truly associated with PASC. Additionally, several diagnoses unique to the test-negative cohort were identified, many of which include symptoms relating to acute respiratory infection. Some specific infection codes, such as B25, *Cytomegaloviral Disease* (CMV) were found significant, which may be a result of reactivation in patients infected with SARS-CoV-2(41). Other codes, such as D89, *Other disorders involving the immune mechanism, not elsewhere classified*, or skin conditions such as dermatitis and eczema (L20-L30), were also enriched in the PASC patients when compared to SARS-CoV-2 negative patients. Autoimmunity following an infection has been documented and warrants further investigation in the pediatric population(42). Alternate findings enriched in this analysis include features plausibly associated with COVID-19 infection including acute myocarditis, hypotension, fever, nasal congestion,

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

epistaxis, tachycardia, and gastrointestinal symptoms such as diarrhea, constipation, and gastroenteritis. More serious conditions such as pulmonary embolism and pulmonary edema were also enriched in this population but may be due to follow up care from an acute COVID-19 infection, and not a new finding in the post-acute period. However, these findings warrant further investigation. Findings such as sleep apnea, or behavioral codes associated with restlessness and agitation and disorientation, may also be plausibly associated with post-acute sequelae of COVID-19 infection.

Compared to previous work identifying pediatric PASC features by comparing SARS-CoV-2 test-positive to test-negative patients in the PEDSnet EHR, we found several differences. Abnormal liver enzymes, allergies, disorders of teeth/gingiva, and bronchiolitis were all reported as potential PASC-associated features in [18] but were not identified as significant cuts in this study when PASC patients were compared to COVID+ or COVID- cohorts, either at the specific code or macro level. Conversely, significant findings in our study which were not present in [18] included sleep disorders, anemias, eye disorders, constipation, sepsis, nutritional deficiencies, hypotension, and various musculoskeletal conditions and soft tissue and muscle disorders.

Finally, we conducted a sensitivity analysis comparing COVID positive to COVID negative patients. By identifying events that occur in the PASC risk window more frequently in COVID-positive patients, we can increase the sensitivity of capturing potential PASC features. This may come at the expense of precision, as we recognize that statistical signals identified in this analysis may correspond to features not truly associated with PASC, either due to confounding resulting from a more heterogeneous population than PASC-diagnosed patients or due to surveillance bias due to COVID-positivity. While there was significant overlap of the features identified in this and the primary analyses, there were also many additional features identified in the COVID-positive to COVID-negative analysis. This discrepancy may be due to either the increased power resulting from the larger cohort in the sensitivity analysis, or the result of confounding or surveillance bias described above. Further investigation of these signals is warranted.

To our knowledge, our study is the first to employ a TreeScan approach in children to detect groups of clinical diagnoses specific to this population, as well as the first to make use of the U09.9 and B94.8 diagnosis codes to do so. Our study is strengthened by multiple comparison groups diagnoses of PASC rather than surveys or case series analyses. We have confirmed previous findings and provide new insights into how diagnosis clusters relate to each other in importance. Further, an advantage of EHR data is the ability to learn from previous clinical history of patients. In our analysis, we included an 18-month washout period to ensure that the signals we were observing were not due to baseline differences in disease prevalence.
Prior work identifying pediatric PASC-associated features [18] relied on a set of diagnosis clusters developed by the study team to group similar diagnoses. Our study addresses this methodological limitation by a broader approach where diagnosis clusters are tested along all levels of the ICD-10-CM hierarchy. Importantly, among features of PASC found in previous work, our study contributes precise, diagnosis code-level results showing how PASC features

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

are identified in clinical practice. This makes our findings ideally suited for development of rules-based phenotypes for identifying pediatric PASC in EHRs.

At the same time, there were several significant limitations to our study. First, we are at risk of misclassifying patients in our PASC-negative comparator cohort by relying solely on the U09.9 and B94.8 codes to identify cases. There are likely children and adolescents in our comparison cohort with complications of COVID-19 who never receive a formal PASC diagnosis and may either present to a health system with symptoms that the clinician did not distinguish as a complication of COVID-19 or may not receive care for their symptoms at all. However, this would likely bias toward the null. Further, the U09.9 code has only been formally implemented since October 2021, and while the B94.8 code likely served as a proxy for PASC prior to this, the likely lack of uniform code usage, particularly early in the pandemic, introduces potential bias with regards to the timing of cohort entry dates in our cohort during the pandemic. Additionally, for PASC-diagnosed and serology test positive patients in our cohort who had no prior COVID diagnoses or positive viral testing, our imputation of cohort entry dates may not accurately reflect timing of COVID infection, and in the case of PASC-diagnosed patients, PASC onset. Second, our ascertainment of other diagnosis codes is similarly subject to the biases of health care delivery. For example, our findings may disproportionately include diagnoses that are required for utilization (*e.g.,* to justify diagnostic testing) or symptoms that are uncommon and previously associated in clinical practice with COVID-19 and may under-detect changes following COVID-19 in relatively common childhood illness that escape coding practices. Third, we have analyzed the pandemic as a single cross-section, and further information may be gained by studying evolution of diagnoses over time since infection, or across different portions of the pandemic. Finally, the cohort is composed of patients followed in large pediatric health systems, and results may reflect the largely urban demographics of these systems or the practice patterns of pediatric specialists.

CONCLUSION

This systematic evaluation of diagnoses in a large, multi-system national cohort of children with PASC provides a valuable addition to our understanding of this condition's protean manifestations. Our results can inform the design of further prospective studies to more deeply investigate the patterns identified here to develop improved clinical practice and to direct the study of the biological basis of PASC.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Table 1. Study population characteristics after matching

	Level	Overall N (%)	SARS CoV-2-negative N (%)	SARS CoV-2-positive N (%)	PASC N (%)
n		13,750	6,250	6,250	1,250
Age at cohort entrance (%)	<1	440 (3.2)	200 (3.2)	200 (3.2)	40 (3.2)
	1-4	1276 (9.3)	580 (9.3)	580 (9.3)	116 (9.3)
	5-11	3542 (25.8)	1610 (25.8)	1610 (25.8)	322 (25.8)
	12-15	4620 (33.6)	2100 (33.6)	2100 (33.6)	420 (33.6)
	16-20	3872 (28.2)	1760 (28.2)	1760 (28.2)	352 (28.2)
Sex (%)	Female	7577 (55.1)	3400 (54.4)	3487 (55.8)	690 (55.2)
	Male	6173 (44.9)	2850 (45.6)	2763 (44.2)	560 (44.8)
Race/ethnicity (%)	Non-Hispanic Asian/PI	422 (3.1)	186 (3.0)	196 (3.1)	40 (3.2)
	Non-Hispanic Black/AA	1255 (9.1)	555 (8.9)	583 (9.3)	117 (9.4)
	Hispanic	2017 (14.7)	905 (14.5)	928 (14.8)	184 (14.7)
	Multiple	566 (4.1)	254 (4.1)	260 (4.2)	52 (4.2)
	Other/Unknown	700 (5.1)	316 (5.1)	312 (5.0)	72 (5.8)
	Non-Hispanic White	8790 (63.9)	4034 (64.5)	3971 (63.5)	785 (62.8)
Cohort entry month (%)	Mar 2020	77 (0.6)	35 (0.6)	35 (0.6)	7 (0.6)
	Apr 2020	132 (1.0)	60 (1.0)	60 (1.0)	12 (1.0)
	May 2020	88 (0.6)	40 (0.6)	40 (0.6)	8 (0.6)
	Jun 2020	165 (1.2)	75 (1.2)	75 (1.2)	15 (1.2)
	Jul 2020	154 (1.1)	70 (1.1)	70 (1.1)	14 (1.1)
	Aug 2020	110 (0.8)	50 (0.8)	50 (0.8)	10 (0.8)
	Sep 2020	231 (1.7)	105 (1.7)	105 (1.7)	21 (1.7)
	Oct 2020	242 (1.8)	110 (1.8)	110 (1.8)	22 (1.8)
	Nov 2020	385 (2.8)	175 (2.8)	175 (2.8)	35 (2.8)
	Dec 2020	396 (2.9)	180 (2.9)	180 (2.9)	36 (2.9)
	Jan 2021	506 (3.7)	230 (3.7)	230 (3.7)	46 (3.7)
	Feb 2021	264 (1.9)	120 (1.9)	120 (1.9)	24 (1.9)
	Mar 2021	451 (3.3)	205 (3.3)	205 (3.3)	41 (3.3)
	Apr 2021	473 (3.4)	215 (3.4)	215 (3.4)	43 (3.4)
	May 2021	341 (2.5)	155 (2.5)	155 (2.5)	31 (2.5)
	Jun 2021	495 (3.6)	225 (3.6)	225 (3.6)	45 (3.6)
	Jul 2021	517 (3.8)	235 (3.8)	235 (3.8)	47 (3.8)
	Aug 2021	792 (5.8)	360 (5.8)	360 (5.8)	72 (5.8)

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	Sep 2021	1474 (10.7)	670 (10.7)	670 (10.7)	134 (10.7)
	Oct 2021	1188 (8.6)	540 (8.6)	540 (8.6)	108 (8.6)
	Nov 2021	1122 (8.2)	510 (8.2)	510 (8.2)	102 (8.2)
	Dec 2021	1958 (14.2)	890 (14.2)	890 (14.2)	178 (14.2)
	Jan 2022	2189 (15.9)	995 (15.9)	995 (15.9)	199 (15.9)
Institution (%)	A	1039 (7.6)	518 (8.3)	437 (7.0)	84 (6.7)
	B	1565 (11.4)	692 (11.1)	715 (11.4)	158 (12.6)
	C	2553 (18.6)	1191 (19.1)	1125 (18.0)	237 (19.0)
	D	397 (2.9)	177 (2.8)	183 (2.9)	37 (3.0)
	E	1048 (7.6)	437 (7.0)	513 (8.2)	98 (7.8)
	F	546 (4.0)	262 (4.2)	224 (3.6)	60 (4.8)
	G	1307 (9.5)	594 (9.5)	600 (9.6)	113 (9.0)
	H	2385 (17.3)	1041 (16.7)	1121 (17.9)	223 (17.8)
	I	2910 (21.2)	1338 (21.4)	1332 (21.3)	240 (19.2)
Test Location (%)	Emergency Department	2179 (15.8)	969 (15.5)	1032 (16.5)	178 (14.2)
	Inpatient	1835 (13.3)	782 (12.5)	891 (14.3)	162 (13.0)
	Other/Unknown	80 (0.6)	37 (0.6)	10 (0.2)	33 (2.6)
	Outpatient Office	6019 (43.8)	2792 (44.7)	2671 (42.7)	556 (44.5)
	Outpatient: Test Only	3637 (26.5)	1670 (26.7)	1646 (26.3)	321 (25.7)
PMCA Index (%)	0	7436 (54.1)	3370 (53.9)	3377 (54.0)	689 (55.1)
	1	4339 (31.6)	1948 (31.2)	2013 (32.2)	378 (30.2)
	2	1975 (14.4)	932 (14.9)	860 (13.8)	183 (14.6)

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Table 2: Summary of significant features in the two comparison cohorts

Major systemic findings, or conditions (general and specific)	
Diseases of the nervous system	migraines, headache syndromes, sleep disorders, insomnia, chronic pain, dysautonomia, post-viral fatigue syndrome, encephalitis, disorders of autonomic nervous system, sleep apnea*, transverse myelitis*
Mental and behavioral disorders	anxiety disorders, severe stress and adjustment disorders, organic including symptomatic mental disorders, obsessive-compulsive disorder, dissociative and conversion disorders*
Diseases of the respiratory system	asthma, ARDS/respiratory failure, diseases of vocal cords and larynx, tonsilitis, exercise induced bronchospasm*, chronic sinusitis*, pulmonary edema*,
Diseases of the circulatory system	arrhythmias, cerebrovascular disease, thromboembolic disease, acute myocarditis*, hypotension*, pulmonary embolism*, tachycardia*,
Disease of the musculoskeletal system and connective tissue	pain in joints, pain in limbs, dorsalgia, myalgias, hypermobility syndrome, soft tissue disorders, muscle disorders, arthropathies, osteopathies, disorders of bone, reactive arthritis, physeal arrest
Diseases of the digestive system	Diseases of oesophagus, stomach, and duodenum, aphagia and dysphagia
Diseases of blood	purpura and other hemorrhagic conditions*, anemias*, sickle cell disorders*
Endocrine, nutritional and metabolic diseases	nutritional deficiencies, obesity and overweight, disorders of mineral metabolism, volume depletion or fluid overload, obesity
Renal	Acute kidney failure*
Diseases of the eye	Disorders of refraction and accommodation*
Immunologic findings	Other disorders involving the immune mechanism not elsewhere classified, cytomegaloviral disease*
Major syndromic findings, or symptoms	
Circulatory and respiratory signs and symptoms	Palpitations, cough, dyspnea, shortness of breath, chest pain, viral pneumonia, asphyxia and hypoxemia*, hyperglycemia*
GI signs/symptoms	abdominal pain, nausea/vomiting, diarrhea*, constipation*, gastroenteritis*
CNS/musculoskeletal signs/symptoms	lack of coordination, abnormalities of gait and mobility
Cognition, perception,	dizziness, smell/taste disturbances, restlessness and agitation*,

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

emotional state	disorientation*, shock*
General signs and symptoms	Fatigue/malaise, headache, weakness, syncope and collapse, pain not elsewhere classified, fever*, enlarged lymph node, nasal congestion*, epistaxis*, noninflammatory disorders of female genital tract*
Skin	Symptoms and signs involving the skin and subcutaneous tissue, dermatitis,* eczema,* urticaria* and erythema*

*Findings significant in PASC vs COVID-negative comparison but not PASC vs COVID-positive comparison

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 3a: Significant branches of the R00-R09 block of the ICD hierarchy with nodes colored by percent excess cases

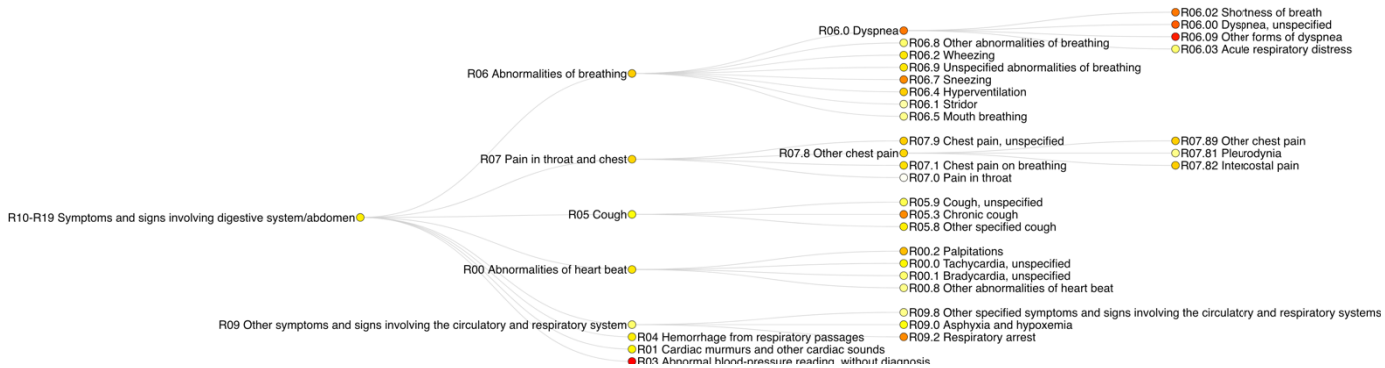

Figure 3b: Significant branches of the R10-R19 block of the ICD hierarchy with nodes colored by percent excess cases

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Table 3a: PASC vs COVID positive comparison TreeScan results

Cut	Node	Tree Level	Log Likelihood Ratio	P value	N cases observed	N cases expected	Percent excess cases
A00-B99 Certain infections and parasitic diseases							
6	B90-B94 Sequelae of infectious and parasitic diseases	2	327.89	0.001	183	30.50	500.00
7	B94 Sequelae of other and unspecified infectious and parasitic diseases	3	327.89	0.001	183	30.50	500.00
8	B94.8 Sequelae of other specified infectious and parasitic diseases	4	326.10	0.001	182	30.33	500.07
14	A00-B99 Certain infections and parasitic diseases	1	157.15	0.001	1,006	167.67	141.55
139	B95-B98 Bacterial, viral and other infectious agents	2	11.37	0.002	259	43.17	71.42
148	B97 Viral agents as the cause of diseases classified elsewhere	3	10.81	0.005	120	20.00	105.00
149	A65-A69 Other spirochaetal diseases	2	10.68	0.009	16	2.67	311.99
158	B25-B34 Other viral diseases	2	10.00	0.014	214	35.67	73.82
D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism							
33	D89.89 Other specified disorders involving the immune mechanism, not elsewhere classified	5	55.21	0.001	44	7.33	432.06
49	D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	1	39.94	0.001	901	150.17	71.81
56	D89 Other disorders involving the immune	3	35.01	0.001	93	15.50	229.03

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	mechanism, not elsewhere classified						
58	D89.8 Other specified disorders involving the immune mechanism, not elsewhere classified	4	32.12	0.001	83	13.83	232.61
72	D80-D89 Certain disorders involving the immune mechanism	2	23.01	0.001	287	47.83	98.62
E00-E90 Endocrine, nutritional and metabolic diseases							
30	E00-E90 Endocrine, nutritional and metabolic diseases	1	60.96	0.001	1,480	246.67	69.05
51	E70-E88 Metabolic disorders	2	39.53	0.001	631	105.17	86.36
112	E87 Other disorders of fluid, electrolyte and acid-base balance	3	15.36	0.001	194	32.33	97.96
113	E65-E68 Obesity and other hyperalimentation	2	15.28	0.001	182	30.33	101.12
119	E66 Overweight and obesity	3	14.12	0.001	179	29.83	97.79
131	E83 Disorders of mineral metabolism	3	12.30	0.001	142	23.67	102.79
141	E86 Volume depletion	3	11.31	0.002	106	17.67	115.05
145	E86.0 Dehydration	4	11.10	0.005	99	16.50	118.18
151	E50-E64 Other nutritional deficiencies	2	10.67	0.009	137	22.83	97.11
F00-F99 Mental and behavioural disorders							
28	F00-F99 Mental and behavioural disorders	1	68.24	0.001	1,776	296.00	66.55
52	F40-F48 Neurotic, stress-related and somatoform disorders	2	39.33	0.001	543	90.50	93.37
88	F41.9 Anxiety disorder, unspecified	4	19.15	0.001	194	32.33	110.33
89	F41 Other anxiety disorders	3	19.05	0.001	264	44.00	93.18
90	F80-F89 Disorders of psychological development	2	18.82	0.001	422	70.33	72.05
144	F00-F09 Organic, including symptomatic,	2	11.11	0.004	35	5.83	208.75

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	mental disorders						
	G00-G99 Diseases of the nervous system						
15	G00-G99 Diseases of the nervous system	1	155.02	0.001	2,238	373.00	91.15
29	G89-G99 Other disorders of the nervous system	2	66.03	0.001	519	86.50	126.59
32	G40-G47 Episodic and paroxysmal disorders	2	59.22	0.001	1,365	227.50	70.99
46	G89.29 Other chronic pain	5	40.51	0.001	201	33.50	162.69
48	G89.2 Chronic pain, not elsewhere classified	4	40.12	0.001	202	33.67	161.36
50	G89 Pain, not elsewhere classified	3	39.59	0.001	268	44.67	137.30
54	G00-G09 Inflammatory diseases of the central nervous system	2	36.23	0.001	61	10.17	293.31
57	G43 Migraine	3	32.23	0.001	301	50.17	115.27
64	G04 Encephalitis, myelitis and encephalomyelitis	3	25.52	0.001	40	6.67	304.80
68	G47 Sleep disorders	3	24.39	0.001	481	80.17	77.12
69	G43.0 Migraine without aura	4	23.67	0.001	116	19.33	163.84
71	G44 Other headache syndromes	3	23.21	0.001	87	14.50	189.66
74	G44.5 Complicated headache syndromes	4	22.55	0.001	22	3.67	390.46
77	G90.9 Disorder of the autonomic nervous system, unspecified	4	20.93	0.001	19	3.17	404.73
79	G90 Disorders of autonomic nervous system	3	20.55	0.001	38	6.33	279.15
82	G04.8 Other encephalitis, myelitis and encephalomyelitis	4	19.85	0.001	24	4.00	350.00
86	G44.52 New daily persistent headache (NDPH)	5	19.31	0.001	18	3.00	400.00
94	G93.3 Postviral fatigue syndrome	4	18.16	0.001	13	2.17	453.00
111	G43.009 Migraine	6	15.37	0.001	55	9.17	194.44

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	without aura, not intractable, without status migrainosus						
114	G93 Other disorders of brain	3	15.25	0.001	174	29.00	103.45
118	G43.00 Migraine without aura, not intractable	5	14.26	0.001	64	10.67	171.79
133	G04.81 Other encephalitis and encephalomyelitis	5	12.11	0.001	17	2.83	324.03
166	G47.9 Sleep disorder, unspecified	4	9.47	0.023	74	12.33	127.09
167	G43.01 Migraine without aura, intractable	5	9.46	0.023	52	8.67	153.75
I00-I99 Diseases of the circulatory system							
35	I00-I99 Diseases of the circulatory system	1	53.46	0.001	948	158.00	81.65
41	I30-I52 Other forms of heart disease	2	45.42	0.001	335	55.83	131.06
87	I49 Other cardiac arrhythmias	3	19.26	0.001	51	8.50	229.41
122	I49.8 Other specified cardiac arrhythmias	4	13.45	0.001	34	5.67	235.10
132	I51 Complications and ill-defined descriptions of heart disease	3	12.12	0.001	59	9.83	164.50
J00-J99 Diseases of the respiratory system							
18	J00-J99 Diseases of the respiratory system	1	99.31	0.001	2,514	419.00	67.54
45	J40-J47 Chronic lower respiratory diseases	2	42.07	0.001	669	111.50	86.55
47	J45 Asthma	3	40.37	0.001	648	108.00	86.11
70	J30-J39 Other diseases of upper respiratory tract	2	23.64	0.001	459	76.50	77.78
75	J45.9 Other and unspecified asthma	4	21.90	0.001	203	33.83	115.78
80	J96-J99 Other diseases of the respiratory system	2	20.15	0.001	374	62.33	79.69
109	J35 Chronic diseases of tonsils and adenoids	3	15.52	0.001	78	13.00	161.54
115	J45.909 Unspecified asthma, uncomplicated	6	14.48	0.001	153	25.50	107.84
121	J45.90 Unspecified	5	13.60	0.001	169	28.17	98.79

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	asthma						
123	J96.0 Acute respiratory failure	4	13.36	0.001	95	15.83	133.73
125	J12.8 Other viral pneumonia	4	12.87	0.001	32	5.33	237.71
129	J45.4 Moderate persistent asthma	4	12.36	0.001	106	17.67	120.71
143	J03 Acute tonsillitis	3	11.17	0.004	26	4.33	246.42
147	J45.40 Moderate persistent asthma, uncomplicated	5	11.00	0.005	84	14.00	128.57
150	J80-J84 Other respiratory diseases principally affecting the interstitium	2	10.67	0.009	39	6.50	192.31
152	J96.01 Acute respiratory failure with hypoxia	5	10.63	0.009	49	8.17	169.28
153	J98 Other respiratory disorders	3	10.57	0.010	150	25.00	92.00
154	J20-J22 Other acute lower respiratory infections	2	10.47	0.011	46	7.67	173.79
157	J96 Respiratory failure, not elsewhere classified	3	10.30	0.012	221	36.83	73.77
159	J45.99 Other asthma	5	9.99	0.014	34	5.67	199.82
165	J12 Viral pneumonia, not elsewhere classified	3	9.48	0.023	45	7.50	166.67
	K00-K93 Diseases of the digestive system						
135	K00-K93 Diseases of the digestive system	1	11.86	0.001	1,776	296.00	26.69
137	K20-K31 Diseases of oesophagus, stomach and duodenum	2	11.40	0.002	318	53.00	64.15
	L00-L99 Diseases of the skin and subcutaneous tissue						
127	L00-L99 Diseases of the skin and subcutaneous tissue	1	12.67	0.001	815	135.83	41.35
	M00-M99 Diseases of the musculoskeletal system and connective tissue						
13	M00-M99 Diseases of the musculoskeletal system and connective tissue	1	170.35	0.001	2,580	430.00	88.84
25	M20-M25 Other joint	2	68.87	0.001	837	139.50	100.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	disorders						
31	M25 Other joint disorder, not elsewhere classified	3	59.85	0.001	545	90.83	116.89
34	M79 Other and unspecified soft tissue disorders, not elsewhere classified	3	53.49	0.001	365	60.83	136.73
37	M70-M79 Other soft tissue disorders	2	50.12	0.001	391	65.17	127.10
39	M25.5 Pain in joint	4	46.27	0.001	393	65.50	121.37
43	M79.1 Myalgia	4	42.95	0.001	87	14.50	265.52
66	M25.50 Pain in unspecified joint	5	25.07	0.001	38	6.33	310.74
67	M79.10 Myalgia, unspecified site	5	24.76	0.001	46	7.67	278.10
81	M02.30 Reiter's disease, unspecified site	5	19.87	0.001	14	2.33	457.94
84	M30-M36 Systemic connective tissue disorders	2	19.49	0.001	82	13.67	177.98
85	M02.3 Reiter's disease	4	19.42	0.001	16	2.67	424.34
92	M79.18 Myalgia, other site	5	18.28	0.001	41	6.83	251.39
93	M35 Other systemic involvement of connective tissue	3	18.18	0.001	33	5.50	281.82
96	M35.7 Hypermobility syndrome	4	18.05	0.001	21	3.50	357.14
98	M02 Postinfective and reactive arthropathies	3	17.71	0.001	17	2.83	394.70
101	M00-M02 Infectious arthropathies	2	17.25	0.001	24	4.00	325.00
103	M79.6 Pain in limb, hand, foot, fingers and toes	4	16.64	0.001	238	39.67	91.58
104	M60-M63 Disorders of muscles	2	16.55	0.001	316	52.67	78.47
106	M86-M90 Other osteopathies	2	15.80	0.001	117	19.50	130.77
107	M62 Other disorders of muscle	3	15.74	0.001	312	52.00	76.92
128	M62.8 Other specified	4	12.60	0.001	278	46.33	72.67

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	disorders of muscle						
138	M89.1 Physeal arrest	4	11.38	0.002	9	1.50	433.33
140	M89 Other disorders of bone	3	11.37	0.002	44	7.33	186.49
146	M79.60 Pain in limb, unspecified	5	11.10	0.005	99	16.50	118.18
156	M50-M54 Other dorsopathies	2	10.37	0.012	243	40.50	70.37
170	M25.57 Pain in ankle and joints of foot	5	9.10	0.024	53	8.83	149.15
172	M54 Dorsalgia	3	9.00	0.025	225	37.50	68.00
R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified							
1	R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	1	696.14	0.001	7,440	1,240.00	107.18
9	R00-R09 Symptoms and signs involving the circulatory and respiratory systems	2	315.26	0.001	1,768	294.67	152.15
10	R06.0 Dyspnea	4	205.81	0.001	282	47.00	327.66
11	R50-R69 General symptoms and signs	2	187.76	0.001	1,975	329.17	108.10
12	R06 Abnormalities of breathing	3	172.25	0.001	564	94.00	204.26
16	R53 Malaise and fatigue	3	130.02	0.001	380	63.33	217.39
17	R53.8 Other malaise and fatigue	4	116.09	0.001	311	51.83	228.00
19	R06.00 Dyspnea, unspecified	5	91.08	0.001	100	16.67	367.91
20	R06.02 Shortness of breath	5	85.46	0.001	116	19.33	329.38
21	R53.83 Other fatigue	5	80.62	0.001	207	34.50	233.33
22	R40-R46 Symptoms and signs involving cognition, perception, emotional state and behaviour	2	77.91	0.001	665	110.83	121.06
23	R10-R19 Symptoms and signs involving the digestive system and abdomen	2	71.22	0.001	1,508	251.33	74.27

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

24	R07 Pain in throat and chest	3	69.87	0.001	280	46.67	182.84
36	R42 Dizziness and giddiness	3	51.34	0.001	137	22.83	228.52
38	R07.9 Chest pain, unspecified	4	46.88	0.001	154	25.67	203.86
40	R06.09 Other forms of dyspnea	5	45.94	0.001	34	5.67	446.74
42	R25-R29 Symptoms and signs involving the nervous and musculoskeletal systems	2	44.32	0.001	272	45.33	144.87
44	R00 Abnormalities of heart beat	3	42.38	0.001	217	36.17	159.88
53	R51 Headache	3	36.56	0.001	208	34.67	150.94
55	R51.9 Headache, unspecified	4	35.57	0.001	193	32.17	154.90
59	R10 Abdominal and pelvic pain	3	31.59	0.001	523	87.17	84.70
60	R00.2 Palpitations	4	30.93	0.001	82	13.67	229.19
61	R05 Cough	3	29.56	0.001	348	58.00	101.72
62	R53.82 Chronic fatigue, unspecified	5	27.98	0.001	33	5.50	354.55
63	R07.89 Other chest pain	5	25.79	0.001	85	14.17	203.46
65	R07.8 Other chest pain	4	25.20	0.001	99	16.50	184.85
73	R29 Other symptoms and signs involving the nervous and musculoskeletal systems	3	22.64	0.001	151	25.17	138.38
76	R05.3 Chronic cough	4	21.81	0.001	34	5.67	305.64
78	R68 Other general symptoms and signs	3	20.62	0.001	110	18.33	156.41
83	R43 Disturbances of smell and taste	3	19.53	0.001	34	5.67	288.01
91	R70-R79 Abnormal findings on examination of blood, without diagnosis	2	18.64	0.001	300	50.00	86.00
95	R11 Nausea and vomiting	3	18.05	0.001	427	71.17	70.02
97	R90-R96 Abnormal findings on diagnostic imaging and in function	2	17.75	0.001	279	46.50	87.10

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	studies, without diagnosis						
99	R68.89 Other general symptoms and signs	5	17.64	0.001	83	13.83	167.53
100	R55 Syncope and collapse	3	17.32	0.001	77	12.83	172.80
102	R68.8 Other general symptoms and signs	4	16.84	0.001	92	15.33	154.40
105	R43.9 Unspecified disturbances of smell and taste	4	16.12	0.001	16	2.67	386.89
108	R41 Other symptoms and signs involving cognitive functions and awareness	3	15.58	0.001	133	22.17	121.02
110	R53.1 Weakness	4	15.42	0.001	68	11.33	173.61
116	R00.0 Tachycardia, unspecified	4	14.44	0.001	99	16.50	136.36
124	R10.84 Generalized abdominal pain	5	13.03	0.001	147	24.50	104.08
126	R53.81 Other malaise	5	12.87	0.001	71	11.83	153.59
130	R27 Other lack of coordination	3	12.34	0.001	42	7.00	200.00
134	R10.8 Other abdominal pain	4	11.97	0.001	160	26.67	94.98
136	R09 Other symptoms and signs involving the circulatory and respiratory system	3	11.60	0.001	271	45.17	70.47
155	R05.9 Cough, unspecified	4	10.38	0.012	194	32.33	79.40
160	R11.10 Vomiting, unspecified	5	9.94	0.015	201	33.50	76.12
161	R46.81 Obsessive-compulsive behavior	5	9.71	0.020	8	1.33	426.32
162	R29.90 Unspecified symptoms and signs involving the nervous system	5	9.69	0.023	10	1.67	379.04
163	R13 Aphagia and dysphagia	3	9.54	0.023	267	44.50	64.04
164	R13.1 Dysphagia	4	9.54	0.023	267	44.50	64.04
168	R59.1 Generalized	4	9.28	0.023	15	2.50	300.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	enlarged lymph nodes						
169	R06.8 Other abnormalities of breathing	4	9.15	0.024	224	37.33	68.77
171	R41.8 Other symptoms and signs involving cognitive functions and awareness	4	9.09	0.024	99	16.50	106.06
U00-U95 Codes for special purposes							
2	U09 Post COVID-19 condition	3	492.73	0.001	275	45.83	500.04
3	U09.9 Post COVID-19 condition, unspecified	4	492.73	0.001	275	45.83	500.04
4	U00-U95 Codes for special purposes	1	408.47	0.001	630	105.00	307.62
5	U00-U49 Provisional assignment of new diseases of uncertain etiology or emergency use	2	408.47	0.001	630	105.00	307.62
26	U07.1 Emergency use of U07.1 | COVID-19	4	68.29	0.001	355	59.17	158.58
27	U07	3	68.29	0.001	355	59.17	158.58
Z00-Z99 Factors influencing health status and contact with health services							
117	Z00-Z99 Factors influencing health status and contact with health services	1	14.42	0.001	1,774	295.67	29.54
120	Z69-Z76 Persons encountering health services in other circumstances	2	13.64	0.001	254	42.33	79.54

Table 2b: PASC vs COVID negative comparison TreeScan results

Cut	Node	Tree Level	Log Likelihood	P value	N cases observed	N cases expected	Percent excess

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

			Ratio				cases
A00-B99 Certain infections and parasitic diseases							
7	B90-B94 Sequelae of infectious and parasitic diseases	2	327.89	0.001	183	30.50	500.00
8	B94 Sequelae of other and unspecified infectious and parasitic diseases	3	327.89	0.001	183	30.50	500.00
9	B94.8 Sequelae of other specified infectious and parasitic diseases	4	326.10	0.001	182	30.33	500.07
14	A00-B99 Certain infections and parasitic diseases	1	245.33	0.001	796	132.67	205.27
75	B95-B98 Bacterial, viral and other infectious agents	2	31.80	0.001	175	29.17	153.69
94	B97 Viral agents as the cause of diseases classified elsewhere	3	26.25	0.001	78	13.00	215.38
117	B25-B34 Other viral diseases	2	22.14	0.001	160	26.67	132.47
177	A30-A49 Other bacterial diseases	2	14.78	0.001	53	8.83	194.45
203	A65-A69 Other spirochaetal diseases	2	12.98	0.001	14	2.33	372.10
207	B97.89 Other viral agents as the cause of diseases classified elsewhere	5	12.54	0.002	48	8.00	187.50
216	B34 Viral infection of unspecified site	3	12.28	0.002	122	20.33	111.51
220	B97.8 Other viral agents as the cause of diseases classified elsewhere	4	12.08	0.002	49	8.17	181.52
232	B25 Cytomegaloviral disease	3	11.38	0.002	9	1.50	433.33
D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism							
23	D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the	1	124.54	0.001	575	95.83	169.23

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	immune mechanism						
40	D89.89 Other specified disorders involving the immune mechanism, not elsewhere classified	5	59.62	0.001	42	7.00	457.14
51	D80-D89 Certain disorders involving the immune mechanism	2	51.91	0.001	198	33.00	187.88
66	D89 Other disorders involving the immune mechanism, not elsewhere classified	3	41.12	0.001	84	14.00	264.29
67	D89.8 Other specified disorders involving the immune mechanism, not elsewhere classified	4	40.07	0.001	72	12.00	283.33
102	D60-D64 Aplastic and other anaemias	2	24.97	0.001	130	21.67	158.42
146	D65-D69 Coagulation defects, purpura and other haemorrhagic conditions	2	17.60	0.001	73	12.17	179.38
157	D55-D59 Haemolytic anaemias	2	16.08	0.001	33	5.50	263.64
161	D64 Other anemias	3	15.91	0.001	77	12.83	165.00
166	D70-D77 Other diseases of blood and blood-forming organs	2	15.70	0.001	81	13.50	159.26
184	D57 Sickle-cell disorders	3	14.39	0.001	22	3.67	308.72
202	D69 Purpura and other hemorrhagic conditions	3	13.02	0.001	47	7.83	193.74
214	D64.9 Anemia, unspecified	4	12.30	0.002	62	10.33	161.37
249	D57.4 Sickle-cell thalassemia	4	10.75	0.005	6	1.00	500.00
255	D84 Other immunodeficiencies	3	10.35	0.005	94	15.67	116.98
	E00-E90 Endocrine, nutritional and metabolic diseases						
25	E00-E90 Endocrine, nutritional and metabolic diseases	1	113.64	0.001	1,205	200.83	107.64
37	E70-E88 Metabolic disorders	2	72.43	0.001	498	83.00	136.14

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

62	E87 Other disorders of fluid, electrolyte and acid-base balance	3	43.15	0.001	118	19.67	225.37
105	E86 Volume depletion	3	24.49	0.001	72	12.00	216.67
114	E86.0 Dehydration	4	22.76	0.001	69	11.50	213.04
139	E65-E68 Obesity and other hyperalimentation	2	18.20	0.001	170	28.33	115.32
148	E66 Overweight and obesity	3	17.20	0.001	166	27.67	113.23
181	E83 Disorders of mineral metabolism	3	14.53	0.001	133	22.17	116.51
222	E50-E64 Other nutritional deficiencies	2	11.80	0.002	132	22.00	104.55
223	E66.9 Obesity, unspecified	4	11.75	0.002	67	11.17	150.67
244	E87.7 Fluid overload	4	11.08	0.002	32	5.33	218.95
254	E87.70 Fluid overload, unspecified	5	10.49	0.005	30	5.00	220.00
	F00-F99 Mental and behavioural disorders						
26	F00-F99 Mental and behavioural disorders	1	107.08	0.001	1,548	258.00	91.09
48	F40-F48 Neurotic, stress-related and somatoform disorders	2	53.47	0.001	483	80.50	117.39
87	F41 Other anxiety disorders	3	27.79	0.001	228	38.00	123.68
88	F41.9 Anxiety disorder, unspecified	4	27.37	0.001	166	27.67	145.75
91	F80-F89 Disorders of psychological development	2	26.88	0.001	377	62.83	92.58
182	F44 Dissociative and conversion disorders	3	14.49	0.001	13	2.17	406.91
217	F00-F09 Organic, including symptomatic, mental disorders	2	12.25	0.002	33	5.50	227.27
221	F90-F98 Behavioural and emotional disorders with onset usually occurring in childhood and adolescence	2	11.80	0.002	165	27.50	92.73
256	F88 Other disorders of psychological	3	10.28	0.005	143	23.83	93.03

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	development						
262	F42 Obsessive-compulsive disorder	3	10.05	0.007	22	3.67	254.22
G00-G99 Diseases of the nervous system							
16	G00-G99 Diseases of the nervous system	1	226.27	0.001	1,937	322.83	120.86
31	G40-G47 Episodic and paroxysmal disorders	2	93.82	0.001	1,177	196.17	98.30
33	G89-G99 Other disorders of the nervous system	2	81.51	0.001	471	78.50	149.68
44	G43 Migraine	3	58.59	0.001	226	37.67	186.70
54	G89 Pain, not elsewhere classified	3	49.64	0.001	240	40.00	165.00
55	G00-G09 Inflammatory diseases of the central nervous system	2	48.47	0.001	50	8.33	380.19
58	G89.29 Other chronic pain	5	45.09	0.001	190	31.67	177.87
59	G89.2 Chronic pain, not elsewhere classified	4	44.65	0.001	191	31.83	176.47
71	G44 Other headache syndromes	3	33.99	0.001	69	11.50	265.22
72	G04 Encephalitis, myelitis and encephalomyelitis	3	33.82	0.001	33	5.50	390.91
80	G47 Sleep disorders	3	30.02	0.001	450	75.00	89.33
81	G43.0 Migraine without aura	4	29.98	0.001	102	17.00	200.00
93	G90 Disorders of autonomic nervous system	3	26.47	0.001	32	5.33	350.28
107	G44.5 Complicated headache syndromes	4	24.19	0.001	21	3.50	414.29
116	G04.8 Other encephalitis, myelitis and encephalomyelitis	4	22.55	0.001	22	3.67	390.46
120	G93.3 Postviral fatigue syndrome	4	21.50	0.001	12	2.00	500.00
132	G44.52 New daily persistent headache (NDPH)	5	19.31	0.001	18	3.00	400.00
138	G93 Other disorders of brain	3	18.25	0.001	162	27.00	118.52
142	G90.9 Disorder of the	4	18.05	0.001	21	3.50	357.14

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	autonomic nervous system, unspecified						
162	G43.01 Migraine without aura, intractable	5	15.81	0.001	39	6.50	238.46
176	G43.009 Migraine without aura, not intractable, without status migrainosus	6	14.88	0.001	56	9.33	189.39
178	G43.00 Migraine without aura, not intractable	5	14.69	0.001	63	10.50	176.19
180	G04.81 Other encephalitis and encephalomyelitis	5	14.54	0.001	15	2.50	380.00
204	G44.8 Other specified headache syndromes	4	12.98	0.001	14	2.33	372.10
212	G47.3 Sleep apnea	4	12.32	0.002	253	42.17	75.48
235	G60-G65 Polyneuropathies and other disorders of the peripheral nervous system	2	11.33	0.002	23	3.83	265.54
237	G44.89 Other headache syndrome	5	11.27	0.002	11	1.83	391.80
241	G43.019 Migraine without aura, intractable, without status migrainosus	6	11.09	0.002	29	4.83	231.26
250	G37.3 Acute transverse myelitis in demyelinating disease of central nervous system	4	10.75	0.005	6	1.00	500.00
280	G43.9 Migraine, unspecified	4	9.16	0.014	39	6.50	176.92
H00-H59 Diseases of the eye and adnexa							
126	H00-H59 Diseases of the eye and adnexa	1	20.56	0.001	790	131.67	54.17
248	H49-H52 Disorders of ocular muscles, binocular movement, accommodation and refraction	2	10.81	0.002	341	56.83	60.13
260	H52 Disorders of refraction and	3	10.14	0.005	231	38.50	71.43

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	accommodation						
I00-I99 Diseases of the circulatory system							
29	I00-I99 Diseases of the circulatory system	1	102.68	0.001	740	123.33	132.71
39	I30-I52 Other forms of heart disease	2	60.44	0.001	292	48.67	165.05
110	I49 Other cardiac arrhythmias	3	23.44	0.001	45	7.50	273.33
141	I49.8 Other specified cardiac arrhythmias	4	18.10	0.001	28	4.67	306.85
154	I95-I99 Other and unspecified disorders of the circulatory system	2	16.34	0.001	50	8.33	212.12
172	I51 Complications and ill-defined descriptions of heart disease	3	15.28	0.001	52	8.67	199.88
173	I95 Hypotension	3	15.24	0.001	37	6.17	240.36
211	I26-I28 Pulmonary heart disease and diseases of pulmonary circulation	2	12.43	0.002	45	7.50	193.33
257	I40 Acute myocarditis	3	10.27	0.005	14	2.33	329.18
258	I26 Pulmonary embolism	3	10.27	0.005	14	2.33	329.18
J00-J99 Diseases of the respiratory system							
15	J00-J99 Diseases of the respiratory system	1	234.13	0.001	1,867	311.17	125.60
35	J40-J47 Chronic lower respiratory diseases	2	75.13	0.001	534	89.00	133.71
36	J45 Asthma	3	73.86	0.001	512	85.33	135.56
63	J30-J39 Other diseases of upper respiratory tract	2	43.01	0.001	370	61.67	120.53
68	J45.9 Other and unspecified asthma	4	39.96	0.001	152	25.33	188.20
70	J96-J99 Other diseases of the respiratory system	2	34.25	0.001	309	51.50	117.48
79	J45.909 Unspecified asthma, uncomplicated	6	30.15	0.001	108	18.00	194.44
85	J12.8 Other viral pneumonia	4	28.52	0.001	19	3.17	467.82
89	J96.0 Acute respiratory failure	4	26.97	0.001	65	10.83	241.64
92	J45.90 Unspecified asthma	5	26.54	0.001	126	21.00	166.67
96	J12 Viral pneumonia, not	3	25.75	0.001	24	4.00	400.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	elsewhere classified						
103	J00-J06 Acute upper respiratory infections	2	24.94	0.001	391	65.17	87.20
104	J09-J18 Influenza and pneumonia	2	24.74	0.001	110	18.33	172.78
111	J35 Chronic diseases of tonsils and adenoids	3	23.34	0.001	62	10.33	229.14
121	J12.82 Pneumonia due to coronavirus disease 2019	5	21.50	0.001	12	2.00	500.00
127	J80-J84 Other respiratory diseases principally affecting the interstitium	2	20.17	0.001	26	4.33	338.80
134	J98 Other respiratory disorders	3	19.04	0.001	118	19.67	144.03
136	J96.01 Acute respiratory failure with hypoxia	5	18.70	0.001	35	5.83	277.36
152	J45.4 Moderate persistent asthma	4	16.84	0.001	92	15.33	154.40
153	J03 Acute tonsillitis	3	16.54	0.001	20	3.33	350.45
155	J96 Respiratory failure, not elsewhere classified	3	16.25	0.001	190	31.67	102.08
170	J45.99 Other asthma	5	15.33	0.001	26	4.33	292.61
188	J45.40 Moderate persistent asthma, uncomplicated	5	14.00	0.001	75	12.50	156.00
200	J80 Acute respiratory distress syndrome	3	13.06	0.001	10	1.67	438.92
208	J45.990 Exercise induced bronchospasm	6	12.54	0.002	19	3.17	310.09
230	J45.30 Mild persistent asthma, uncomplicated	5	11.41	0.002	94	15.67	123.36
236	J45.3 Mild persistent asthma	4	11.30	0.002	110	18.33	112.77
242	J32 Chronic sinusitis	3	11.09	0.002	29	4.83	231.26
247	J20-J22 Other acute lower respiratory infections	2	10.91	0.002	45	7.50	180.00
265	J03.9 Acute tonsillitis, unspecified	4	9.92	0.008	12	2.00	350.00
269	J35.3 Hypertrophy of tonsils with hypertrophy of adenoids	4	9.77	0.008	17	2.83	288.69
270	J81 Pulmonary edema	3	9.71	0.010	8	1.33	426.32

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

271	J35.0 Chronic tonsillitis and adenoiditis	4	9.69	0.013	10	1.67	379.04
272	J35.01 Chronic tonsillitis	5	9.69	0.013	10	1.67	379.04
K00-K93 Diseases of the digestive system							
53	K00-K93 Diseases of the digestive system	1	49.78	0.001	1,366	227.67	64.71
125	K55-K64 Other diseases of intestines	2	20.89	0.001	485	80.83	70.73
147	K59 Other functional intestinal disorders	3	17.52	0.001	390	65.00	72.31
150	K59.0 Constipation	4	17.06	0.001	375	62.50	72.80
169	K20-K31 Diseases of oesophagus, stomach and duodenum	2	15.57	0.001	291	48.50	79.38
171	K59.00 Constipation, unspecified	5	15.28	0.001	211	35.17	93.35
198	K52 Other and unspecified noninfective gastroenteritis and colitis	3	13.24	0.001	70	11.67	157.07
213	K31.8 Other specified diseases of stomach and duodenum	4	12.30	0.002	30	5.00	240.00
263	K00-K14 Diseases of oral cavity, salivary glands and jaws	2	10.04	0.007	140	23.33	92.88
264	K52.9 Noninfective gastroenteritis and colitis, unspecified	4	10.03	0.007	54	9.00	155.56
L00-L99 Diseases of the skin and subcutaneous tissue							
82	L00-L99 Diseases of the skin and subcutaneous tissue	1	29.50	0.001	672	112.00	71.43
179	L49-L54 Urticaria and erythema	2	14.59	0.001	44	7.33	213.78
189	L20-L30 Dermatitis and eczema	2	13.89	0.001	201	33.50	91.04
261	L30 Other and unspecified dermatitis	3	10.06	0.005	47	7.83	168.20
279	L30.9 Dermatitis, unspecified	4	9.16	0.014	39	6.50	176.92
M00-M99 Diseases of the musculoskeletal system and connective tissue							
18	M00-M99 Diseases of the musculoskeletal system	1	172.11	0.001	2,571	428.50	89.50

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	and connective tissue						
38	M79 Other and unspecified soft tissue disorders, not elsewhere classified	3	61.70	0.001	341	56.83	153.39
41	M70-M79 Other soft tissue disorders	2	58.99	0.001	363	60.50	144.63
49	M25 Other joint disorder, not elsewhere classified	3	53.27	0.001	571	95.17	107.00
50	M20-M25 Other joint disorders	2	52.61	0.001	918	153.00	82.35
52	M25.5 Pain in joint	4	51.53	0.001	375	62.50	132.00
61	M79.1 Myalgia	4	43.72	0.001	86	14.33	269.85
83	M79.10 Myalgia, unspecified site	5	29.36	0.001	41	6.83	324.60
90	M86-M90 Other osteopathies	2	26.97	0.001	89	14.83	203.44
100	M02.3 Reiter's disease	4	25.08	0.001	14	2.33	500.86
101	M02 Postinfective and reactive arthropathies	3	25.08	0.001	14	2.33	500.86
106	M00-M02 Infectious arthropathies	2	24.43	0.001	19	3.17	436.28
112	M02.30 Reiter's disease, unspecified site	5	23.29	0.001	13	2.17	499.08
123	M79.6 Pain in limb, hand, foot, fingers and toes	4	21.35	0.001	217	36.17	110.12
128	M50-M54 Other dorsopathies	2	20.15	0.001	194	32.33	113.42
130	M54 Dorsalgia	3	19.76	0.001	172	28.67	119.74
133	M89 Other disorders of bone	3	19.04	0.001	32	5.33	294.00
140	M25.50 Pain in unspecified joint	5	18.10	0.001	47	7.83	232.06
151	M30-M36 Systemic connective tissue disorders	2	17.03	0.001	88	14.67	159.03
156	M60-M63 Disorders of muscles	2	16.21	0.001	318	53.00	77.36
165	M79.18 Myalgia, other site	5	15.74	0.001	45	7.50	220.00
168	M62 Other disorders of muscle	3	15.57	0.001	313	52.17	76.35
183	M79.60 Pain in limb,	5	14.45	0.001	88	14.67	145.40

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	unspecified						
185	M25.56 Pain in knee	5	14.35	0.001	103	17.17	132.96
194	M35 Other systemic involvement of connective tissue	3	13.42	0.001	40	6.67	214.84
199	M35.7 Hypermobility syndrome	4	13.17	0.001	26	4.33	269.52
201	M24.9 Joint derangement, unspecified	4	13.06	0.001	10	1.67	438.92
210	M62.8 Other specified disorders of muscle	4	12.44	0.002	279	46.50	72.04
233	M89.1 Physeal arrest	4	11.38	0.002	9	1.50	433.33
282	M25.561 Pain in right knee	6	9.08	0.017	46	7.67	160.76
N00-N99 Diseases of the genitourinary system							
137	N17 Acute kidney failure	3	18.45	0.001	38	6.33	263.35
158	N00-N99 Diseases of the genitourinary system	1	15.99	0.001	615	102.50	54.15
192	N17.9 Acute kidney failure, unspecified	4	13.45	0.001	34	5.67	235.10
227	N17-N19 Renal failure	2	11.55	0.002	64	10.67	153.05
259	N80-N98 Noninflammatory disorders of female genital tract	2	10.25	0.005	186	31.00	80.65
R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified							
1	R00-R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	1	1,145.85	0.001	5,963	993.83	158.49
4	R00-R09 Symptoms and signs involving the circulatory and respiratory systems	2	508.76	0.001	1,357	226.17	228.51
10	R50-R69 General symptoms and signs	2	299.48	0.001	1,606	267.67	155.91
13	R06.0 Dyspnea	4	245.38	0.001	250	41.67	382.36
17	R06 Abnormalities of breathing	3	214.14	0.001	494	82.33	247.38
19	R53 Malaise and fatigue	3	164.37	0.001	328	54.67	267.66
20	R53.8 Other malaise and	4	149.00	0.001	265	44.17	284.88

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	fatigue						
21	R40-R46 Symptoms and signs involving cognition, perception, emotional state and behaviour	2	135.61	0.001	507	84.50	189.94
22	R07 Pain in throat and chest	3	131.20	0.001	189	31.50	319.05
24	R10-R19 Symptoms and signs involving the digestive system and abdomen	2	122.47	0.001	1,253	208.83	109.74
27	R06.00 Dyspnea, unspecified	5	106.60	0.001	90	15.00	420.00
28	R06.02 Shortness of breath	5	104.66	0.001	101	16.83	393.17
30	R53.83 Other fatigue	5	96.11	0.001	185	30.83	273.01
32	R00 Abnormalities of heart beat	3	85.44	0.001	143	23.83	294.46
34	R07.9 Chest pain, unspecified	4	81.42	0.001	108	18.00	333.33
42	R51 Headache	3	58.90	0.001	160	26.67	226.21
43	R42 Dizziness and giddiness	3	58.64	0.001	126	21.00	257.14
45	R00.2 Palpitations	4	57.94	0.001	54	9.00	400.00
46	R51.9 Headache, unspecified	4	55.40	0.001	151	25.17	225.78
47	R05 Cough	3	53.96	0.001	267	44.50	162.92
56	R06.09 Other forms of dyspnea	5	48.36	0.001	33	5.50	463.64
57	R25-R29 Symptoms and signs involving the nervous and musculoskeletal systems	2	47.51	0.001	263	43.83	153.25
60	R07.8 Other chest pain	4	44.08	0.001	70	11.67	302.74
64	R07.89 Other chest pain	5	42.30	0.001	62	10.33	316.26
65	R11 Nausea and vomiting	3	41.47	0.001	318	53.00	128.30
69	R10 Abdominal and pelvic pain	3	38.64	0.001	488	81.33	97.96
73	R43 Disturbances of smell and taste	3	32.90	0.001	24	4.00	450.00
74	R50 Fever of other and unknown origin	3	32.54	0.001	166	27.67	160.21
76	R70-R79 Abnormal	2	31.70	0.001	245	40.83	127.77

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	findings on examination of blood, without diagnosis						
77	R09 Other symptoms and signs involving the circulatory and respiratory system	3	30.64	0.001	189	31.50	144.44
78	R00.0 Tachycardia, unspecified	4	30.15	0.001	66	11.00	254.55
84	R90-R96 Abnormal findings on diagnostic imaging and in function studies, without diagnosis	2	28.55	0.001	233	38.83	124.05
86	R53.81 Other malaise	5	27.84	0.001	45	7.50	300.00
95	R05.3 Chronic cough	4	26.19	0.001	30	5.00	360.00
97	R53.82 Chronic fatigue, unspecified	5	25.68	0.001	35	5.83	328.82
99	R41 Other symptoms and signs involving cognitive functions and awareness	3	25.46	0.001	105	17.50	180.00
108	R68 Other general symptoms and signs	3	23.85	0.001	102	17.00	176.47
109	R50.9 Fever, unspecified	4	23.46	0.001	127	21.17	155.08
113	R19 Other symptoms and signs involving the digestive system and abdomen	3	23.28	0.001	117	19.50	161.54
118	R05.9 Cough, unspecified	4	21.87	0.001	146	24.33	138.39
119	R11.10 Vomiting, unspecified	5	21.77	0.001	150	25.00	136.00
122	R29 Other symptoms and signs involving the nervous and musculoskeletal systems	3	21.37	0.001	155	25.83	132.29
124	R20-R23 Symptoms and signs involving the skin and subcutaneous tissue	2	20.91	0.001	195	32.50	115.38
129	R43.9 Unspecified disturbances of smell and taste	4	19.87	0.001	14	2.33	457.94
131	R55 Syncope and collapse	3	19.58	0.001	72	12.00	191.67

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

135	R68.8 Other general symptoms and signs	4	18.79	0.001	87	14.50	168.97
143	R68.89 Other general symptoms and signs	5	18.05	0.001	82	13.67	170.67
144	R11.1 Vomiting	4	17.92	0.001	179	29.83	111.20
145	R53.1 Weakness	4	17.72	0.001	63	10.50	195.24
149	R06.8 Other abnormalities of breathing	4	17.18	0.001	182	30.33	107.72
159	R09.02 Hypoxemia	5	15.96	0.001	57	9.50	194.74
160	R09.0 Asphyxia and hypoxemia	4	15.96	0.001	57	9.50	194.74
163	R10.8 Other abdominal pain	4	15.76	0.001	144	24.00	116.67
164	R19.7 Diarrhea, unspecified	4	15.74	0.001	74	12.33	167.64
167	R11.0 Nausea	4	15.58	0.001	71	11.83	170.50
174	R09.8 Other specified symptoms and signs involving the circulatory and respiratory systems	4	15.11	0.001	127	21.17	122.01
175	R10.84 Generalized abdominal pain	5	15.01	0.001	139	23.17	115.80
187	R29.90 Unspecified symptoms and signs involving the nervous system	5	14.33	0.001	8	1.33	501.50
190	R41.8 Other symptoms and signs involving cognitive functions and awareness	4	13.69	0.001	83	13.83	145.84
191	R63	3	13.68	0.001	232	38.67	83.60
193	R27 Other lack of coordination	3	13.42	0.001	40	6.67	214.84
195	R45 Symptoms and signs involving emotional state	3	13.36	0.001	114	19.00	121.05
196	R13.1 Dysphagia	4	13.27	0.001	243	40.50	80.25
197	R13 Aphagia and dysphagia	3	13.27	0.001	243	40.50	80.25
205	R29.9 Unspecified symptoms and signs involving the nervous and musculoskeletal	4	12.88	0.001	12	2.00	400.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	systems						
206	R45.1 Restlessness and agitation	4	12.88	0.001	12	2.00	400.00
209	R79 Other abnormal findings of blood chemistry	3	12.52	0.002	137	22.83	105.87
215	R04 Hemorrhage from respiratory passages	3	12.28	0.002	39	6.50	207.69
218	R59 Enlarged lymph nodes	3	12.25	0.002	36	6.00	216.67
224	R70	3	11.74	0.002	15	2.50	340.00
225	R91.8 Other nonspecific abnormal finding of lung field	4	11.73	0.002	28	4.67	242.61
226	R93 Abnormal findings on diagnostic imaging of other body structures	3	11.64	0.002	50	8.33	176.11
228	R46 Symptoms and signs involving appearance and behavior	3	11.53	0.002	86	14.33	130.29
231	R09.81 Nasal congestion	5	11.41	0.002	94	15.67	123.36
234	R94.31 Abnormal electrocardiogram [ECG] [EKG]	5	11.35	0.002	61	10.17	155.65
238	R41.0 Disorientation, unspecified	4	11.27	0.002	11	1.83	391.80
239	R47-R49 Symptoms and signs involving speech and voice	2	11.26	0.002	72	12.00	141.67
243	R91 Abnormal findings on diagnostic imaging of lung	3	11.08	0.002	32	5.33	218.95
245	R10.9 Unspecified abdominal pain	4	11.05	0.002	111	18.50	110.81
246	R94.3 Abnormal results of cardiovascular function studies	4	10.98	0.002	62	10.33	151.69
251	R46.8 Other symptoms and signs involving appearance and behavior	4	10.70	0.005	85	14.17	125.83
252	R20 Disturbances of skin sensation	3	10.67	0.005	39	6.50	192.31
253	R73.9 Hyperglycemia,	4	10.52	0.005	27	4.50	233.33

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	unspecified						
266	R57 Shock, not elsewhere classified	3	9.92	0.008	12	2.00	350.00
268	R30-R39 Symptoms and signs involving the urinary system	2	9.89	0.008	162	27.00	85.19
273	R94 Abnormal results of function studies	3	9.68	0.013	125	20.83	96.83
275	R26	3	9.50	0.013	20	3.33	260.36
277	R04.0 Epistaxis	4	9.40	0.013	32	5.33	200.19
278	R59.1 Generalized enlarged lymph nodes	4	9.28	0.014	15	2.50	300.00
281	R73 Elevated blood glucose level	3	9.09	0.016	64	10.67	134.30
283	R11.2 Nausea with vomiting, unspecified	4	9.01	0.018	68	11.33	129.48
U00-U95 Codes for special purposes							
2	U00-U49 Provisional assignment of new diseases of uncertain etiology or emergency use	2	749.53	0.001	431	71.83	495.85
3	U00-U95 Codes for special purposes	1	749.53	0.001	431	71.83	495.85
5	U09 Post COVID-19 condition	3	492.73	0.001	275	45.83	500.04
6	U09.9 Post COVID-19 condition, unspecified	4	492.73	0.001	275	45.83	500.04
11	U07.1 Emergency use of U07.1 | COVID-19	4	274.14	0.001	153	25.50	500.00
12	U07	3	259.86	0.001	156	26.00	488.46
Z00-Z99 Factors influencing health status and contact with health services							
98	Z00-Z99 Factors influencing health status and contact with health services	1	25.59	0.001	1,622	270.33	41.68
115	Z69-Z76 Persons encountering health services in other circumstances	2	22.63	0.001	212	35.33	115.11
229	Z89 Acquired absence of limb	3	11.44	0.002	13	2.17	360.83
240	Z99.81 Dependence on	5	11.17	0.002	38	6.33	200.16

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

	supplemental oxygen						

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

Acknowledgments

The authors gratefully acknowledge the contributions of Miranda Higginbotham.

Funding Source

This research was funded by the National Institutes of Health (NIH) Agreement OT2HL161847-01 as part of the Researching COVID to Enhance Recovery (RECOVER) program of research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the RECOVER Program, the NIH or other funders.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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. RECOVER: Researching COVID to Enhance Recovery.
2. Younger DS. Post-acute sequelae of SARS-CoV-2 infection (PASC): peripheral, autonomic, and central nervous system features in a child. Neurol Sci. 2021;42(10):3959-63. doi:10.1007/s10072-021-05345-5
3. Pfaff ER, Girvin AT, Bennett TD, Bhatia A, Brooks IM, Deer RR, et al. Identifying who has long COVID in the USA: a machine learning approach using N3C data. Lancet Digit Health. 2022;4(7):e532-e41. doi:10.1016/S2589-7500(22)00048-6
4. Munipalli B, Seim L, Dawson NL, Knight D, Dabrh AMA. Post-acute sequelae of COVID-19 (PASC): a meta-narrative review of pathophysiology, prevalence, and management. SN Compr Clin Med. 2022;4(1):90. doi:10.1007/s42399-022-01167-4
5. Reese J, Blau H, Bergquist T, Loomba JJ, Callahan T, Laraway B, et al. Generalizable Long COVID Subtypes: Findings from the NIH N3C and RECOVER Programs. medRxiv. 2022. doi:10.1101/2022.05.24.22275398
6. Yoo SM, Liu TC, Motwani Y, Sim MS, Viswanathan N, Samras N, et al. Factors Associated with Post-Acute Sequelae of SARS-CoV-2 (PASC) After Diagnosis of Symptomatic COVID-19 in the Inpatient and Outpatient Setting in a Diverse Cohort. J Gen Intern Med. 2022;37(8):1988-95. doi:10.1007/s11606-022-07523-3
7. Zhang H, Zang C, Xu Z, Zhang Y, Xu J, Bian J, et al. Machine Learning for Identifying Data-Driven Subphenotypes of Incident Post-Acute SARS-CoV-2 Infection Conditions with Large Scale Electronic Health Records: Findings from the RECOVER Initiative. medRxiv. 2022. doi:10.1101/2022.05.21.22275412
8. Behnood SA, Shafran R, Bennett SD, Zhang AXD, O'Mahoney LL, Stephenson TJ, et al. Persistent symptoms following SARS-CoV-2 infection amongst children and young people: A meta-analysis of controlled and uncontrolled studies. J Infect. 2022;84(2):158-70. doi:10.1016/j.jinf.2021.11.011
9. Palacios S, Krivchenia K, Eisner M, Young B, Ramilo O, Mejias A, et al. Long-term pulmonary sequelae in adolescents post-SARS-CoV-2 infection. Pediatr Pulmonol. 2022. doi:10.1002/ppul.26059
10. Ashkenazi-Hoffnung L, Shmueli E, Ehrlich S, Ziv A, Bar-On O, Birk E, et al. Long COVID in Children: Observations From a Designated Pediatric Clinic. Pediatr Infect Dis J. 2021;40(12):e509-e11. doi:10.1097/INF.0000000000003285
11. Asadi-Pooya AA, Nemati M, Nemati H. 'Long COVID': Symptom persistence in children hospitalised for COVID-19. J Paediatr Child Health. 2022. doi:10.1111/jpc.16120
12. Buonsenso D, Munblit D, De Rose C, Sinatti D, Ricchiuto A, Carfi A, et al. Preliminary evidence on long COVID in children. Acta Paediatrica. 2021;110(7):2208-11.
13. Fainardi V, Meoli A, Chiopris G, Motta M, Skenderaj K, Grandinetti R, et al. Long COVID in Children and Adolescents. Life (Basel). 2022;12(2). doi:10.3390/life12020285
14. Roge I, Smane L, Kivite-Urtane A, Pucuka Z, Racko I, Klavina L, et al. Comparison of Persistent Symptoms After COVID-19 and Other Non-SARS-CoV-2 Infections in Children. Frontiers in Pediatrics. 2021;9. doi:10.3389/fped.2021.752385

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

15. Thallapureddy K, Thallapureddy K, Zerda E, Suresh N, Kamat D, Rajasekaran K, et al. Long-Term Complications of COVID-19 Infection in Adolescents and Children. Curr Pediatr Rep. 2022;10(1):11-7. doi:10.1007/s40124-021-00260-x
16. Morrow AK, Ng R, Vargas G, Jashar DT, Henning E, Stinson N, et al. Postacute/Long COVID in Pediatrics: Development of a Multidisciplinary Rehabilitation Clinic and Preliminary Case Series. Am J Phys Med Rehabil. 2021;100(12):1140-7. doi:10.1097/PHM.0000000000001896
17. Brackel CLH, Lap CR, Buddingh EP, van Houten MA, van der Sande L, Langereis EJ, et al. Pediatric long-COVID: An overlooked phenomenon? Pediatr Pulmonol. 2021;56(8):2495-502. doi:10.1002/ppul.25521
18. Rao S, Lee GM, Razzaghi H, Lorman V, Mejias A, Pajor NM, et al. Clinical features and burden of post-acute sequelae of SARS-CoV-2 infection in children and adolescents: an exploratory EHR-based cohort study from the RECOVER program. medRxiv. 2022. doi:10.1101/2022.05.24.22275544
19. aapmr. CDC Announces Approval of ICD-10 Code for Post-Acute Sequelae of COVID-19. 2021.
20. Mejias A, Schuchard J, Rao S, Bennett TD, Jhaveri R, Thacker D, et al. Leveraging serology testing to identify children at risk for post-acute sequelae of SARS-CoV-2 infection: An EHR-based cohort study from the RECOVER program. medRxiv. 2022:2022.06.20.22276645. doi:10.1101/2022.06.20.22276645
21. Wang SV, Jin Y, Fireman B, Gruber S, He M, Wyss R, et al. Relative Performance of Propensity Score Matching Strategies for Subgroup Analyses. American Journal of Epidemiology. 2018;187(8):1799-807. doi:10.1093/aje/kwy049
22. Simon TD, Haaland W, Hawley K, Lambka K, Mangione-Smith R. Development and Validation of the Pediatric Medical Complexity Algorithm (PMCA) Version 3.0. Acad Pediatr. 2018;18(5):577-80. doi:10.1016/j.acap.2018.02.010
23. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 2:69-80. doi:10.1002/pds.3263
24. Sun YX, Wang M, Yang MF, Zhan SY. [Review on tree-based scan statistic in drug and vaccine safety monitoring]. Zhonghua Liu Xing Bing Xue Za Zhi. 2021;42(7):1286-91. doi:10.3760/cma.j.cn112338-20201103-01297
25. Templeton AR, Maxwell T, Posada D, Stengård JH, Boerwinkle E, Sing CF. Tree scanning: a method for using haplotype trees in phenotype/genotype association studies. Genetics. 2005;169(1):441-53. doi:10.1534/genetics.104.030080
26. Kulldorff M. TreeScan User Guide, v2.0. 2020.
27. CDC. ICD-10-CM Tabular List of Diseases and Injuries. 2022.
28. Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. BMJ. 2021;374:n1648. doi:10.1136/bmj.n1648
29. Rando HM, Bennett TD, Byrd JB, Bramante C, Callahan TJ, Chute CG, et al. Challenges in defining Long COVID: Striking differences across literature, Electronic Health Records, and patient-reported information. medRxiv. 2021. doi:10.1101/2021.03.20.21253896

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted December 11, 2022. 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 .

30. Estiri H, Strasser ZH, Brat GA, Semenov YR, Consortium for Characterization of C-bEHR, Patel CJ, et al. Evolving phenotypes of non-hospitalized patients that indicate long COVID. BMC Med. 2021;19(1):249. doi:10.1186/s12916-021-02115-0

31. National Institute for Health and Care Excellence: Clinical Guidelines. COVID-19 rapid guideline: managing the long-term effects of COVID-19. London: National Institute for Health and Care Excellence (NICE)
Copyright © NICE 2020.; 2020.

32. Asadi-Pooya AA, Nemati H, Shahisavandi M, Akbari A, Emami A, Lotfi M, et al. Long COVID in children and adolescents. World J Pediatr. 2021;17(5):495-9. doi:10.1007/s12519-021-00457-6

33. Osmanov IM, Spiridonova E, Bobkova P, Gamirova A, Shikhaleva A, Andreeva M, et al. Risk factors for post-COVID-19 condition in previously hospitalised children using the ISARIC Global follow-up protocol: a prospective cohort study. European Respiratory Journal. 2022;59(2):2101341. doi:10.1183/13993003.01341-2021

34. Radtke T, Ulyte A, Puhan MA, Kriemler S. Long-term Symptoms After SARS-CoV-2 Infection in Children and Adolescents. JAMA. 2021. doi:10.1001/jama.2021.11880

35. Borch L, Holm M, Knudsen M, Ellermann-Eriksen S, Hagstroem S. Long COVID symptoms and duration in SARS-CoV-2 positive children - a nationwide cohort study. Eur J Pediatr. 2022;181(4):1597-607. doi:10.1007/s00431-021-04345-z

36. Molteni E, Sudre CH, Canas LS, Bhopal SS, Hughes RC, Antonelli M, et al. Illness duration and symptom profile in symptomatic UK school-aged children tested for SARS-CoV-2. Lancet Child Adolesc Health. 2021;5(10):708-18. doi:10.1016/S2352-4642(21)00198-X

37. Erol N, Alpinar A, Erol C, Sari E, Alkan K. Intriguing new faces of Covid-19: persisting clinical symptoms and cardiac effects in children. Cardiol Young. 2022;32(7):1085-91. doi:10.1017/s1047951121003693

38. Lopez-Leon S, Wegman-Ostrosky T, Ayuzo Del Valle NC, Perelman C, Sepulveda R, Rebolledo PA, et al. Long-COVID in children and adolescents: a systematic review and meta-analyses. Sci Rep. 2022;12(1):9950. doi:10.1038/s41598-022-13495-5

39. Groff D, Sun A, Ssentongo AE, Ba DM, Parsons N, Poudel GR, et al. Short-term and Long-term Rates of Postacute Sequelae of SARS-CoV-2 Infection: A Systematic Review. JAMA Netw Open. 2021;4(10):e2128568. doi:10.1001/jamanetworkopen.2021.28568

40. Barrett CE KA, Alvarez P, et al. Risk for Newly Diagnosed Diabetes >30 Days After SARS-CoV-2 Infection Among Persons Aged <18 Years — United States, March 1, 2020–June 28, 20212022.

41. Naendrup JH, Garcia Borrega J, Eichenauer DA, Shimabukuro-Vornhagen A, Kochanek M, Böll B. Reactivation of EBV and CMV in Severe COVID-19-Epiphenomena or Trigger of Hyperinflammation in Need of Treatment? A Large Case Series of Critically ill Patients. J Intensive Care Med. 2022;37(9):1152-8. doi:10.1177/08850666211053990

42. Dotan A, Muller S, Kanduc D, David P, Halpert G, Shoenfeld Y. The SARS-CoV-2 as an instrumental trigger of autoimmunity. Autoimmun Rev. 2021;20(4):102792. doi:10.1016/j.autrev.2021.102792