medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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], 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], Hanieh Razzaghi, MPH[*1]

*Co-senior authorship

Affiliations:

[1]Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA, 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NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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 February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

[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

[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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

[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: 4416

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 conditions and symptoms associated with pediatric PASC.

Materials and Methods

We used a propensity-matched cohort design comparing children identified using the new PASC ICD10CM diagnosis code (U09.9) (N=1309) to children with (N=6545) and without (N=6545) 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 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 February 13, 2023. 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 .

LAY SUMMARY

Pediatric long COVID in children does not currently have a precise clinical definition, in part due to its widely varying presentation in kids. By comparing children diagnosed with long COVID to children who had COVID-19 but were not diagnosed with long COVID, this study identified several groups of symptoms and conditions that are associated with pediatric long COVID. These findings can be used towards developing a precise definition of long COVID in children for use in future studies.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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. (16)

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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

the lack of consensus of specific symptoms that define PASC in children necessitates a more data-driven approach for knowledge discovery.(17)

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 (18). 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 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

for consent was waived. The PEDSnet COVID-19 Database Version week 141 was used, which includes clinical histories through December 1, 2022 of patients who have been tested for SARS-CoV-2, diagnosed with COVID-19, or received a COVID-19 vaccine[1]. 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 or an interface terminology (IMO) term in the EHR with the following strings: ('post' and 'acute' and 'covid') or ('complication' and 'covid'). As evidence for PASC, we also admitted a B94.8 (Sequelae of other specified infectious and parasitic diseases) 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,

---

[1] More information on the PEDSnet RECOVER cohort is available at https://github.com/PEDSnet/Data_Models_Public/blob/master/PEDSnet/docs/RECOVER%20Cohort.md

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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(19). 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 (only 533 of 1309 patients with PASC [40.7%] had evidence of SARS-CoV-2 infection prior to their earliest PASC diagnosis). 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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. Further, to ensure that patients had a history of care at their institution, we required that all patients had at least two encounters in the health system in the 18 months prior to cohort entry.

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 December 1, 2022. We used a propensity score matched approach to balance covariates between cases and each of the comparison groups to minimize potential confounding (20). 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) (21) 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

cohort entry month and age group (22). 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(23) and disease surveillance and characterization(23, 24), 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(25). 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(26) 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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 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 spanning from 18 months prior to cohort entry to 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.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

Study population and matching

Between March 1, 2020, and June 22, 2022, there was a total of 14,399 patients identified for inclusion into the 3 cohorts, with PASC cases accounting for 1,309 patients (Table 1). Older children and females were more likely to be included in the PASC cohort than younger children and males (overall: 54.9% vs 45.1%). The most common cohort entry months were in the fall of 2021. A majority of children in the PASC cohort (55.8%) had a chronic condition. Following 5:1 nearest neighbor matching, both comparison cohorts had excellent balance (|SMD|<0.1 for all 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.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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		14,399	6,545	6,545	1,309
Age at cohort entrance (%)	<1	363 (2.5)	165 (2.5)	165 (2.5)	33 (2.5)
	1-4	1309 (9.1)	595 (9.1)	595 (9.1)	119 (9.1)
	5-11	4708 (32.7)	2140 (32.7)	2140 (32.7)	428 (32.7)
	12-15	4213 (29.3)	1915 (29.3)	1915 (29.3)	383 (29.3)
	16-20	3806 (26.4)	1730 (26.4)	1730 (26.4)	346 (26.4)
Sex (%)	Female	7903 (54.9)	3544 (54.1)	3640 (55.6)	719 (54.9)
	Male	6496 (45.1)	3001 (45.9)	2905 (44.4)	590 (45.1)
Race/ethnicity (%)	Non-Hispanic Asian/PI	519 (3.6)	245 (3.7)	228 (3.5)	46 (3.5)
	Non-Hispanic Black/AA	1425 (9.9)	623 (9.5)	674 (10.3)	128 (9.8)
	Hispanic	1698 (11.8)	748 (11.4)	793 (12.1)	157 (12.0)
	Multiple	593 (4.1)	258 (3.9)	282 (4.3)	53 (4.0)
	Other/Unknown	1275 (8.9)	540 (8.3)	607 (9.3)	128 (9.8)
	Non-Hispanic White	8889 (61.7)	4131 (63.1)	3961 (60.5)	797 (60.9)
Cohort entry month (%)	Mar 2020	77 (0.5)	35 (0.5)	35 (0.5)	7 (0.5)
	Apr 2020	132 (0.9)	60 (0.9)	60 (0.9)	12 (0.9)
	May 2020	66 (0.5)	30 (0.5)	30 (0.5)	6 (0.5)
	Jun 2020	110 (0.8)	50 (0.8)	50 (0.8)	10 (0.8)
	Jul 2020	121 (0.8)	55 (0.8)	55 (0.8)	11 (0.8)
	Aug 2020	110 (0.8)	50 (0.8)	50 (0.8)	10 (0.8)
	Sep 2020	165 (1.1)	75 (1.1)	75 (1.1)	15 (1.1)
	Oct 2020	198 (1.4)	90 (1.4)	90 (1.4)	18 (1.4)
	Nov 2020	341 (2.4)	155 (2.4)	155 (2.4)	31 (2.4)
	Dec 2020	341 (2.4)	155 (2.4)	155 (2.4)	31 (2.4)

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	Jan 2021	429 (3.0)	195 (3.0)	195 (3.0)	39 (3.0)
	Feb 2021	242 (1.7)	110 (1.7)	110 (1.7)	22 (1.7)
	Mar 2021	352 (2.4)	160 (2.4)	160 (2.4)	32 (2.4)
	Apr 2021	385 (2.7)	175 (2.7)	175 (2.7)	35 (2.7)
	May 2021	286 (2.0)	130 (2.0)	130 (2.0)	26 (2.0)
	Jun 2021	385 (2.7)	175 (2.7)	175 (2.7)	35 (2.7)
	Jul 2021	418 (2.9)	190 (2.9)	190 (2.9)	38 (2.9)
	Aug 2021	671 (4.7)	305 (4.7)	305 (4.7)	61 (4.7)
	Sep 2021	1111 (7.7)	505 (7.7)	505 (7.7)	101 (7.7)
	Oct 2021	770 (5.3)	350 (5.3)	350 (5.3)	70 (5.3)
	Nov 2021	957 (6.6)	435 (6.6)	435 (6.6)	87 (6.6)
	Dec 2021	1584 (11.0)	720 (11.0)	720 (11.0)	144 (11.0)
	Jan 2022	1837 (12.8)	835 (12.8)	835 (12.8)	167 (12.8)
	Feb 2022	1023 (7.1)	465 (7.1)	465 (7.1)	93 (7.1)
	Mar 2022	858 (6.0)	390 (6.0)	390 (6.0)	78 (6.0)
	Apr 2022	693 (4.8)	315 (4.8)	315 (4.8)	63 (4.8)
	May 2022	737 (5.1)	335 (5.1)	335 (5.1)	67 (5.1)
Institution (%)	**A**	1116 (7.8)	562 (8.6)	453 (6.9)	101 (7.7)
	B	1507 (10.5)	664 (10.1)	686 (10.5)	157 (12.0)
	C	2724 (18.9)	1275 (19.5)	1187 (18.1)	262 (20.0)
	D	469 (3.3)	212 (3.2)	218 (3.3)	39 (3.0)
	E	1058 (7.3)	429 (6.6)	541 (8.3)	88 (6.7)
	F	801 (5.6)	333 (5.1)	391 (6.0)	77 (5.9)
	G	1130 (7.8)	525 (8.0)	495 (7.6)	110 (8.4)
	H	2670 (18.5)	1235 (18.9)	1201 (18.3)	234 (17.9)
	I	2924 (20.3)	1310 (20.0)	1373 (21.0)	241 (18.4)
Test Location (%)	**ED**	2087 (14.5)	863 (13.2)	1055 (16.1)	169 (12.9)
	Inpatient	1446 (10.0)	627 (9.6)	695 (10.6)	124 (9.5)
	Other/Unknown	372 (2.6)	139 (2.1)	189 (2.9)	44 (3.4)
	Outpatient Office	6702 (46.5)	3130 (47.8)	2940 (44.9)	632 (48.3)

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

PMCA Index (%)	Outpatient: Test Only	3792 (26.3)	1786 (27.3)	1666 (25.5)	340 (26.0)
	Non-chronic	6297 (43.7)	2802 (42.8)	2917 (44.6)	578 (44.2)
	Chronic non-complex	5605 (38.9)	2572 (39.3)	2535 (38.7)	498 (38.0)
	Complex chronic	2497 (17.3)	1171 (17.9)	1093 (16.7)	233 (17.8)

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

Tree-based scan statistic

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 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 subbranch of the R00-R99 branch is visualized in Figure 1.

The PASC vs SARS-CoV-2 negative comparison showed significance for 304 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.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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, encephalopathy*, 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, mood disorders
Diseases of the respiratory system	asthma, ARDS/respiratory failure, diseases of vocal cords and larynx, tonsilitis, exercise induced bronchospasm*, pulmonary edema*
Diseases of the circulatory system	arrhythmias, cerebrovascular disease, thromboembolic disease, hypotension, tachycardia*, cardiac murmurs*
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, osteomyelitis*
Diseases of the digestive system	Diseases of oesophagus, stomach, and duodenum, aphagia and dysphagia, diseases of liver*
Diseases of blood	anemias*, sickle cell disorders*
Endocrine, nutritional and metabolic diseases	nutritional deficiencies*, obesity and overweight, disorders of mineral metabolism*, volume depletion or fluid overload, obesity, hyperkalemia*
Renal / Urinary	Acute kidney failure*, signs and symptoms involving the urinary system, hematuria
Diseases of the eye	Disorders of refraction and accommodation*
Immunologic findings	Other disorders involving the immune mechanism not elsewhere

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	classified, cytomegaloviral disease*
Major syndromic findings, or symptoms	
Circulatory and respiratory signs and symptoms	Palpitations, cough, dyspnea, shortness of breath, chest pain, viral and bacterial pneumonia, snoring, asphyxia and hypoxemia, hyperglycemia*
GI signs/symptoms	abdominal pain, nausea/vomiting, diarrhea*, constipation*, gastroenteritis*, pressure ulcer*
CNS/musculoskeletal signs/symptoms Cognition, perception, emotional state	lack of coordination, abnormalities of gait and mobility dizziness, smell/taste disturbances, restlessness and agitation*, disorientation*, depressive episode
General signs and symptoms	Fatigue/malaise, muscle weakness, headache, weakness, syncope and collapse, pain not elsewhere classified, fever*, enlarged lymph nodes*, nasal congestion*, epistaxis, noninflammatory disorders of female genital tract*, other inflammation of vagina and vulva*, visual disturbances
Skin	Symptoms and signs involving the skin and subcutaneous tissue, eczema,* urticaria* and erythema*, paresthesia of skin*

*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 February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

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(27) .

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 finding is not surprising given the heterogeneity of previously reported manifestations of PASC and the difficulty in gaining clinical consensus to define PASC(28, 29).

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, 16), shortness of breath(16,

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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), and other breathing abnormalities(31); chest pain(10, 32, 33); malaise and fatigue(14, 30, 33, 34); cough(14); and abnormalities of heartbeat including palpitations and tachycardia(35, 36). However, within this category we also detected novel signals such as hematuria as well as non-specific symptoms and signs involving the urinary system. Additionally, within this category we also detected novel signals such as hematuria as well as non-specific symptoms and signs involving the urinary system.

Neurologic diagnoses such as headache and migraine were recurring, with several specific codes enriched in the cases. Sleep disturbances and, specifically, sleep apnea, were also present (12, 31).(12, 31). Generalized pain was a key feature in our musculoskeletal system findings, which include non-localized myalgias, joint pain, knee and leg pain, hip pain, upper abdomen and epigastric 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 (37). 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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. The branches of the ICD hierarchy that are a priori unlikely to be related to COVID-19, e.g. injuries (S00-T98), external causes of morbidity and mortality (V01-Y98), and congenital malformations, deformations, and chromosomal abnormalities (Q00-Q99) serve as a negative control outcome; the relative dearth of diagnoses from these branches serves as a check on residual bias from unmeasured confounding.

We also compared children with a PASC diagnosis code to SARS-CoV-2 test-negative patients. Most 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(38). 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

following an infection has been documented and warrants further investigation in the pediatric population(39). Alternate findings enriched in this analysis include features plausibly associated with COVID-19 infection including acute myocarditis, hypotension, fever, nasal congestion, 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.

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 Rao et al(17), 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 the results of Rao et al (17), and included sleep disorders, anemias, eye disorders, constipation, sepsis, nutritional deficiencies, hypotension, and various musculoskeletal conditions and soft tissue and muscle disorders.

Finally, we conducted two sensitivity analyses: the first comparing COVID positive to COVID negative patients, and the second comparing PASC-positive to COVID positive patients pre- and post- omicron. 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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. In our analysis stratified by pre- and post- omicron variant, we found that the pre-omicron era was enriched for metabolic (e.g., E70-E88: metabolic disorders), immunologic (e.g., D80-D89: Certain disorders involving the immune mechanism), and circulatory system (e.g., I95: hypotension) codes, as well as disturbances of smell and taste, while the post-omicron era was enriched for conditions affecting the urinary system (e.g., R31: hematuria) and multiple migraine diagnoses. However, further analyses and additional methodologic rigor to ensure adequate power as well as balance of visit utilization and other potential confounders between the cohorts is warranted and beyond the scope of the current analysis.In our analysis stratified by pre-omicron and omicron variants, we found that the pre-omicron era was enriched for metabolic (e.g., E70-E88: metabolic disorders), immunologic (e.g., D80-D89: Certain disorders involving the immune mechanism), and circulatory system (e.g., I95: hypotension) codes, as well as disturbances of smell and taste, while the omicron era was enriched for conditions affecting the urinary system (e.g., R31: hematuria) and multiple migraine diagnoses. However, further analyses and additional methodologic rigor to ensure adequate power as well as balance of potential confounders between the cohorts is warranted and beyond the scope of the current analysis.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

Strengths of our study include using a tree-based scan statistic approach in children to detect groups of clinical diagnoses specific to this population, as well as the use of the U09.9 and B94.8 diagnosis codes to do so. Our study is further 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(17) 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 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

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

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 thus our results concern PASC-associated features in aggregate over the time period of the study; the course of PASC may differ by variant and further information may be gained by studying evolution of diagnoses across different portions of the pandemic or over time since infection. Finally, the cohort comprises 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. Only about half of these hospitals offer robust primary care networks, and therefore children seen in tertiary care centers exclusively may not be representative of all children in the US. This is complicated by the fact that children may only visit these health systems occasionally (e.g., ED visits) and receive the majority of their care elsewhere. Nonetheless, despite these limitations, in previous studies, we have successfully replicated similar rates of health outcomes

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

in our database as have been reported in data representative of the general pediatric population.(40, 41) In the COVID-19context, we have reported that 76% of our patient population have no chronic conditions, 12% have chronic conditions, and 11% have complex-chronic conditions,(42) which are similar to patients found in medicated populations(43). Further, the risk of this bias altering the findings of the current study relies on differential impact on cases and controls, and we have mitigated this risk by matching on medical complexity and demographic factors. Finally, we required at least two encounters with the health system 7 days to 18 months prior to cohort entry to ensure history of clinical care within the network.

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 February 13, 2023. 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							
7	B94 Sequelae of other and unspecified infectious and parasitic diseases	3	281.31	0.001	157	26.17	499.92
8	B90-B94 Sequelae of infectious and parasitic diseases	2	281.31	0.001	157	26.17	499.92
9	B94.8 Sequelae of other specified infectious and parasitic diseases	4	275.93	0.001	154	25.67	499.92
17	A00-B99 Certain infections and parasitic diseases	1	120.32	0.001	1,119	186.50	115.55
152	A65-A69 Other spirochaetal diseases	2	11.74	0.002	15	2.50	340.00
182	B95-B98 Bacterial, viral and other infectious agents	2	9.95	0.024	301	50.17	61.45
D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism							
44	D89.89 Other specified disorders involving the immune mechanism, not elsewhere classified	5	48.29	0.001	46	7.67	395.44
62	D89 Other disorders involving the immune mechanism, not elsewhere classified	3	27.93	0.001	116	19.33	179.36
65	D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	1	26.84	0.001	941	156.83	56.86
69	D89.8 Other specified disorders involving the immune mechanism, not elsewhere	4	24.44	0.001	104	17.33	176.98

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	classified						
116	D80-D89 Certain disorders involving the immune mechanism	2	14.83	0.001	331	55.17	72.20
E00-E90 Endocrine, nutritional and metabolic diseases							
39	E00-E90 Endocrine, nutritional and metabolic diseases	1	52.09	0.001	1,571	261.83	61.56
59	E70-E88 Metabolic disorders	2	31.77	0.001	671	111.83	74.37
88	E65-E68 Obesity and other hyperalimentation	2	19.65	0.001	192	32.00	112.50
91	E66 Overweight and obesity	3	18.66	0.001	188	31.33	110.66
136	E66.9 Obesity, unspecified	4	12.92	0.001	78	13.00	146.15
147	E87 Other disorders of fluid, electrolyte and acid-base balance	3	12.03	0.002	185	30.83	88.13
171	E86.0 Dehydration	4	10.57	0.019	101	16.83	113.90
174	E86 Volume depletion	3	10.31	0.020	106	17.67	109.39
F00-F99 Mental and behavioural disorders							
21	F00-F99 Mental and behavioural disorders	1	97.90	0.001	2,037	339.50	74.96
49	F40-F48 Neurotic, stress-related and somatoform disorders	2	45.19	0.001	632	105.33	92.73
70	F41 Other anxiety disorders	3	24.34	0.001	277	46.17	103.60
74	F80-F89 Disorders of psychological development	2	22.51	0.001	506	84.33	71.94
89	F41.9 Anxiety disorder, unspecified	4	19.64	0.001	196	32.67	111.20
101	F44 Dissociative and conversion disorders	3	16.71	0.001	27	4.50	300.00
129	F30-F39 Mood [affective] disorders	2	13.88	0.001	279	46.50	76.34
143	F00-F09 Organic, including symptomatic, mental disorders	2	12.30	0.002	30	5.00	240.00
179	F90-F98 Behavioural and	2	10.11	0.022	200	33.33	77.02

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 disorders with onset usually occurring in childhood and adolescence						
180	F32 Depressive episode	3	10.01	0.022	183	30.50	80.33
203	F41.1 Generalized anxiety disorder	4	9.03	0.042	36	6.00	183.33
G00-G99 Diseases of the nervous system							
12	G00-G99 Diseases of the nervous system	1	199.49	0.001	2,560	426.67	97.11
23	G40-G47 Episodic and paroxysmal disorders	2	89.80	0.001	1,640	273.33	80.37
28	G89-G99 Other disorders of the nervous system	2	83.68	0.001	570	95.00	136.84
37	G89 Pain, not elsewhere classified	3	57.51	0.001	296	49.33	159.48
41	G43 Migraine	3	50.63	0.001	337	56.17	138.56
43	G89.2 Chronic pain, not elsewhere classified	4	48.80	0.001	225	37.50	169.33
45	G89.29 Other chronic pain	5	48.22	0.001	223	37.17	169.03
55	G47 Sleep disorders	3	36.02	0.001	581	96.83	85.89
57	G90 Disorders of autonomic nervous system	3	32.80	0.001	38	6.33	358.14
58	G44 Other headache syndromes	3	32.26	0.001	107	17.83	202.86
63	G43.0 Migraine without aura	4	27.64	0.001	137	22.83	162.81
68	G90.9 Disorder of the autonomic nervous system, unspecified	4	25.05	0.001	17	2.83	465.37
71	G00-G09 Inflammatory diseases of the central nervous system	2	23.34	0.001	62	10.33	229.14
73	G44.5 Complicated headache syndromes	4	22.88	0.001	26	4.33	361.89
90	G04 Encephalitis, myelitis and encephalomyelitis	3	19.22	0.001	37	6.17	272.77
98	G43.01 Migraine without aura,	5	16.91	0.001	49	8.17	218.24

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	intractable						
99	G44.52 New daily persistent headache (NDPH)	5	16.87	0.001	22	3.67	335.97
115	G04.8 Other encephalitis, myelitis and encephalomyelitis	4	14.85	0.001	24	4.00	300.00
119	G90.1 Familial dysautonomia [Riley-Day]	4	14.75	0.001	11	1.83	446.45
125	G93 Other disorders of brain	3	14.10	0.001	175	29.17	98.83
127	G43.019 Migraine without aura, intractable, without status migrainosus	6	13.99	0.001	39	6.50	223.08
132	G47.3 Sleep apnea	4	13.60	0.001	321	53.50	70.09
141	G43.009 Migraine without aura, not intractable, without status migrainosus	6	12.71	0.001	75	12.50	148.00
146	G43.00 Migraine without aura, not intractable	5	12.06	0.002	88	14.67	131.77
160	G93.3 Postviral fatigue syndrome	4	11.27	0.007	11	1.83	391.80
167	G44.8 Other specified headache syndromes	4	10.80	0.007	21	3.50	271.43
183	G43.80 Other migraine, not intractable	5	9.92	0.026	12	2.00	350.00
188	G44.89 Other headache syndrome	5	9.77	0.026	17	2.83	288.69
I00-I99 Diseases of the circulatory system							
47	I00-I99 Diseases of the circulatory system	1	46.29	0.001	893	148.83	78.06
87	I49 Other cardiac arrhythmias	3	19.75	0.001	39	6.50	269.23
97	I30-I52 Other forms of heart disease	2	17.11	0.001	261	43.50	88.51
128	I95.1 Orthostatic hypotension	4	13.96	0.001	20	3.33	320.42
135	I49.8 Other specified cardiac arrhythmias	4	13.17	0.001	26	4.33	269.52

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

178	I95-I99 Other and unspecified disorders of the circulatory system	2	10.17	0.022	68	11.33	138.31
198	I51 Complications and ill-defined descriptions of heart disease	3	9.16	0.041	39	6.50	176.92
J00-J99 Diseases of the respiratory system							
25	J00-J99 Diseases of the respiratory system	1	86.96	0.001	3,044	507.33	56.90
53	J40-J47 Chronic lower respiratory diseases	2	41.76	0.001	885	147.50	74.24
54	J45 Asthma	3	41.35	0.001	852	142.00	75.35
75	J45.9 Other and unspecified asthma	4	22.14	0.001	254	42.33	103.17
86	J30-J39 Other diseases of upper respiratory tract	2	19.77	0.001	603	100.50	61.19
92	J12.8 Other viral pneumonia	4	18.35	0.001	23	3.83	343.86
94	J96-J99 Other diseases of the respiratory system	2	18.06	0.001	391	65.17	73.39
95	J45.909 Unspecified asthma, uncomplicated	6	17.67	0.001	180	30.00	110.00
111	J12 Viral pneumonia, not elsewhere classified	3	15.81	0.001	39	6.50	238.46
113	J45.90 Unspecified asthma	5	15.08	0.001	212	35.33	92.47
122	J12.82 Pneumonia due to coronavirus disease 2019	5	14.49	0.001	13	2.17	406.91
139	J38 Diseases of vocal cords and larynx, not elsewhere classified	3	12.80	0.001	54	9.00	177.78
144	J03 Acute tonsillitis	3	12.30	0.002	30	5.00	240.00
150	J98 Other respiratory disorders	3	11.90	0.002	152	25.33	97.39
153	J96.0 Acute respiratory failure	4	11.61	0.003	82	13.67	134.09
156	J03.91 Acute recurrent tonsillitis, unspecified	5	11.38	0.006	9	1.50	433.33

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

175	J45.4 Moderate persistent asthma	4	10.22	0.021	156	26.00	88.46
184	J09-J18 Influenza and pneumonia	2	9.92	0.026	210	35.00	74.29
194	J15 Bacterial pneumonia, not elsewhere classified	3	9.33	0.038	26	4.33	223.33
199	J45.40 Moderate persistent asthma, uncomplicated	5	9.10	0.042	132	22.00	90.91
K00-K93 Diseases of the digestive system							
165	K20-K31 Diseases of oesophagus, stomach and duodenum	2	10.94	0.007	387	64.50	56.59
181	K00-K93 Diseases of the digestive system	1	10.00	0.022	2,054	342.33	22.69
L00-L99 Diseases of the skin and subcutaneous tissue							
96	L00-L99 Diseases of the skin and subcutaneous tissue	1	17.45	0.001	966	161.00	44.72
M00-M99 Diseases of the musculoskeletal system and connective tissue							
15	M00-M99 Diseases of the musculoskeletal system and connective tissue	1	183.48	0.001	3,258	543.00	81.58
22	M20-M25 Other joint disorders	2	93.84	0.001	1,038	173.00	105.20
31	M25 Other joint disorder, not elsewhere classified	3	67.30	0.001	658	109.67	112.46
36	M25.5 Pain in joint	4	59.36	0.001	462	77.00	127.27
46	M79 Other and unspecified soft tissue disorders, not elsewhere classified	3	47.13	0.001	476	79.33	110.51
50	M79.1 Myalgia	4	43.85	0.001	111	18.50	235.14
52	M70-M79 Other soft tissue disorders	2	42.76	0.001	527	87.83	99.25
72	M25.50 Pain in unspecified joint	5	23.03	0.001	51	8.50	252.94
77	M79.18 Myalgia, other site	5	21.61	0.001	56	9.33	232.26
78	M60-M63 Disorders of muscles	2	21.37	0.001	402	67.00	79.10

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

82	M79.10 Myalgia, unspecified site	5	20.46	0.001	52	8.67	234.49
83	M62 Other disorders of muscle	3	20.15	0.001	396	66.00	77.27
84	M25.56 Pain in knee	5	20.03	0.001	171	28.50	121.05
93	M02.3 Reiter's disease	4	18.16	0.001	13	2.17	453.00
102	M50-M54 Other dorsopathies	2	16.59	0.001	268	44.67	85.81
103	M35 Other systemic involvement of connective tissue	3	16.47	0.001	38	6.33	247.55
104	M02.30 Reiter's disease, unspecified site	5	16.45	0.001	12	2.00	450.00
105	M30-M36 Systemic connective tissue disorders	2	16.32	0.001	97	16.17	147.37
107	M02 Postinfective and reactive arthropathies	3	16.12	0.001	14	2.33	415.02
110	M54 Dorsalgia	3	15.84	0.001	242	40.33	88.45
112	M62.8 Other specified disorders of muscle	4	15.76	0.001	343	57.17	73.17
114	M00-M02 Infectious arthropathies	2	15.05	0.001	19	3.17	341.64
121	M21 Other acquired deformities of limbs	3	14.56	0.001	177	29.50	100.00
126	M79.6 Pain in limb, hand, foot, fingers and toes	4	14.06	0.001	318	53.00	71.70
130	M24 Other specific joint derangements	3	13.77	0.001	156	26.00	103.85
133	M24.9 Joint derangement, unspecified	4	13.23	0.001	16	2.67	349.44
137	M25.55 Pain in hip	5	12.87	0.001	71	11.83	153.59
140	M86-M90 Other osteopathies	2	12.78	0.001	144	24.00	104.17
151	M35.7 Hypermobility syndrome	4	11.87	0.002	25	4.17	259.71
154	M79.60 Pain in limb, unspecified	5	11.55	0.003	129	21.50	104.65
161	M25.551 Pain in right hip	6	11.09	0.007	29	4.83	231.26

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

172	M25.561 Pain in right knee	6	10.42	0.020	86	14.33	123.31
186	M62.81 Muscle weakness (generalized)	5	9.77	0.026	180	30.00	80.00
189	M89.1 Physeal arrest	4	9.69	0.034	10	1.67	379.04
192	M79.66 Pain in lower leg	5	9.40	0.037	32	5.33	200.19
193	M21.8 Other specified acquired deformities of limbs	4	9.37	0.037	23	3.83	239.43
196	M89 Other disorders of bone	3	9.27	0.040	49	8.17	157.04
Q00-Q99 Congenital malformations, deformations and chromosomal abnormalities							
191	Q79 Congenital malformations of musculoskeletal system, not elsewhere classified	3	9.52	0.036	59	9.83	144.15
197	Q00-Q99 Congenital malformations, deformations and chromosomal abnormalities	1	9.23	0.041	746	124.33	36.73
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	816.06	0.001	8,383	1,397.17	109.49
6	R00-R09 Symptoms and signs involving the circulatory and respiratory systems	2	365.38	0.001	2,038	339.67	152.60
10	R06.0 Dyspnea	4	242.71	0.001	381	63.50	304.72
11	R06 Abnormalities of breathing	3	207.49	0.001	706	117.67	199.99
13	R50-R69 General symptoms and signs	2	195.76	0.001	2,236	372.67	103.40
14	R53 Malaise and fatigue	3	185.81	0.001	403	67.17	255.81
16	R53.8 Other malaise and fatigue	4	179.60	0.001	338	56.33	276.35
18	R40-R46 Symptoms and signs involving cognition, perception, emotional state and behaviour	2	106.54	0.001	724	120.67	137.01
19	R53.83 Other fatigue	5	105.25	0.001	230	38.33	254.81

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

20	R06.02 Shortness of breath	5	99.42	0.001	148	24.67	313.46
24	R10-R19 Symptoms and signs involving the digestive system and abdomen	2	88.31	0.001	1,710	285.00	77.89
26	R07 Pain in throat and chest	3	86.65	0.001	281	46.83	205.36
27	R06.00 Dyspnea, unspecified	5	86.01	0.001	104	17.33	350.09
29	R06.09 Other forms of dyspnea	5	79.13	0.001	81	13.50	381.48
30	R42 Dizziness and giddiness	3	67.75	0.001	175	29.17	232.53
32	R07.9 Chest pain, unspecified	4	62.74	0.001	151	25.17	241.68
35	R25-R29 Symptoms and signs involving the nervous and musculoskeletal systems	2	60.00	0.001	300	50.00	162.00
38	R51 Headache	3	53.53	0.001	234	39.00	174.36
40	R53.82 Chronic fatigue, unspecified	5	51.50	0.001	48	8.00	400.00
42	R51.9 Headache, unspecified	4	48.87	0.001	208	34.67	176.90
48	R00 Abnormalities of heart beat	3	45.42	0.001	206	34.33	170.90
51	R10 Abdominal and pelvic pain	3	42.83	0.001	648	108.00	88.89
56	R00.2 Palpitations	4	34.35	0.001	74	12.33	256.85
60	R29 Other symptoms and signs involving the nervous and musculoskeletal systems	3	28.87	0.001	187	31.17	140.62
61	R53.81 Other malaise	5	28.50	0.001	60	10.00	260.00
64	R07.89 Other chest pain	5	27.49	0.001	88	14.67	206.75
66	R07.8 Other chest pain	4	25.91	0.001	104	17.33	182.75
67	R05 Cough	3	25.83	0.001	412	68.67	86.40
76	R43 Disturbances of smell and taste	3	21.89	0.001	39	6.50	284.62
79	R55 Syncope and collapse	3	21.24	0.001	88	14.67	179.48
80	R11 Nausea and vomiting	3	20.83	0.001	454	75.67	73.12

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

81	R41 Other symptoms and signs involving cognitive functions and awareness	3	20.57	0.001	128	21.33	143.79
100	R00.0 Tachycardia, unspecified	4	16.84	0.001	92	15.33	154.40
106	R09 Other symptoms and signs involving the circulatory and respiratory system	3	16.24	0.001	340	56.67	74.70
108	R29.9 Unspecified symptoms and signs involving the nervous and musculoskeletal systems	4	16.12	0.001	16	2.67	386.89
109	R10.1 Pain localized to upper abdomen	4	15.91	0.001	128	21.33	125.04
117	R10.84 Generalized abdominal pain	5	14.81	0.001	180	30.00	100.00
118	R90-R96 Abnormal findings on diagnostic imaging and in function studies, without diagnosis	2	14.79	0.001	278	46.33	79.15
120	R10.8 Other abdominal pain	4	14.65	0.001	189	31.50	96.83
123	R29.90 Unspecified symptoms and signs involving the nervous system	5	14.49	0.001	13	2.17	406.91
131	R70-R79 Abnormal findings on examination of blood, without diagnosis	2	13.68	0.001	298	49.67	73.14
134	R26	3	13.17	0.001	26	4.33	269.52
138	R27 Other lack of coordination	3	12.83	0.001	38	6.33	215.96
142	R06.8 Other abnormalities of breathing	4	12.46	0.002	261	43.50	74.71
145	R68 Other general symptoms and signs	3	12.27	0.002	130	21.67	107.66
148	R30-R39 Symptoms and signs involving the urinary system	2	11.96	0.002	314	52.33	66.25
149	R04 Hemorrhage from respiratory	3	11.95	0.002	56	9.33	167.95

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	passages						
155	R53.1 Weakness	4	11.55	0.003	64	10.67	153.05
158	R09.0 Asphyxia and hypoxemia	4	11.31	0.006	106	17.67	115.05
159	R09.02 Hypoxemia	5	11.31	0.006	106	17.67	115.05
162	R05.9 Cough, unspecified	4	11.06	0.007	212	35.33	78.32
163	R31 Hematuria	3	11.05	0.007	48	8.00	175.00
164	R19 Other symptoms and signs involving the digestive system and abdomen	3	10.97	0.007	195	32.50	81.54
166	R29.898 Other symptoms and signs involving the musculoskeletal system	6	10.86	0.007	96	16.00	118.75
168	R10.13 Epigastric pain	5	10.76	0.007	81	13.50	129.63
169	R11.0 Nausea	4	10.59	0.019	97	16.17	116.45
170	R29.89 Other symptoms and signs involving the musculoskeletal system	5	10.59	0.019	97	16.17	116.45
173	R29.8 Other symptoms and signs involving the nervous and musculoskeletal systems	4	10.37	0.020	126	21.00	100.00
176	R13 Aphagia and dysphagia	3	10.22	0.021	299	49.83	62.55
177	R13.1 Dysphagia	4	10.22	0.021	299	49.83	62.55
185	R68.89 Other general symptoms and signs	5	9.82	0.026	96	16.00	112.50
187	R26.2 Difficulty in walking, not elsewhere classified	4	9.77	0.026	17	2.83	288.69
190	R06.83 Snoring	5	9.61	0.035	113	18.83	101.81
195	R04.0 Epistaxis	4	9.27	0.040	49	8.17	157.04
200	R46 Symptoms and signs involving appearance and behavior	3	9.10	0.042	132	22.00	90.91
201	R41.8 Other symptoms and signs involving cognitive functions and	4	9.08	0.042	95	15.83	108.46

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	awareness						
202	R05.3 Chronic cough	4	9.03	0.042	36	6.00	183.33
S00-T98 Injury, poisoning and certain other consequences of external causes							
157	S72.3 Fracture of shaft of femur	4	11.38	0.006	9	1.50	433.33
U00-U95 Codes for special purposes							
2	U09 Post COVID-19 condition	3	614.57	0.001	343	57.17	499.97
3	U09.9 Post COVID-19 condition, unspecified	4	614.57	0.001	343	57.17	499.97
4	U00-U95 Codes for special purposes	1	509.50	0.001	657	109.50	338.36
5	U00-U49 Provisional assignment of new diseases of uncertain etiology or emergency use	2	509.50	0.001	657	109.50	338.36
33	U07.1 Emergency use of U07.1 | COVID-19	4	62.65	0.001	314	52.33	161.80
34	U07	3	62.65	0.001	314	52.33	161.80
Z00-Z99 Factors influencing health status and contact with health services							
85	Z69-Z76 Persons encountering health services in other circumstances	2	19.96	0.001	268	44.67	94.76
124	Z00-Z99 Factors influencing health status and contact with health services	1	14.19	0.001	2,063	343.83	27.10

Table 3b: PASC vs COVID negative comparison TreeScan results

Cut	Node	Tree Level	Log Likelihood Ratio	P value	N cases observed	N cases expected	Percer excess cases
A00-B99 Certain infections and parasitic diseases							

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

10	B90-B94 Sequelae of infectious and parasitic diseases	2	281.31	0.001	157	26.17	499.92
11	B94 Sequelae of other and unspecified infectious and parasitic diseases	3	281.31	0.001	157	26.17	499.92
12	B94.8 Sequelae of other specified infectious and parasitic diseases	4	275.93	0.001	154	25.67	499.92
15	A00-B99 Certain infections and parasitic diseases	1	239.90	0.001	797	132.83	202.64
76	B95-B98 Bacterial, viral and other infectious agents	2	35.38	0.001	190	31.67	155.76
96	B25-B34 Other viral diseases	2	27.51	0.001	162	27.00	148.15
116	B97 Viral agents as the cause of diseases classified elsewhere	3	23.01	0.001	81	13.50	196.30
128	B34 Viral infection of unspecified site	3	20.49	0.001	121	20.17	147.89
214	B97.2 Coronavirus as the cause of diseases classified elsewhere	4	12.54	0.002	7	1.17	498.29
215	B97.29 Other coronavirus as the cause of diseases classified elsewhere	5	12.54	0.002	7	1.17	498.29
217	A30-A49 Other bacterial diseases	2	12.54	0.002	48	8.00	187.50
240	B97.89 Other viral agents as the cause of diseases classified elsewhere	5	11.38	0.003	54	9.00	166.67
246	B34.8 Other viral infections of unspecified site	4	11.14	0.006	18	3.00	300.00
252	B97.8 Other viral agents as the cause of diseases classified elsewhere	4	10.98	0.006	55	9.17	161.72
286	B25 Cytomegaloviral disease	3	9.69	0.020	10	1.67	379.04
D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism							
33	D50-D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	1	91.55	0.001	623	103.83	136.93

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

50	D89.89 Other specified disorders involving the immune mechanism, not elsewhere classified	5	55.61	0.001	42	7.00	442.86
55	D89 Other disorders involving the immune mechanism, not elsewhere classified	3	48.34	0.001	83	13.83	290.46
63	D89.8 Other specified disorders involving the immune mechanism, not elsewhere classified	4	43.65	0.001	73	12.17	294.41
65	D80-D89 Certain disorders involving the immune mechanism	2	42.19	0.001	221	36.83	157.94
158	D64.9 Anemia, unspecified	4	16.91	0.001	49	8.17	218.24
161	D60-D64 Aplastic and other anaemias	2	16.73	0.001	129	21.50	127.91
185	D64 Other anemias	3	14.59	0.001	70	11.67	165.64
200	D55-D59 Haemolytic anaemias	2	13.41	0.001	37	6.17	224.15
210	D65-D69 Coagulation defects, purpura and other haemorrhagic conditions	2	12.71	0.001	75	12.50	148.00
211	D57 Sickle-cell disorders	3	12.64	0.001	24	4.00	275.00
E00-E90 Endocrine, nutritional and metabolic diseases							
30	E00-E90 Endocrine, nutritional and metabolic diseases	1	98.96	0.001	1,294	215.67	96.13
42	E70-E88 Metabolic disorders	2	73.49	0.001	491	81.83	138.30
71	E87 Other disorders of fluid, electrolyte and acid-base balance	3	40.29	0.001	105	17.50	231.43
119	E86 Volume depletion	3	22.80	0.001	72	12.00	208.33
123	E86.0 Dehydration	4	22.21	0.001	70	11.67	208.48
129	E65-E68 Obesity and other hyperalimentation	2	20.43	0.001	189	31.50	115.87
133	E66 Overweight and obesity	3	19.94	0.001	183	30.50	116.39
203	E87.7 Fluid overload	4	13.17	0.001	26	4.33	269.52
208	E66.9 Obesity, unspecified	4	12.92	0.001	78	13.00	146.15

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

212	E87.70 Fluid overload, unspecified	5	12.64	0.001	24	4.00	275.00
239	E87.5 Hyperkalemia	4	11.38	0.003	9	1.50	433.33
261	E83 Disorders of mineral metabolism	3	10.61	0.012	129	21.50	100.00
270	E50-E64 Other nutritional deficiencies	2	10.33	0.013	147	24.50	91.84
F00-F99 Mental and behavioural disorders							
27	F00-F99 Mental and behavioural disorders	1	108.72	0.001	1,973	328.83	80.64
51	F40-F48 Neurotic, stress-related and somatoform disorders	2	54.50	0.001	590	98.33	106.45
81	F41 Other anxiety disorders	3	30.21	0.001	254	42.33	122.06
94	F80-F89 Disorders of psychological development	2	27.72	0.001	475	79.17	83.15
111	F41.9 Anxiety disorder, unspecified	4	24.05	0.001	180	30.00	130.00
171	F44 Dissociative and conversion disorders	3	15.83	0.001	28	4.67	285.44
189	F00-F09 Organic, including symptomatic, mental disorders	2	14.49	0.001	27	4.50	277.78
242	F80 Specific developmental disorders of speech and language	3	11.29	0.003	114	19.00	110.53
253	F69 Unspecified disorder of adult personality and behavior	3	10.75	0.011	6	1.00	500.00
267	F90-F98 Behavioural and emotional disorders with onset usually occurring in childhood and adolescence	2	10.45	0.013	198	33.00	78.79
275	F32 Depressive episode	3	10.19	0.013	182	30.33	81.34
294	F43 Reaction to severe stress, and adjustment disorders	3	9.45	0.022	222	37.00	70.27
G00-G99 Diseases of the nervous system							
13	G00-G99 Diseases of the nervous system	1	259.71	0.001	2,311	385.17	118.35

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

25	G40-G47 Episodic and paroxysmal disorders	2	122.60	0.001	1,475	245.83	100.55
29	G89-G99 Other disorders of the nervous system	2	101.22	0.001	520	86.67	159.61
48	G43 Migraine	3	57.17	0.001	318	53.00	152.83
52	G47 Sleep disorders	3	53.89	0.001	501	83.50	115.57
53	G89 Pain, not elsewhere classified	3	52.73	0.001	309	51.50	148.54
57	G89.2 Chronic pain, not elsewhere classified	4	47.17	0.001	229	38.17	164.61
59	G89.29 Other chronic pain	5	46.20	0.001	228	38.00	163.16
73	G44 Other headache syndromes	3	39.27	0.001	95	15.83	241.12
78	G00-G09 Inflammatory diseases of the central nervous system	2	31.56	0.001	51	8.50	300.00
82	G43.0 Migraine without aura	4	30.11	0.001	131	21.83	174.85
99	G90 Disorders of autonomic nervous system	3	26.46	0.001	44	7.33	295.63
101	G44.5 Complicated headache syndromes	4	25.75	0.001	24	4.00	400.00
106	G93 Other disorders of brain	3	24.61	0.001	138	23.00	152.17
112	G04 Encephalitis, myelitis and encephalomyelitis	3	23.84	0.001	32	5.33	331.52
131	G47.3 Sleep apnea	4	20.12	0.001	284	47.33	92.27
135	G44.52 New daily persistent headache (NDPH)	5	19.39	0.001	20	3.33	380.48
144	G43.01 Migraine without aura, intractable	5	18.74	0.001	46	7.67	238.98
148	G47.9 Sleep disorder, unspecified	4	18.19	0.001	75	12.50	180.00
149	G04.8 Other encephalitis, myelitis and encephalomyelitis	4	18.05	0.001	21	3.50	357.14
159	G90.9 Disorder of the autonomic nervous system, unspecified	4	16.87	0.001	22	3.67	335.97
168	G93.3 Postviral fatigue syndrome	4	16.13	0.001	9	1.50	500.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

176	G43.019 Migraine without aura, intractable, without status migrainosus	6	15.24	0.001	37	6.17	240.36
207	G43.00 Migraine without aura, not intractable	5	13.01	0.001	85	14.17	139.94
209	G93.4 Other and unspecified encephalopathy	4	12.87	0.001	64	10.67	162.42
223	G43.009 Migraine without aura, not intractable, without status migrainosus	6	12.37	0.002	76	12.67	144.67
231	G44.89 Other headache syndrome	5	11.74	0.002	15	2.50	340.00
235	G44.8 Other specified headache syndromes	4	11.62	0.002	20	3.33	290.39
241	G47.30 Sleep apnea, unspecified	5	11.38	0.003	54	9.00	166.67
258	G04.81 Other encephalitis and encephalomyelitis	5	10.68	0.012	16	2.67	311.99
272	G90.1 Familial dysautonomia [Riley-Day]	4	10.27	0.013	14	2.33	329.18
292	G93.40 Encephalopathy, unspecified	5	9.50	0.022	20	3.33	260.36
297	G25 Other extrapyramidal and movement disorders	3	9.40	0.023	32	5.33	200.19
H00-H59 Diseases of the eye and adnexa							
238	H53.1 Subjective visual disturbances	4	11.44	0.002	13	2.17	360.83
256	H53.14 Visual discomfort	5	10.75	0.011	6	1.00	500.00
291	H53-H54 Visual disturbances and blindness	2	9.59	0.020	138	23.00	91.30
298	H53 Visual disturbances	3	9.38	0.023	114	19.00	100.00
I00-I99 Diseases of the circulatory system							
32	I00-I99 Diseases of the circulatory system	1	95.20	0.001	682	113.67	133.13
75	I30-I52 Other forms of heart disease	2	35.57	0.001	193	32.17	154.90
151	I95.1 Orthostatic hypotension	4	17.71	0.001	17	2.83	394.70

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

157	I95-I99 Other and unspecified disorders of the circulatory system	2	16.93	0.001	52	8.67	211.42
163	I95 Hypotension	3	16.47	0.001	38	6.33	247.55
194	I49 Other cardiac arrhythmias	3	14.11	0.001	48	8.00	200.00
213	I60-I69 Cerebrovascular diseases	2	12.58	0.001	79	13.17	142.98
226	I10-I19 Hypertensive diseases	2	12.03	0.002	131	21.83	106.14
228	I26-I28 Pulmonary heart disease and diseases of pulmonary circulation	2	11.95	0.002	46	7.67	186.83
247	I49.8 Other specified cardiac arrhythmias	4	11.09	0.006	29	4.83	231.26
290	I51 Complications and ill-defined descriptions of heart disease	3	9.61	0.020	38	6.33	184.36
J00-J99 Diseases of the respiratory system							
14	J00-J99 Diseases of the respiratory system	1	246.45	0.001	2,185	364.17	118.58
35	J40-J47 Chronic lower respiratory diseases	2	86.15	0.001	682	113.67	126.09
36	J45 Asthma	3	83.40	0.001	661	110.17	126.01
64	J96-J99 Other diseases of the respiratory system	2	43.41	0.001	282	47.00	140.43
67	J30-J39 Other diseases of upper respiratory tract	2	41.35	0.001	480	80.00	102.50
69	J45.9 Other and unspecified asthma	4	40.97	0.001	193	32.17	167.33
74	J09-J18 Influenza and pneumonia	2	35.85	0.001	122	20.33	200.05
79	J45.909 Unspecified asthma, uncomplicated	6	31.00	0.001	139	23.17	171.90
85	J45.90 Unspecified asthma	5	29.52	0.001	160	26.67	154.97
103	J96 Respiratory failure, not elsewhere classified	3	25.29	0.001	154	25.67	145.42
114	J00-J06 Acute upper respiratory infections	2	23.47	0.001	460	76.67	77.38
117	J45.4 Moderate persistent asthma	4	22.92	0.001	111	18.50	164.86

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

118	J96.0 Acute respiratory failure	4	22.82	0.001	57	9.50	236.84
121	J12.8 Other viral pneumonia	4	22.55	0.001	20	3.33	410.51
122	J12 Viral pneumonia, not elsewhere classified	3	22.38	0.001	31	5.17	325.53
126	J45.40 Moderate persistent asthma, uncomplicated	5	21.38	0.001	91	15.17	176.86
134	J12.82 Pneumonia due to coronavirus disease 2019	5	19.71	0.001	11	1.83	501.09
139	J98 Other respiratory disorders	3	19.14	0.001	125	20.83	140.04
166	J38 Diseases of vocal cords and larynx, not elsewhere classified	3	16.32	0.001	47	7.83	219.28
174	J45.30 Mild persistent asthma, uncomplicated	5	15.34	0.001	130	21.67	121.50
188	J03 Acute tonsillitis	3	14.49	0.001	27	4.50	277.78
196	J35 Chronic diseases of tonsils and adenoids	3	13.85	0.001	65	10.83	167.77
199	J96.01 Acute respiratory failure with hypoxia	5	13.45	0.001	34	5.67	235.10
216	J22 Unspecified acute lower respiratory infection	3	12.54	0.002	7	1.17	498.29
224	J45.99 Other asthma	5	12.25	0.002	33	5.50	227.27
232	J80-J84 Other respiratory diseases principally affecting the interstitium	2	11.73	0.002	28	4.67	242.61
236	J45.3 Mild persistent asthma	4	11.60	0.002	166	27.67	91.54
244	J20-J22 Other acute lower respiratory infections	2	11.26	0.006	41	6.83	192.83
254	J15.21 Pneumonia due to staphylococcus aureus	5	10.75	0.011	6	1.00	500.00
255	J15.2 Pneumonia due to staphylococcus	4	10.75	0.011	6	1.00	500.00
257	J38.3 Other diseases of vocal cords	4	10.68	0.012	16	2.67	311.99
262	J15 Bacterial pneumonia, not	3	10.61	0.013	24	4.00	250.00

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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						
264	J45.2 Mild intermittent asthma	4	10.55	0.013	180	30.00	83.33
284	J03.0 Streptococcal tonsillitis	4	9.71	0.017	8	1.33	426.32
285	J03.01 Acute recurrent streptococcal tonsillitis	5	9.71	0.017	8	1.33	426.32
288	J03.91 Acute recurrent tonsillitis, unspecified	5	9.69	0.020	10	1.67	379.04
295	J45.20 Mild intermittent asthma, uncomplicated	5	9.44	0.022	143	23.83	88.84
K00-K93 Diseases of the digestive system							
77	K00-K93 Diseases of the digestive system	1	35.37	0.001	1,701	283.50	48.15
155	K20-K31 Diseases of oesophagus, stomach and duodenum	2	17.20	0.001	343	57.17	76.67
179	K55-K64 Other diseases of intestines	2	14.87	0.001	625	104.17	51.68
225	K59.0 Constipation	4	12.19	0.002	505	84.17	52.07
249	K59 Other functional intestinal disorders	3	11.01	0.006	521	86.83	48.57
250	K70-K77 Diseases of liver	2	10.98	0.006	62	10.33	151.69
274	K31.8 Other specified diseases of stomach and duodenum	4	10.23	0.013	50	8.33	164.11
304	K31 Other diseases of stomach and duodenum	3	8.97	0.031	57	9.50	142.11
L00-L99 Diseases of the skin and subcutaneous tissue							
62	L00-L99 Diseases of the skin and subcutaneous tissue	1	44.06	0.001	766	127.67	82.50
138	L80-L99 Other disorders of the skin and subcutaneous tissue	2	19.14	0.001	198	33.00	109.09
222	L89 Pressure ulcer	3	12.42	0.002	27	4.50	255.56
233	L20-L30 Dermatitis and eczema	2	11.69	0.002	248	41.33	74.21
268	L49-L54 Urticaria and erythema	2	10.41	0.013	53	8.83	160.48

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

M00-M99 Diseases of the musculoskeletal system and connective tissue							
18	M00-M99 Diseases of the musculoskeletal system and connective tissue	1	215.57	0.001	3,089	514.83	91.52
38	M79 Other and unspecified soft tissue disorders, not elsewhere classified	3	81.52	0.001	370	61.67	170.80
39	M70-M79 Other soft tissue disorders	2	79.67	0.001	402	67.00	161.19
41	M20-M25 Other joint disorders	2	77.45	0.001	1,113	185.50	91.37
46	M25.5 Pain in joint	4	58.78	0.001	464	77.33	126.30
49	M25 Other joint disorder, not elsewhere classified	3	57.07	0.001	701	116.83	99.44
60	M79.1 Myalgia	4	45.14	0.001	109	18.17	241.22
70	M79.6 Pain in limb, hand, foot, fingers and toes	4	40.67	0.001	211	35.17	158.74
80	M86-M90 Other osteopathies	2	30.38	0.001	95	15.83	209.54
90	M79.10 Myalgia, unspecified site	5	28.35	0.001	42	7.00	314.29
93	M50-M54 Other dorsopathies	2	27.90	0.001	220	36.67	126.34
97	M54 Dorsalgia	3	26.87	0.001	197	32.83	131.50
100	M60-M63 Disorders of muscles	2	26.43	0.001	375	62.50	92.00
102	M62 Other disorders of muscle	3	25.69	0.001	366	61.00	91.80
107	M25.50 Pain in unspecified joint	5	24.50	0.001	49	8.17	267.20
125	M79.60 Pain in limb, unspecified	5	22.11	0.001	96	16.00	175.00
137	M25.56 Pain in knee	5	19.22	0.001	174	29.00	117.24
142	M30-M36 Systemic connective tissue disorders	2	18.96	0.001	90	15.00	166.67
143	M62.8 Other specified disorders of muscle	4	18.80	0.001	325	54.17	82.76
153	M79.18 Myalgia, other site	5	17.23	0.001	64	10.67	190.53
162	M35 Other systemic involvement of connective tissue	3	16.47	0.001	38	6.33	247.55

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

164	M02.30 Reiter's disease, unspecified site	5	16.45	0.001	12	2.00	450.00
167	M00-M02 Infectious arthropathies	2	16.28	0.001	18	3.00	366.67
169	M24 Other specific joint derangements	3	16.00	0.001	147	24.50	116.33
183	M89 Other disorders of bone	3	14.60	0.001	38	6.33	231.75
186	M02.3 Reiter's disease	4	14.54	0.001	15	2.50	380.00
187	M02 Postinfective and reactive arthropathies	3	14.54	0.001	15	2.50	380.00
190	M99 Biomechanical lesions, not elsewhere classified	3	14.33	0.001	8	1.33	501.50
191	M99-M99 Biomechanical lesions, not elsewhere classified	2	14.33	0.001	8	1.33	501.50
192	M99.0 Segmental and somatic dysfunction	4	14.33	0.001	8	1.33	501.50
204	M62.81 Muscle weakness (generalized)	5	13.11	0.001	163	27.17	98.75
221	M79.66 Pain in lower leg	5	12.42	0.002	27	4.50	255.56
245	M35.7 Hypermobility syndrome	4	11.17	0.006	26	4.33	246.42
263	M86 Osteomyelitis	3	10.58	0.013	36	6.00	200.00
271	M24.9 Joint derangement, unspecified	4	10.27	0.013	19	3.17	278.55
277	M79.605 Pain in left leg	6	10.08	0.013	37	6.17	191.73
281	M79.662 Pain in left lower leg	6	9.92	0.015	12	2.00	350.00
287	M89.1 Physeal arrest	4	9.69	0.020	10	1.67	379.04
300	M25.561 Pain in right knee	6	9.34	0.024	90	15.00	113.33
301	M21 Other acquired deformities of limbs	3	9.30	0.024	205	34.17	72.67
N00-N99 Diseases of the genitourinary system							
130	N00-N99 Diseases of the genitourinary system	1	20.35	0.001	768	128.00	54.69

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

195	N17 Acute kidney failure	3	13.99	0.001	39	6.50	223.08
220	N76 Other inflammation of vagina and vulva	3	12.42	0.002	27	4.50	255.56
278	N17.9 Acute kidney failure, unspecified	4	9.99	0.013	34	5.67	199.82
293	N80-N98 Noninflammatory disorders of female genital tract	2	9.46	0.022	263	43.83	64.27
303	N70-N77 Inflammatory diseases of female pelvic organs	2	9.03	0.031	36	6.00	183.33
Q00-Q99 Congenital malformations, deformations and chromosomal abnormalities							
182	Q79 Congenital malformations of musculoskeletal system, not elsewhere classified	3	14.63	0.001	47	7.83	206.51
198	Q79.6 Ehlers-Danlos syndromes	4	13.45	0.001	43	7.17	206.83
202	Q00-Q99 Congenital malformations, deformations and chromosomal abnormalities	1	13.18	0.001	700	116.67	45.71
260	Q20-Q28 Congenital malformations of the circulatory system	2	10.62	0.012	93	15.50	119.35
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	1,327.81	0.001	6,736	1,122.67	160.72
6	R00-R09 Symptoms and signs involving the circulatory and respiratory systems	2	608.75	0.001	1,533	255.50	235.81
7	R06.0 Dyspnea	4	334.81	0.001	306	51.00	403.92
8	R50-R69 General symptoms and signs	2	326.26	0.001	1,791	298.50	153.94
9	R06 Abnormalities of breathing	3	313.01	0.001	546	91.00	287.91
19	R53 Malaise and fatigue	3	197.77	0.001	387	64.50	270.54
20	R53.8 Other malaise and fatigue	4	189.63	0.001	326	54.33	290.21
21	R10-R19 Symptoms and signs involving the digestive system and	2	154.74	0.001	1,400	233.33	117.29

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	abdomen						
22	R40-R46 Symptoms and signs involving cognition, perception, emotional state and behaviour	2	142.00	0.001	628	104.67	173.24
23	R07 Pain in throat and chest	3	137.91	0.001	209	34.83	310.57
24	R06.02 Shortness of breath	5	137.06	0.001	119	19.83	414.37
26	R53.83 Other fatigue	5	121.77	0.001	209	34.83	290.47
28	R06.00 Dyspnea, unspecified	5	103.08	0.001	92	15.33	408.81
31	R06.09 Other forms of dyspnea	5	96.99	0.001	71	11.83	449.45
34	R07.9 Chest pain, unspecified	4	88.76	0.001	120	20.00	330.00
37	R42 Dizziness and giddiness	3	81.90	0.001	155	25.83	275.53
40	R51 Headache	3	77.98	0.001	188	31.33	241.53
43	R51.9 Headache, unspecified	4	71.08	0.001	167	27.83	244.95
44	R00 Abnormalities of heart beat	3	70.07	0.001	160	26.67	248.71
45	R10 Abdominal and pelvic pain	3	62.08	0.001	564	94.00	117.02
47	R05 Cough	3	57.90	0.001	295	49.17	160.32
54	R25-R29 Symptoms and signs involving the nervous and musculoskeletal systems	2	52.31	0.001	321	53.50	144.86
56	R07.89 Other chest pain	5	47.31	0.001	62	10.33	335.62
58	R07.8 Other chest pain	4	46.89	0.001	72	12.00	308.33
61	R11 Nausea and vomiting	3	44.98	0.001	344	57.33	128.50
66	R00.2 Palpitations	4	41.77	0.001	65	10.83	306.28
68	R09 Other symptoms and signs involving the circulatory and respiratory system	3	41.14	0.001	238	39.67	149.56
72	R53.82 Chronic fatigue, unspecified	5	40.04	0.001	57	9.50	321.05
84	R06.8 Other abnormalities of breathing	4	29.75	0.001	188	31.33	142.58
86	R50 Fever of other and unknown	3	29.25	0.001	175	29.17	146.83

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

	origin						
87	R90-R96 Abnormal findings on diagnostic imaging and in function studies, without diagnosis	2	29.08	0.001	216	36.00	130.56
88	R00.0 Tachycardia, unspecified	4	28.77	0.001	68	11.33	244.22
89	R53.81 Other malaise	5	28.50	0.001	60	10.00	260.00
91	R70-R79 Abnormal findings on examination of blood, without diagnosis	2	28.31	0.001	230	38.33	124.37
92	R20-R23 Symptoms and signs involving the skin and subcutaneous tissue	2	28.25	0.001	200	33.33	134.02
95	R29 Other symptoms and signs involving the nervous and musculoskeletal systems	3	27.58	0.001	191	31.83	135.63
98	R05.9 Cough, unspecified	4	26.70	0.001	150	25.00	152.00
104	R55 Syncope and collapse	3	25.15	0.001	80	13.33	207.58
105	R43 Disturbances of smell and taste	3	24.64	0.001	36	6.00	316.67
108	R09.0 Asphyxia and hypoxemia	4	24.49	0.001	72	12.00	216.67
109	R09.02 Hypoxemia	5	24.49	0.001	72	12.00	216.67
110	R50.9 Fever, unspecified	4	24.46	0.001	142	23.67	149.26
113	R19 Other symptoms and signs involving the digestive system and abdomen	3	23.76	0.001	144	24.00	145.83
115	R30-R39 Symptoms and signs involving the urinary system	2	23.09	0.001	254	42.33	105.53
120	R41 Other symptoms and signs involving cognitive functions and awareness	3	22.70	0.001	122	20.33	155.78
127	R68 Other general symptoms and signs	3	20.63	0.001	103	17.17	162.09
132	R10.8 Other abdominal pain	4	20.08	0.001	167	27.83	122.78
136	R11.10 Vomiting, unspecified	5	19.32	0.001	162	27.00	122.22

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

140	R10.9 Unspecified abdominal pain	4	19.12	0.001	107	17.83	152.38
141	R10.84 Generalized abdominal pain	5	19.04	0.001	163	27.17	120.83
145	R06.83 Snoring	5	18.63	0.001	84	14.00	171.43
146	R13 Aphagia and dysphagia	3	18.48	0.001	250	41.67	94.38
147	R13.1 Dysphagia	4	18.48	0.001	250	41.67	94.38
150	R09.8 Other specified symptoms and signs involving the circulatory and respiratory systems	4	17.74	0.001	160	26.67	117.47
152	R20 Disturbances of skin sensation	3	17.38	0.001	34	5.67	270.37
154	R11.1 Vomiting	4	17.22	0.001	198	33.00	103.03
160	R46 Symptoms and signs involving appearance and behavior	3	16.74	0.001	103	17.17	144.61
165	R29.90 Unspecified symptoms and signs involving the nervous system	5	16.45	0.001	12	2.00	450.00
170	R46.8 Other symptoms and signs involving appearance and behavior	4	15.86	0.001	102	17.00	141.18
173	R68.8 Other general symptoms and signs	4	15.36	0.001	89	14.83	149.49
175	R10.1 Pain localized to upper abdomen	4	15.34	0.001	130	21.67	121.50
177	R68.89 Other general symptoms and signs	5	15.13	0.001	79	13.17	158.16
178	R11.2 Nausea with vomiting, unspecified	4	15.13	0.001	62	10.33	180.74
181	R29.9 Unspecified symptoms and signs involving the nervous and musculoskeletal systems	4	14.75	0.001	17	2.83	359.36
184	R11.0 Nausea	4	14.59	0.001	84	14.00	150.00
193	R19.7 Diarrhea, unspecified	4	14.33	0.001	92	15.33	141.36
197	R52 Pain, unspecified	3	13.61	0.001	49	8.17	193.76
201	R63 Symptoms and signs concerning food and fluid intake	3	13.24	0.001	274	45.67	75.17

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

205	R53.1 Weakness	4	13.11	0.001	60	10.00	170.00
206	R20.2 Paresthesia of skin	4	13.06	0.001	10	1.67	438.92
219	R13.10 Dysphagia, unspecified	5	12.46	0.002	98	16.33	126.58
227	R10.13 Epigastric pain	5	12.03	0.002	77	12.83	141.62
229	R04 Hemorrhage from respiratory passages	3	11.95	0.002	56	9.33	167.95
234	R05.3 Chronic cough	4	11.67	0.002	31	5.17	228.82
237	R31 Hematuria	3	11.49	0.002	47	7.83	180.97
243	R09.81 Nasal congestion	5	11.29	0.003	118	19.67	108.44
248	R26	3	11.09	0.006	29	4.83	231.26
251	R10.3 Pain localized to other parts of lower abdomen	4	10.98	0.006	148	24.67	94.57
265	R40 Somnolence, stupor and coma	3	10.52	0.013	33	5.50	209.09
266	R01 Cardiac murmurs and other cardiac sounds	3	10.52	0.013	27	4.50	233.33
269	R46.89 Other symptoms and signs involving appearance and behavior	5	10.35	0.013	94	15.67	116.98
273	R79 Other abnormal findings of blood chemistry	3	10.25	0.013	139	23.17	94.22
276	R41.8 Other symptoms and signs involving cognitive functions and awareness	4	10.10	0.013	91	15.17	117.53
279	R59 Enlarged lymph nodes	3	9.99	0.013	34	5.67	199.82
280	R43.9 Unspecified disturbances of smell and taste	4	9.92	0.015	12	2.00	350.00
282	R93 Abnormal findings on diagnostic imaging of other body structures	3	9.84	0.016	51	8.50	158.82
283	R46.81 Obsessive-compulsive behavior	5	9.71	0.017	8	1.33	426.32
289	R04.0 Epistaxis	4	9.65	0.020	48	8.00	162.50
299	R01.1 Cardiac murmur, unspecified	4	9.37	0.023	23	3.83	239.43

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

S00-T98 Injury, poisoning and certain other consequences of external causes							
302	T86 Complications of transplanted organs and tissue	3	9.19	0.027	60	10.00	140.00
U00-U95 Codes for special purposes							
2	U00-U49 Provisional assignment of new diseases of uncertain etiology or emergency use	2	853.05	0.001	481	80.17	498.73
3	U00-U95 Codes for special purposes	1	853.05	0.001	481	80.17	498.73
4	U09 Post COVID-19 condition	3	614.57	0.001	343	57.17	499.97
5	U09.9 Post COVID-19 condition, unspecified	4	614.57	0.001	343	57.17	499.97
16	U07	3	239.73	0.001	138	23.00	495.65
17	U07.1 Emergency use of U07.1 | COVID-19	4	239.73	0.001	138	23.00	495.65
V01-Y98 External causes of morbidity and mortality							
230	V01-Y98 External causes of morbidity and mortality	1	11.89	0.002	190	31.67	86.30
296	V00-V99 Transport accidents	2	9.41	0.022	156	26.00	84.62
Z00-Z99 Factors influencing health status and contact with health services							
83	Z00-Z99 Factors influencing health status and contact with health services	1	29.88	0.001	1,843	307.17	42.27
124	Z69-Z76 Persons encountering health services in other circumstances	2	22.15	0.001	258	43.00	102.33
156	Z99.8 Dependence on other enabling machines and devices	4	17.06	0.001	81	13.50	166.67
172	Z99.81 Dependence on supplemental oxygen	5	15.74	0.001	45	7.50	220.00
180	Z89 Acquired absence of limb	3	14.75	0.001	11	1.83	446.45
218	Z99 Dependence on enabling machines and devices, not elsewhere classified	3	12.52	0.002	133	22.17	107.49

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

259	Z77-Z99 Persons with potential health hazards related to family and personal history and certain conditions influencing health status	2	10.62	0.012	579	96.50	45.08

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

FUNDING STATEMENT

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.

COMPETING INTERESTS STATEMENT

Dr. Rao reports prior grant support from GSK and Biofire and is a consultant for Sequiris.

Dr. Jhaveri is a consultant for AstraZeneca, Seqirus, Dynavax, receives an editorial stipend from Elsevier and Pediatric Infectious Diseases Society and royalties from Up To Date/Wolters Kluwer.

Dr. Mejias reports funding from Janssen and Merck for research support; Janssen, Merck and Sanofi-Pasteur for Advisory Board participation, and Sanofi-Pasteru and AstraZeneca for CME lectures.

Dr. Patel reports funding from the National Institute of Health and Bayer Pharmaceuticals

Dr Brill received support from Novartis and Regeneron Pharmaceuticals within in the last year

Dr. Chen receives consulting support from GSK

Dr Bailey has received grants from Patient-Centered Outcomes Research Institute

All other authors have nothing to disclose.

CONTRIBUTORSHIP STATEMENT

Dr. Bailey, Dr. Lorman and Ms. Razzaghi had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

Dr. Lee, Dr. Lorman, and Ms. Razzaghi were responsible for the study conceptualization and design.

Dr. Lee, Dr. Lorman, Ms. Razzaghi, and Dr. Chen were involved in statistical analyses.

Dr. Lee, Dr. Lorman, and Ms. Razzaghi contributed to the drafting of the manuscript.

All authors contributed to interpretation of the data and review of the manuscript.

Dr. Rao, Dr. Lee, Dr. Bailey, and Dr. Forrest were involved in obtaining funding for the study.

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.

DATA AVAILABILITY STATEMENT

The results reported here are based on detailed individual-level patient data compiled as part of the RECOVER program. Due to the high risk of reidentification based on the number of unique patterns in the date, patient privacy regulations prohibit us from releasing the data publicly. The data are maintained in a secure enclave, with access managed by the program coordinating center to remain compliant with regulatory and program requirements. Please direct requests to access the data, either for reproduction of the work reported here or for other purposes, to recover@chop.edu.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions of Miranda Higginbotham.

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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. https://recovercovid.org. Accessed July 10, 2022 2022.
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. doi:https://doi.org/10.1111/apa.15870
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 February 13, 2023. 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 .

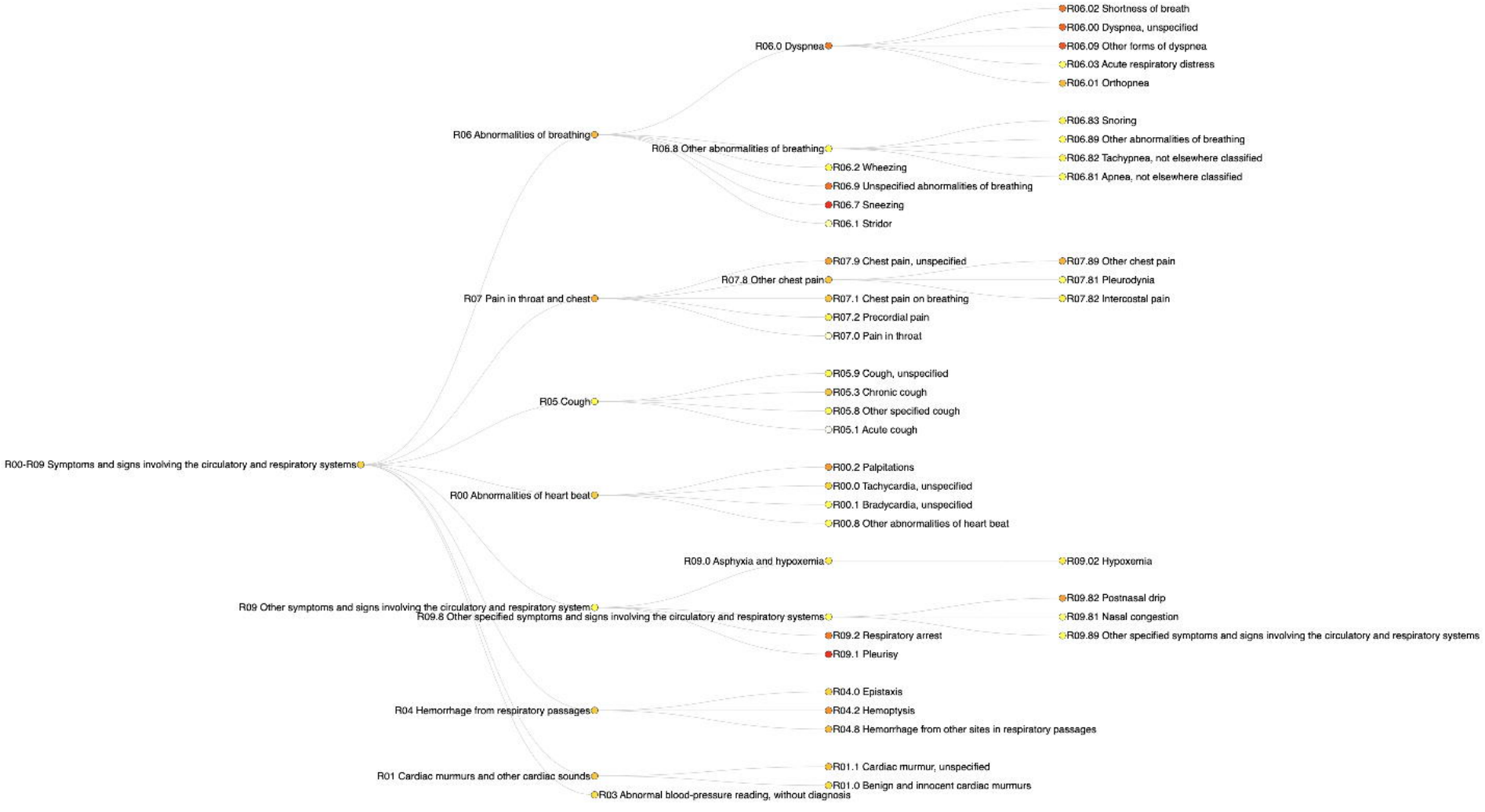
15. Thallapureddy K, Thallapureddy K, Zerda E, Suresh N, Kamat D, Rajasekaran K, et al. Long-Term Complications of COVID-19 Infection in Adolescents and Children. Curr Pediatr Rep. 2022;10(1):11-7. doi:10.1007/s40124-021-00260-x
16. Brackel CLH, Lap CR, Buddingh EP, van Houten MA, van der Sande L, Langereis EJ, et al. Pediatric long-COVID: An overlooked phenomenon? Pediatr Pulmonol. 2021;56(8):2495-502. doi:10.1002/ppul.25521
17. 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
18. aapmr. CDC Announces Approval of ICD-10 Code for Post-Acute Sequelae of COVID-19. 2021. https://www.aapmr.org/members-publications/member-news/member-news-details/2021/07/20/cdc-announces-approval-of-icd-10-code-for-post-acute-sequelae-of-covid-19. Accessed July 8, 2022.
19. 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
20. 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
21. 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
22. 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
23. 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
24. 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
25. Kulldorff M. TreeScan User Guide, v2.0. 2020.
26. CDC. ICD-10-CM Tabular List of Diseases and Injuries. 2022. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD10CM/2022/icd10cm-tabular-2022-April-1.pdf. Accessed August 1, 2022 2022.
27. 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
28. 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 February 13, 2023. 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 .

29. 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
30. 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
31. 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
32. 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
33. 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
34. 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
35. 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
36. 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
37. 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.
38. 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
39. 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
40. Goodwin Davies AJ, Xiao R, Razzaghi H, Bailey LC, Utidjian L, Gluck C, et al. Skeletal Outcomes in Children and Young Adults with Glomerular Disease. J Am Soc Nephrol. 2022;33(12):2233-46. doi:10.1681/ASN.2021101372
41. Cooper C, Dennison EM, Leufkens HG, Bishop N, van Staa TP. Epidemiology of childhood fractures in Britain: a study using the general practice research database. J Bone Miner Res. 2004;19(12):1976-81. doi:10.1359/JBMR.040902
42. Forrest CB, Burrows EK, Mejias A, Razzaghi H, Christakis D, Jhaveri R, et al. Severity of Acute COVID-19 in Children <18 Years Old March 2020 to December 2021. Pediatrics. 2022;149(4). doi:10.1542/peds.2021-055765

medRxiv preprint doi: https://doi.org/10.1101/2022.12.08.22283158; this version posted February 13, 2023. 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 .

43. Simon TD, Cawthon ML, Stanford S, Popalisky J, Lyons D, Woodcox P, et al. Pediatric medical complexity algorithm: a new method to stratify children by medical complexity. Pediatrics. 2014;133(6):e1647-54. doi:10.1542/peds.2013-3875


R00-R09 Symptoms and signs involving the circulatory and respiratory systems
R06 Abnormalities of breathing
R06.0 Dyspnea
R06.02 Shortness of breath
R06.00 Dyspnea, unspecified
R06.09 Other forms of dyspnea
R06.03 Acute respiratory distress
R06.01 Orthopnea
R06.8 Other abnormalities of breathing
R06.83 Snoring
R06.89 Other abnormalities of breathing
R06.82 Tachypnea, not elsewhere classified
R06.81 Apnea, not elsewhere classified
R06.2 Wheezing
R06.9 Unspecified abnormalities of breathing
R06.7 Sneezing
R06.1 Stridor
R07 Pain in throat and chest
R07.9 Chest pain, unspecified
R07.8 Other chest pain
R07.89 Other chest pain
R07.81 Pleurodynia
R07.82 Intercostal pain
R07.1 Chest pain on breathing
R07.2 Precordial pain
R07.0 Pain in throat
R05 Cough
R05.9 Cough, unspecified
R05.3 Chronic cough
R05.8 Other specified cough
R05.1 Acute cough
R00 Abnormalities of heart beat
R00.2 Palpitations
R00.0 Tachycardia, unspecified
R00.1 Bradycardia, unspecified
R00.8 Other abnormalities of heart beat
R09 Other symptoms and signs involving the circulatory and respiratory system
R09.0 Asphyxia and hypoxemia
R09.02 Hypoxemia
R09.8 Other specified symptoms and signs involving the circulatory and respiratory systems
R09.82 Postnasal drip
R09.81 Nasal congestion
R09.89 Other specified symptoms and signs involving the circulatory and respiratory systems
R09.2 Respiratory arrest
R09.1 Pleurisy
R04 Hemorrhage from respiratory passages
R04.0 Epistaxis
R04.2 Hemoptysis
R04.8 Hemorrhage from other sites in respiratory passages
R01 Cardiac murmurs and other cardiac sounds
R01.1 Cardiac murmur, unspecified
R01.0 Benign and innocent cardiac murmurs
R03 Abnormal blood-pressure reading, without diagnosis