

1 **Title:** *EHR-based Case Identification of Pediatric Long COVID: A Report from the RECOVER EHR Cohort*

2 **Authors:**

3 Morgan Botdorf, PhD¹, Kimberley Dickinson, MS¹, Vitaly Lorman PhD¹, Hanieh Razzaghi, PhD¹, Nicole Marchesani, MS¹, Suchitra
4 Rao, MBBS, MSCS², Colin Rogerson, MD³, Miranda Higginbotham, MHA¹, Asuncion Mejias, MD, PhD, MsCS⁴, Daria Salyakina,
5 PhD⁵, Deepika Thacker, MD⁶, Dima Dandachi, MD, MPH⁷, Dimitri A Christakis MD MPH⁸, Emily Taylor, M.A⁹, Hayden Schwenk,
6 MD, MPH¹⁰, Hiroki Morizono, PhD¹¹, Jonathan Cogen MD, MPH¹², Nathan M Pajor MD, MS¹³, Ravi Jhaveri, MD¹⁴, Christopher B.
7 Forrest, MD, PhD¹, L. Charles Bailey, MD, PhD¹, *on behalf of the RECOVER Consortium*

8
9 **Affiliations:** □ □ □

10 ¹ Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA □

11 ² Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital Colorado, Denver, CO

12 ³ Division of Critical Care, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

13 ⁴ Division of Infectious Diseases, Department of Pediatrics, Nationwide Children's Hospital and The Ohio State University, Columbus,
14 OH

15 ⁵ Center for Precision Medicine, Nicklaus Children's Hospital, Miami, FL

16 ⁶ Nemours Cardiac Center, Alfred I. duPont Hospital for Children, Wilmington, DE.

17 ⁷ Division of Infectious Diseases, Department of Medicine, University of Missouri-Columbia, Columbia, MO

18 ⁸ Center for Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA

19 ⁹ RECOVER Patient, Caregiver, or Community Representative New York, NY, USA

20 ¹⁰ Division of Pediatric Infectious Diseases, Stanford School of Medicine, Palo Alto, CA

21 ¹¹ Center for Genetic Medicine Research, Children's National Hospital, Washington, DC

22 ¹² Division of Pulmonary and Sleep Medicine, Department of Pediatrics, Seattle Children's Hospital, University of Washington, Seattle,
23 WA

24 ¹³ Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine,
25 Cincinnati OH

26 ¹⁴ Division of Infectious Diseases, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

27
28 **Authorship Statement**

29 Authorship has been determined according to ICMJE recommendations.

30
31 **Corresponding Author:**

32 Name: Morgan Botdorf

33 Affiliation: Applied Clinical Research Center, Children's Hospital of Philadelphia

34 Email: botdorfm@chop.edu

35 **Abbreviated Title:** Identifying Children with Long COVID

36 **Funding Source:** This research was funded by the National Institutes of Health (NIH) Agreement OTA OT2HL161847-01 as part of
37 the Researching COVID to Enhance Recovery (RECOVER) program of research.

38 **Disclaimer:** This content is solely the responsibility of the authors and does not necessarily represent the official views of the
39 RECOVER Initiative, the NIH, or other funders.

40

41 **Key Words:** PEDSnet, Post-acute sequelae SARS-CoV-2 infection, Long COVID, Chronic COVID-19 Syndrome, Late sequelae of
42 COVID-19, Long haul COVID, Long-term COVID-19, Post COVID syndrome, Post-acute COVID-19, Rule-based phenotyping,
43 Electronic health records, Electronic phenotyping, Chart review

44 **Abbreviations:** PASC—post-acute sequelae of SARS-CoV-2 infection; COVID-19—coronavirus disease 2019; SARS-CoV-2—
45 severe acute respiratory syndrome coronavirus 2; PCR—polymerase chain reaction; EHR—electronic health record; MIS-C—
46 multisystem inflammatory syndrome in children; ICD-10—International Classification of Diseases, version 10.

47

48 **Abstract**

49 **Objective**

50 Long COVID, marked by persistent, recurring, or new symptoms post-COVID-19 infection, impacts children's well-being yet lacks a
51 unified clinical definition. This study evaluates the performance of an empirically derived Long COVID case identification algorithm,
52 or computable phenotype, with manual chart review in a pediatric sample. This approach aims to facilitate large-scale research efforts
53 to understand this condition better.

54 **Methods**

55 The algorithm, composed of diagnostic codes empirically associated with Long COVID, was applied to a cohort of pediatric patients
56 with SARS-CoV-2 infection in the RECOVER PCORnet EHR database. The algorithm classified 31,781 patients with conclusive,
57 probable, or possible Long COVID and 307,686 patients without evidence of Long COVID. A chart review was performed on a
58 subset of patients (n=651) to determine the overlap between the two methods. Instances of discordance were reviewed to understand
59 the reasons for differences.

60 **Results**

61 The sample comprised 651 pediatric patients (339 females, $M_{age} = 10.10$ years) across 16 hospital systems. Results showed moderate
62 overlap between phenotype and chart review Long COVID identification (accuracy = 0.62, PPV = 0.49, NPV = 0.75); however, there
63 were also numerous cases of disagreement. No notable differences were found when the analyses were stratified by age at infection or
64 era of infection. Further examination of the discordant cases revealed that the most common cause of disagreement was the clinician

65 reviewers' tendency to attribute Long COVID-like symptoms to prior medical conditions. The performance of the phenotype
66 improved when prior medical conditions were considered (accuracy = 0.71, PPV = 0.65, NPV = 0.74).

67 **Conclusions**

68 Although there was moderate overlap between the two methods, the discrepancies between the two sources are likely attributed to the
69 lack of consensus on a Long COVID clinical definition. It is essential to consider the strengths and limitations of each method when
70 developing Long COVID classification algorithms.

71

72

73

74

75

76

77

78

79

80

81

82 **Introduction**

83 Long COVID, also known as post-acute sequelae of SARS-CoV-2 infection (PASC), is a significant health concern
84 characterized by ongoing, relapsing, or new symptoms emerging four or more weeks after the acute infection phase¹. While post-viral
85 syndromes like chronic fatigue syndrome following mononucleosis are well-documented in children²⁻³, understanding the clinical
86 manifestations of Long COVID in pediatric patients remains incomplete. The variability of symptoms in children compared to adults
87 complicates diagnosis and treatment⁴⁻⁹. Symptoms can range from fatigue and headache to loss of taste and smell and chest pain⁴⁻⁹.
88 Although rare, diagnosed conditions associated with Long COVID include myocarditis, myositis, postural tachycardia syndrome
89 (POTS), and myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), among other conditions¹⁰. Despite certain symptoms
90 and conditions clearly attributable to a SARS-CoV-2 infection, like multisystem inflammatory syndrome (MIS-C), much remains to
91 be understood about others¹¹⁻¹². These symptoms and conditions impose a substantial burden on children and their families, leading to
92 missed school and the need for service referrals¹³⁻¹⁴. This highlights the importance of improved detection and treatment strategies.

93 Identifying children who suffer from Long COVID in research studies is crucial to better understand this disorder and ensuring
94 timely detection and treatments in clinical settings. However, this task is challenging due to the inconsistency and heterogeneity of
95 associated symptoms. To address this challenge, researchers have used large observational cohort studies that use repositories of
96 electronic health record (EHR) data to identify patients^{5, 8, 9, 15, 16}. These studies have primarily relied on EHR-based diagnosis codes¹⁵⁻
97 ¹⁶. The ICD-10-CM U09.9 code, introduced in October 2021¹⁷⁻¹⁸, allows clinicians to assign a Long COVID diagnosis; however, its
98 utilization remains inconsistent and potentially biased across patients and healthcare settings¹⁶. Additionally, relying solely on this

99 code may not adequately capture all patients due to the variety of symptoms associated with Long COVID. This poses a risk of
100 misclassification if researchers exclusively use the U09.9 code for phenotyping.

101 To improve identification of patients with Long COVID, computable phenotyping techniques, which involve developing a set
102 of rules to identify patients with a disorder, have been used in Long COVID studies. Long COVID phenotypes for adult¹⁹ and
103 pediatric²⁰ patients have been developed using machine-learning approaches that leverage large numbers of clinical features. For
104 example, in a recent pediatric study, a machine learning algorithm demonstrated high precision in classifying both general and MIS-C-
105 specific forms of PASC, with recall rates of up to 70%²⁰. Training these supervised learning models requires a labeled cohort of
106 patients who likely have Long COVID based on healthcare utilization or Long COVID diagnosis codes. Since there is no gold-
107 standard definition of Long COVID, it is difficult to produce an unbiased labeled training set, which limits the generalizability of the
108 models.

109 In this study, we aimed to 1) identify children with Long COVID by utilizing a rules-based computable phenotype approach
110 and 2) assess the performance of this computable phenotype for Long COVID in a subset of children. This approach involves
111 analyzing specific diagnosis coding and symptoms that occur more frequently after a COVID-19 infection. By doing so, we can more
112 accurately identify a larger number of children with Long COVID. In addition, we have included clinician reviews of patient charts to
113 gain a comprehensive understanding of patients' experience with Long COVID. This combined approach represents a significant step
114 in the automation of Long COVID clinical phenotypes using EHR data in the absence of a consensus definition.

115 **Methods**

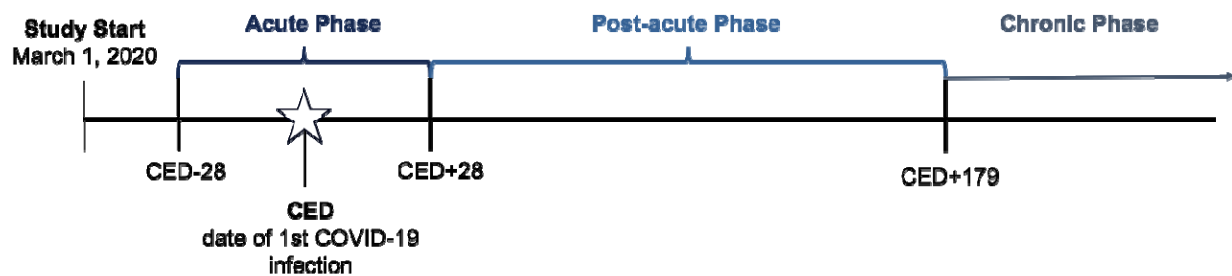
116 *Data Source*

117 This retrospective cohort study is part of the NIH Researching COVID to Enhance Recovery (RECOVER) Initiative, which
118 seeks to understand, treat, and prevent the post-acute sequelae of SARS-CoV-2 infection²¹. The RECOVER PCORnet EHR cohort
119 includes clinical data from patients in 40 hospital systems across the United States. Data were extracted from version 6 of the pediatric
120 RECOVER database, comprising more than 9 million children who were tested for SARS-CoV-2, diagnosed with COVID-19, or
121 received a COVID-19 vaccine between 2019 and December 2022. Institutional Review Board (IRB) approval was obtained under
122 Biomedical Research Alliance of New York (BRANY) protocol #21-08-508. BRANY waived the need for consent and HIPAA
123 authorization.

124 *Study Population*

125 Inclusion criteria for our pediatric sample were as follows: 1) SARS-CoV-2 infection confirmed via clinical diagnosis or PCR,
126 antigen, or qualifying serology test²² between March 2020 and December 2022, 2) age less than 21 years at first COVID-19 infection,
127 and 3) at least two contacts with the healthcare system (at least one being in-person or telehealth) to ensure adequate follow-up during
128 the post-acute phase (28-179 days following infection). We defined clinically meaningful time periods surrounding the initial COVID-
129 19 infection as shown in Figure 1. The acute phase spanned until the 27th day post-infection. The post-acute phase, which was the
130 primary focus of our analyses, spanned from day 28 through day 179 post-infection, ensuring that symptoms were not directly related
131 to the acute COVID-19 infection. For patients with a specific COVID-19 diagnosis or viral test, the initial infection date was the date

132 of diagnosis or test. For patients with diagnoses indicating “history of” or “complication of” COVID-19 or with a positive serology
133 test, we used 28 days prior to the earliest diagnosis or test evidence of COVID-19 as a proxy for initial infection date.



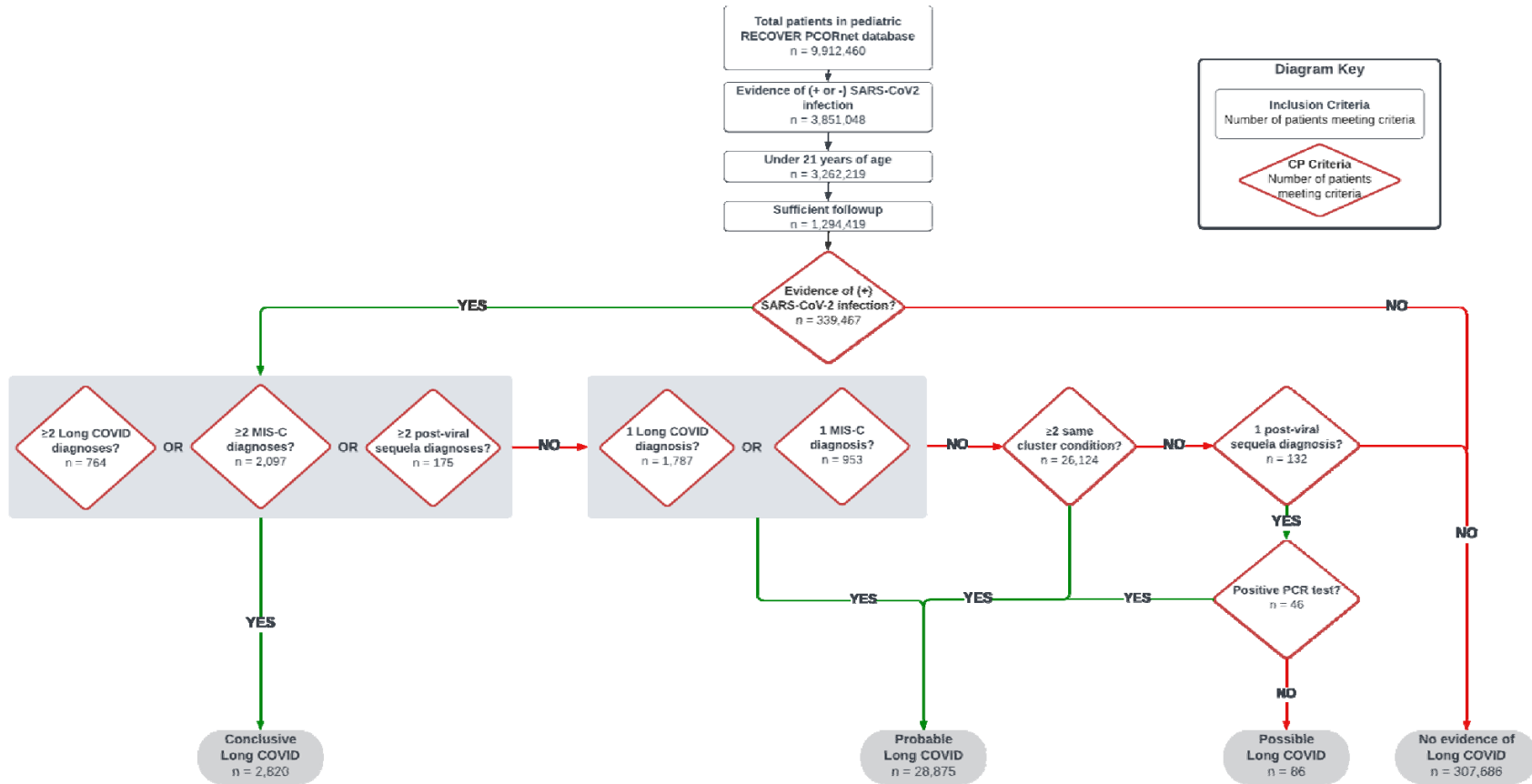
*CED = cohort entry date

134
135 **Figure 1. Study Timeline.**

136
137 *Phenotype classification*

138 Patients were identified as having conclusive, probable, or possible Long COVID according to the algorithm described in
139 Figure 2, which used criteria documented in the EHR in the post-acute period. The algorithm accounts for diagnoses of Long COVID
140 (ICD-10-CM code U09.9), diagnoses of MIS-C (ICD-10-CM code M35.81), diagnoses of sequelae of specified infectious and
141 parasitic diseases (ICD-10-CM code B94.8), and 23 diagnosis clusters identified as probable indicators of Long COVID based on our
142 prior work^{5,9} (Supplemental File 1). The diagnosis clusters were formed using a data mining approach that identified conditions more
143 common in U09.9-diagnosed patients than in non-U09.9 diagnosed COVID-19+ patients in the post-acute period⁵. Clinicians then

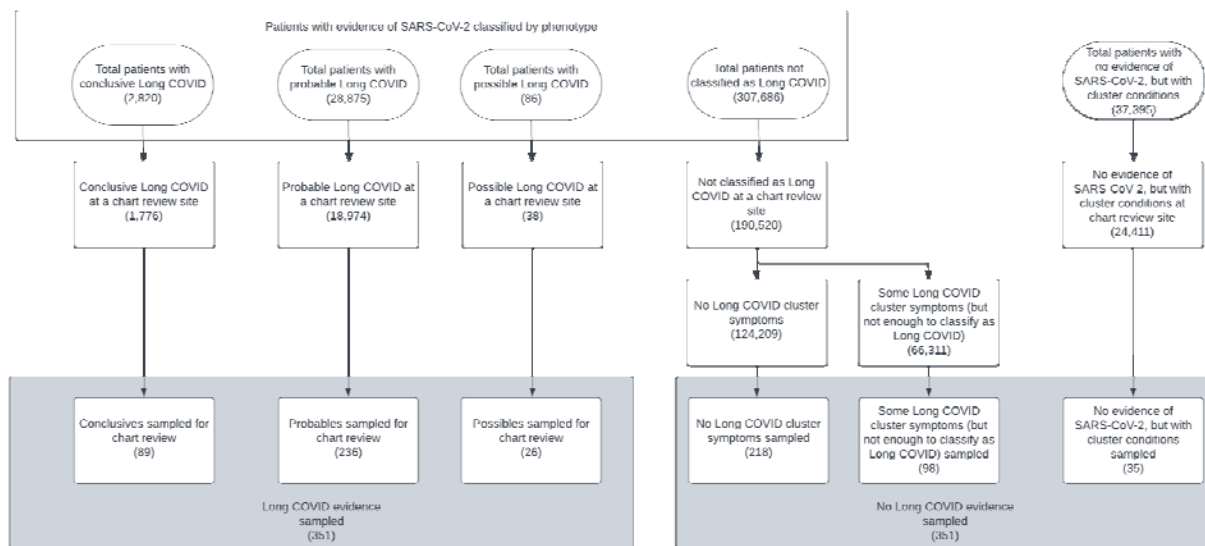
144 reviewed the diagnosis codes to create clusters of ICD-10-CM codes. Clusters included abdominal pain, abnormal liver enzymes,
145 acute kidney injury, acute respiratory distress syndrome, arrhythmias, autonomic dysfunction, cardiovascular signs/symptoms, changes
146 in taste/smell, chest pain, cognitive function, generalized pain, fatigue/malaise, fever, fluid/electrolyte balance, headache, heart
147 disease, myocarditis, musculoskeletal symptoms, myositis, respiratory signs/symptoms, and thrombophlebitis/thromboembolism. Any
148 patient with two or more diagnoses within the same cluster separated by at least 28 days during the post-acute period was labeled as
149 having probable Long COVID, regardless of whether the patient had a specific Long COVID or MIS-C diagnosis code. Figure 2
150 depicts the steps applied to classify patients according to the certainty of them having Long COVID. Any patient with conclusive,
151 probable, or possible Long COVID detected by the phenotype was labeled as “Long COVID Evidence” and all others were labelled as
152 “No Long COVID Evidence”.



153
154 **Figure 2.** Flow chart depicting attrition and algorithm used to identify patients with Long COVID.

155 *Chart Review Sampling*

156 A manual chart review was performed on a subset of the study population at 16
 157 institutions. We sampled 702 patients split between the Long COVID Evidence and No Long
 158 COVID Evidence groups to ensure there was adequate representation across sites. The sampling
 159 strategy is laid out in Figure 3. Approximately 22 Long COVID Evidence patients were
 160 randomly sampled per institution. Each Long COVID Evidence sampled patient was matched
 161 1:1 without replacement with a No Long COVID Evidence patient using exact matching on
 162 institution, age at time of infection, calendar quarter of infection, and acute period hospitalization
 163 (yes/no). Ninety percent of the No Long COVID Evidence sample had SARS-CoV-2 infection
 164 while the remaining ten percent (35 patients) were patients with no evidence of SARS-CoV-2 but
 165 with at least two diagnoses of cluster conditions separated by 28 to 150 days. The latter group
 166 were additional patients included in the chart review to gather insight on the attribution of cluster
 167 diagnoses to conditions other than SARS-CoV-2 infection. A total of 651 children were
 168 ultimately included in analyses based on additional exclusions which will be discussed. The
 169 sampling strategy is laid out in Figure 3.



171 **Figure 3.** *Sampling strategy for the chart review cohort. The numbers reported in parentheses*
172 *represent sample sizes.*

173

174 *Chart Review Procedure*

175 Clinical research teams from each participating institution conducted chart reviews using
176 a REDCap²³ (Research Electronic Data Capture) instrument with questions including
177 information on COVID-19 diagnoses and testing, demographics, COVID-19 prevention and
178 treatment strategies, vaccines, functional outcomes, and conditions post COVID-19 captured in
179 the patient's medical record. Each site had between 1 and 5 reviewers for a total of 44 reviewers
180 across sites. Table 1 contains a summary of patient information extracted from chart review, and
181 the full case report form is included in Supplemental File 2.

182 A secondary review was completed by a clinician who reviewed information extracted by
183 the primary chart reviewer and answered questions regarding the level of confidence with which
184 Long COVID could be assigned to the patient. The clinician was first asked if the patient met
185 criteria for Long COVID based on the NIH definition²¹ which describes Long COVID as signs,
186 symptoms, and conditions that continue or develop after initial COVID-19 or SARS-CoV-2
187 infection, are present four weeks [28 days] or more after the initial phase of infection; may be
188 multisystemic; and may present with a relapsing-remitting pattern and progression or worsening
189 over time, with the possibility of severe and life-threatening events even months or years after
190 infection. They were then asked if the patient met criteria for Long COVID based on the
191 computable phenotype definition. The response to these questions (i.e., conclusive, probable,
192 possible, no evidence) was used to assess concordance with the computable phenotype. The first

193 question, which analyses focused on, asked the clinician to exercise clinical judgment, while the
194 second question was focused on assessing the validity of the structured EHR data.

195 For ease of assessing the performance of the computable phenotype compared with chart
196 review, patients were collapsed into four overlapping groups: computable phenotype-positive
197 (CP-positive), computable phenotype-negative (CP-negative), clinician review-positive (CR-
198 positive), and clinician review-negative (CR-negative). Patients identified by the phenotype as
199 having conclusive or probable Long COVID were placed in the CP-positive group. Conversely,
200 patients found to have possible evidence or no evidence of Long COVID were placed in the CP-
201 negative group. Patients with possible Long COVID were included in the CP-negative sample as
202 the study team concluded that having only one post-viral sequelae code without a positive PCR
203 test to confirm SARS-CoV-2 infection did not provide enough evidence to conclude that the
204 patient's post-viral sequelae was caused by Long COVID. On the other hand, when examining
205 the chart review, we determined that the reasons reviewers used to classify patients as possible
206 for Long COVID were more similar to a positive than a negative Long COVID classification.
207 Therefore, the CR-positive group consisted of patients labeled as conclusive, probable, or
208 possible for Long COVID by the clinician reviewer. In contrast, the CR-negative group consisted
209 of patients labeled as having no evidence of Long COVID by the clinician reviewer.

210 *Performance Assessment*

211 The performance of the computable phenotype was evaluated across various key metrics
212 including sensitivity, specificity, positive predictive value (PPV), and negative predictive value
213 (NPV). Additionally, we examined the accuracy and F1 score of the phenotype. Accuracy
214 assesses the proportion of CR-positive patients who were also CP-positive. The F1 score
215 combines precision and recall providing insight into the overall effectiveness of the phenotype.

216 Next, we assessed whether concordance between the phenotype and clinician review
217 classification differed by age (i.e., under vs. over 12 years old), variant period (i.e., alpha, delta,
218 omicron), and number of symptom clusters through stratified analyses. Analyses were conducted
219 using R version 4.1.2 (2021-11-01; 24).

220 We conducted an assessment to identify discrepancies between the phenotype and chart
221 review identification of Long COVID. We reviewed cases where the phenotype identified Long
222 COVID, but the chart review did not, and vice versa. To understand the reasons behind these
223 discrepancies, we reviewed the chart review form for each discordant patient, along with the
224 clinician reviewer's explanation for assigning or not assigning Long COVID. We then generated
225 themes that accounted for the discrepancy and assigned those themes to the remaining cases. We
226 next aimed to modify our model based the most common themes and perform a sensitivity
227 analysis to assess whether the performance of our modified model was superior to our original
228 model.

229 To describe the sample and investigate the impact of patients with complex medical
230 histories on the performance of the phenotype, we used the Pediatric Medical Complexity
231 Algorithm²⁵ (PMCA). We determined the chronic condition status of each patient by applying
232 the more conservative version of the algorithm. This version requires that a patient have one
233 diagnosis of a progressive or malignant condition or at least two diagnoses per body system for
234 at least two body systems in the three years prior to the SARS-CoV-2 infection.

235 **Results**

236 Among patients with a positive SARS-CoV-2 infection (1,007,867), the computable
237 phenotype detected Long COVID in 31,781 (3.2%) patients. Seven hundred and two patients
238 were included in the chart review sample; however, sixteen charts were not completed due to

239 limitations in the chart reviewers' access to records and were excluded from the sample. In
 240 addition, the 35 patients with no evidence of SARS-CoV-2 infection according to the computable
 241 phenotype were not included in comparative analyses as a full chart review was not completed
 242 on them. Thus, our final sample consisted of 651 patients. Sample demographics and descriptive
 243 statistics are presented in Table 1. Sociodemographic characteristics were similar among those
 244 classified with and without Long COVID by the phenotype and by the clinician chart reviewer
 245 (Supplemental Table S1).

246
 247

248 **Table 1.** Demographics of children with and without Long COVID based on the computable
 249 phenotype definition.

| | CP-Detected Long COVID (N=318) | No CP-Detected Long COVID (N=333) | Overall (N=651) |
|--|--------------------------------------|---|--------------------|
| <i>Approx. CED age (years)</i> | | | |
| Mean (SD) | 10.10 (6.32) | 10.10 (6.28) | 10.10 (6.30) |
| Median [Min, Max] | 11.0 [0, 21.0] | 10.3 [0.1, 21.0] | 10.9 [0, 21.0] |
| <i>CED Age Group (years)</i> | | | |
| <1 | 24 (7.5%) | 25 (7.5%) | 49 (7.5%) |
| 1-4 | 70 (22.0%) | 76 (22.8%) | 146 (22.4%) |
| 5-9 | 52 (16.4%) | 55 (16.5%) | 107 (16.4%) |
| 10-15 | 98 (30.8%) | 100 (30.0%) | 198 (30.4%) |
| 16-20 | 74 (23.3%) | 77 (23.1%) | 151 (23.2%) |
| <i>Patient Sex</i> | | | |
| Male | 147 (46.2%) | 167 (50.2%) | 314 (48.2%) |
| Female | 171 (53.8%) | 166 (49.8%) | 337 (51.8%) |
| <i>Race</i> | | | |
| Asian/Native Hawaiian/Pacific Islander | 12 (3.8%) | 14 (4.2%) | 26 (4.0%) |
| Black | 57 (17.9%) | 58 (17.4%) | 115 (17.7%) |
| White | 175 (55.0%) | 182 (54.7%) | 357 (54.8%) |
| Multiracial | 12 (3.8%) | 9 (2.7%) | 21 (3.2%) |
| Unknown | 62 (19.5%) | 70 (21.0%) | 132 (20.3%) |

| <i>Ethnicity</i> | | | |
|------------------|-------------|-------------|-------------|
| Hispanic | 78 (24.5%) | 99 (29.7%) | 177 (27.2%) |
| Non-Hispanic | 214 (67.3%) | 210 (63.1%) | 424 (65.1%) |
| Unknown | 26 (8.2%) | 24 (7.2%) | 50 (7.7%) |
| <i>Payer*</i> | | | |
| Private | 139 (43.7%) | 143 (42.9%) | 282 (43.3%) |
| Public | 124 (39.0%) | 134 (40.2%) | 258 (39.6%) |
| Other/Unknown | 55 (17.3%) | 56 (16.8%) | 111 (17.1%) |

250 *Note. CP = computable phenotype. *=at time of COVID-19 infection.*

251

252 There was 73.33% agreement between the responses to the two questions the clinician

253 chart reviewers answered. Table 2 presents statistics assessing the performance of the

254 computable phenotype with chart review. The two methods had substantial but incomplete

255 overlap. Analyses assessing whether concordance differed by selected variables showed similar

256 results across age, era associated with infection, and number of symptom clusters (Supplemental

257 Tables S2-4).

258

259 **Table 2.** Statistics comparing computable phenotype and clinician review identification of Long

260 COVID.

| | CR-Positive (N = 239) | CR-Negative (N = 412) | | | | |
|--------------------------------|----------------------------------|----------------------------------|--------------------|------------|------------|-----------|
| CP-Positive (N = 318) | 156 | 162 | | | | |
| CP-Negative (N = 333) | 83 | 250 | | | | |
| Performance Statistics* | | | | | | |
| | Accuracy | Sensitivity | Specificity | PPV | NPV | F1 |
| | 0.624 | 0.653 | 0.607 | 0.491 | 0.751 | 0.560 |

261 *Note. PPV = positive predictive value. NPV = negative predictive value. CP = computable*

262 *phenotype. CR = chart review.*

263 **Reported as CP relative to CR.*

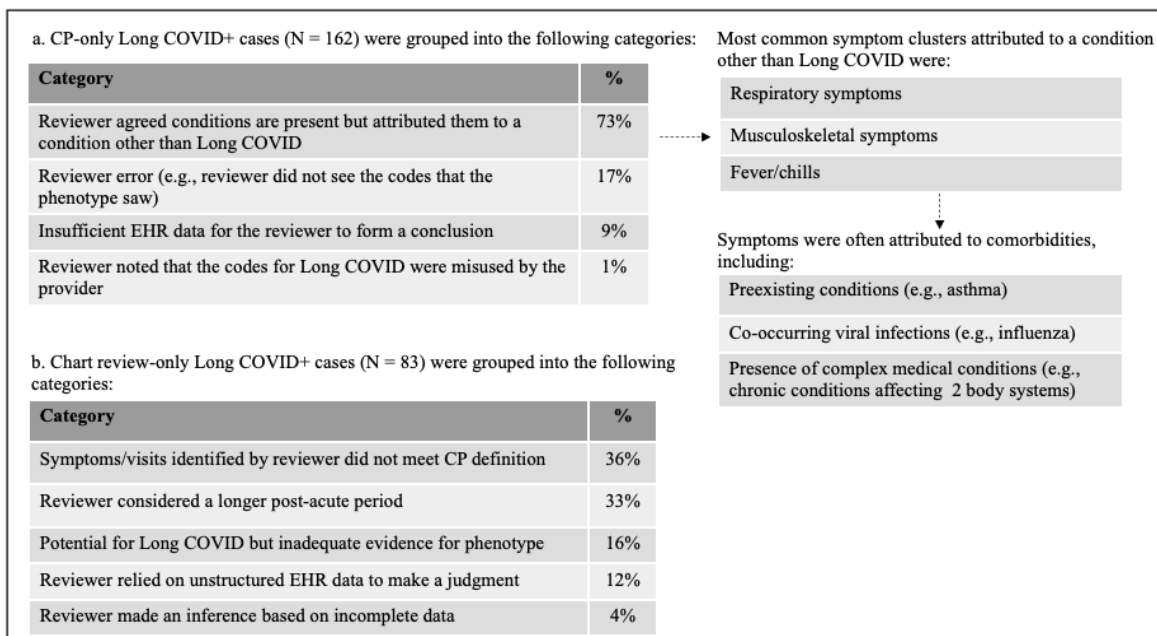
264

265 *Computable Phenotype-Only and Clinician Review-Only Long COVID Positive Review*

266 A review was conducted to assess the reasons for disagreement between the two methods.
267 The initial focus was cases where the phenotype identified Long COVID, but the chart review
268 did not (CP+/CR- cases) as there were many of these cases. Results are presented in Figure 4. In
269 most CP+/CR- cases, the clinician reviewer agreed with the symptoms the computable
270 phenotype identified but attributed those symptoms to another viral infection or preexisting
271 disorder (Figure 4a). This was especially true for symptoms common to other respiratory
272 infections and symptoms with occurrences both pre and post COVID-19 infection. In other less
273 common cases, the reviewer did not see the diagnostic codes that the phenotype saw, or the
274 reviewer made a conclusion based on incomplete information.

275 An assessment of CR+/CP- cases showed that in many cases the reviewer considered
276 symptoms, visits, and time frames that differed from our phenotype (Figure 4b). For example,
277 clinician reviewers considered symptoms beyond 180 days post-infection (up to 11 months in
278 some cases) while the computable phenotype considered symptoms in the one month to 180 days
279 following infection. In addition, reviewers considered symptoms beyond those included in the
280 computable phenotype definition, such as mental health symptoms. For example, clinician
281 reviewers used clinical judgment and identified anxiety as a qualifying symptom, but it is not
282 included in the phenotype definition due to inconsistent recording in structured data.

283



284
285
286
287
288

Figure 4. *Qualitative review of a.) CP+/CR- Long COVID patients and b.) CR+/CP- Long COVID patients.*

289 *Sensitivity Analysis*

290 The review of CP+/CR- patients showed that comorbidities were a large factor
291 contributing to discordance between the two methods. Given the difficulty of distinguishing
292 symptoms due to preexisting conditions and symptoms due to Long COVID, we performed a
293 sensitivity analysis to assess concordance with a modified model in which preexisting conditions
294 and comorbidities were accounted for. First, we censored prior symptoms. In other words, we
295 assessed phenotype classification when non-incident diagnoses were excluded. Patients who met
296 criteria for the computable phenotype with only preexisting symptom clusters that persisted after
297 their COVID-19 diagnosis were labeled as having no evidence of Long COVID (n = 32 patients).
298 Second, given the high prevalence of symptoms reported in the setting of non-COVID-19
299 respiratory infections, we excluded respiratory or fever cluster diagnoses that occurred 2 weeks
300 prior to or after a diagnosis of a non-COVID-19 respiratory infection. This led to exclusion of 31

301 patients who did not otherwise meet the criteria for probable long COVID. Finally, it was
302 difficult to attribute post-acute symptoms to a COVID-19 diagnosis versus underlying medical
303 conditions in patients with multiple chronic medical conditions. Therefore, we identified patients
304 with complex medical histories using the PMCA²⁵. We reclassified patients with a complex
305 chronic condition as having no evidence of Long COVID (n = 67 patients) as we believed that
306 our phenotype could not accurately identify these patients as having Long COVID.

307 Concordance was assessed again after incorporating these alterations into a modified
308 model, and results showed a higher positive predictive value and specificity but lower sensitivity
309 (Table 3). The negative predictive value remained high. Given the difficulty in adequately
310 attributing diagnoses to Long COVID in patients with complex medical histories, we also
311 completed a sensitivity analysis where these patients were removed from the sample, and
312 performance was reassessed (Table 3). Results showed similar performance to the modified
313 model overall, but a higher F1 score.

314

315

316

317 **Table 3.** Statistics comparing modified computable phenotype and clinician review identification
318 of Long COVID.

| | Modified | |
|---|----------------------------------|----------------------------------|
| | CR-Positive (N = 239) | CR-Negative (N = 412) |
| Modified CP-Positive (N=188) | 123 | 65 |
| Modified CP-Negative (N=463) | 116 | 347 |
| No Medically Complex Patients | | |
| | CR-Positive (N = 210) | CR-Negative (N = 308) |

| | | | | | | |
|---------------------------------------|-----------------|--------------------|--------------------|------------|------------|-----------|
| Modified CP-Positive (N = 188) | 123 | 65 | | | | |
| Modified-CP Negative (N = 330) | 87 | 243 | | | | |
| Performance Statistics* | | | | | | |
| | Accuracy | Sensitivity | Specificity | PPV | NPV | F1 |
| Modified | 0.722 | 0.515 | 0.842 | 0.654 | 0.749 | 0.576 |
| No Medically Complex Patients | 0.707 | 0.586 | 0.789 | 0.654 | 0.736 | 0.618 |

319 Note. PPV = positive predictive value. NPV = negative predictive value. CP = computable
320 phenotype. CR = chart review.

321 *Reported as CP relative to CR

322

323 Discussion

324 We conducted a study to assess the performance of a rules-based computable phenotype
325 for identifying pediatric patients with Long COVID in a large EHR database compared to
326 clinician chart review. Results showed moderate overlap between the two methods. Specifically,
327 the computable phenotype was moderately sensitive in detecting patients with Long COVID and
328 specific in detecting those without Long COVID, in comparison to chart review. However, there
329 were several cases where the methods disagreed, with some patients being classified as having
330 Long COVID by the phenotype but not by chart review, and vice versa. The main reason for
331 these discrepancies was due to underlying comorbidities and subsequent respiratory infections.

332 Patients with comorbidities posed a challenge for the computable phenotype and the
333 clinician reviewer. This was likely due to the lack of clinical guidelines for attribution and the
334 difficulty in discerning exacerbation of preexisting symptoms. Clinicians were more likely to
335 attribute post-COVID-19 symptoms to preexisting conditions when comorbidities were present,
336 which likely resulted in the misattribution of Long COVID symptoms and may have been
337 influenced by the provider involved in clinical care. However, the overlap between the two
338 methods increased when the CP accounted for preexisting medical conditions by focusing on

339 incident diagnoses and censoring existing conditions. Nevertheless, because our phenotype was
340 not initially designed to assess exacerbation of preexisting conditions, we caution against its use
341 to diagnose Long COVID in patients with medical complexity. Another source of disagreement
342 between the computable phenotype and chart review stemmed from subsequent non-COVID-19
343 respiratory infections, which are common in children. Although there may be an increased risk of
344 secondary infections due to SARS-CoV-2²⁶⁻²⁷, the symptoms are caused by a different agent.
345 Therefore, we removed these circumstances as indicators of Long COVID in our phenotype.

346 An analysis of chart review-only positives (i.e., those the clinician reviewer classified as
347 having Long COVID that the phenotype did not) showed that differences in the computable
348 phenotype guidelines and the clinician's framework for identifying Long COVID were the main
349 reasons for discordance. While the computable phenotype only assessed symptoms up to 180
350 days post SARS-CoV-2 infection, clinicians may use a longer time window in practice. This
351 suggests an extended time frame for assessing post-acute symptoms may be necessary, but also
352 may increase the risk of later onset of symptoms not being clearly attributable to a SARS-CoV-2
353 infection. In addition, clinician reviewers identified conditions beyond those included in our
354 phenotype as providing evidence for Long COVID. For example, the computable phenotype did
355 not consider mental health conditions due to reporting inconsistency of these conditions using
356 diagnosis codes and the difficulty in distinguishing between biologic and social causes of mental
357 health conditions. However, clinicians tended to include them in their framework for identifying
358 Long COVID. Therefore, constructing computable phenotypes that incorporate subphenotypes of
359 interest (e.g., physiological vs psychological manifestations of Long COVID) may be useful in
360 accounting for different manifestations of Long COVID.

361 Many of the differences in identifying Long COVID are due to the lack of a clear and
362 consistently used definition for Long COVID. The novelty of Long COVID in children, as well
363 as the overlap of symptoms with other acute and chronic disorders such as headache and fatigue,
364 contribute to these differences. Similar difficulties have been encountered when defining other
365 post-viral syndromes. Although chart review is often viewed as the best method for detecting
366 patients with a specific disorder, the lack of a consistent definition for Long COVID by
367 healthcare providers poses significant challenges. Moreover, the chart review and EHR research
368 are prone to errors such as biases, missed codes, misdiagnoses, incomplete information due to
369 fragmented care, and a lack of availability of a unique medical code for certain Long COVID-
370 related conditions. For example, a unique ICD code for POTS did not exist until October 2022²⁸,
371 and many clinicians remain unaware of its existence, making it difficult to pick up the presence
372 of POTS in the phenotype or chart review. Therefore, comparing our computable phenotype with
373 chart review provided insight into the clinician's view of a patient's status, but it did not allow us
374 to validate against a gold standard, as we cannot confidently conclude that either method
375 accurately detects Long COVID.

376 Our two-pronged approach to identifying Long COVID using clinician chart review and a
377 computable phenotype is a strength of the study as previous research that used diagnosis codes or
378 machine learning algorithms did not incorporate a review of patient charts¹⁹⁻²⁰. By incorporating
379 both methods, we were able to qualitatively review cases of discordance. In addition, we focused
380 on pediatric patients, in whom Long COVID is understudied. Research suggests that Long
381 COVID has a lower prevalence in children; however, the current diagnostic tools may not be
382 sensitive enough to detect all cases. Our study design is a strength as it uses data-driven
383 symptom clusters for identifying Long COVID and specific diagnosis codes. This approach

384 allowed us to capture patients who may not have a clear Long COVID presentation but have
385 Long COVID-like symptoms. Although it allowed for our phenotype to be more inclusive, it also
386 resulted in a computable phenotype that was less inclusive than a clinician may be when
387 subjectively assessing whether a patient has Long COVID. As more symptoms of Long COVID
388 are identified, it may be necessary to update the symptom clusters.

389 Our study has limitations, but it also brings attention to some significant areas for future
390 work to focus on. Due to the challenge of distinguishing the progression of a chronic condition
391 from symptom exacerbation due to COVID-19 using EHR data and chart review, we were
392 unable to evaluate the worsening of preexisting symptoms. Instead, we examined differences in
393 concordance after excluding pre-COVID-19 symptoms. This approach provided increased
394 certainty that symptoms were due to Long COVID but may have been too restrictive. Future
395 research should consider cluster-specific washout windows and develop reliable methods to
396 identify patients with Long COVID-related worsening of preexisting conditions. Additionally,
397 our sampling strategy focused on edge cases and rare occurrences to develop and refine the
398 phenotype. This approach was useful for identifying patients with a range of Long COVID-
399 related symptoms and diagnoses, but limits generalizability and underestimates the performance
400 of the phenotype. Future iterations should use random sampling to obtain a more generalizable
401 patient sample. Finally, because our study was based on EHR data and we imposed a two-visit
402 requirement in the post-acute period, our sample may be biased towards patients who have the
403 means to obtain healthcare at the population level.

404 **Conclusion**

405 This study describes a computable phenotype approach to identify children with Long
406 COVID in EHR data. Our study highlights the complexity of identifying and diagnosing Long

407 COVID due to its heterogeneity and overlap with other conditions, which leads to substantial
408 differences observed across methods. To address this challenge, future work could include
409 additional data sources, such as unstructured data, and further refine algorithms with clinical
410 expertise to develop a reliable definition of Long COVID. It is also essential to develop a revised
411 phenotype that can identify Long COVID through the worsening of pre-existing conditions. The
412 development of a reliable CP for Long COVID in children allows for studying large data
413 networks, which has future applications for both observational studies and clinical trials. Further
414 research assessing the presentation of Long COVID in children and the interplay between Long
415 COVID and comorbidities is vital to continue to understand this emerging chronic illness and
416 evaluate interventions that can prevent or mitigate its effects.

417

418 **Acknowledgements:**

419 This study is part of the NIH Researching COVID to Enhance Recovery (RECOVER) Initiative,
420 which seeks to understand, treat, and prevent the post-acute sequelae of SARS-CoV-2 infection
421 (PASC). For more information on RECOVER, visit <https://recovercovid.org/>

422

423 We would like to thank the National Community Engagement Group (NCEG), all patient,
424 caregiver, and community Representatives, and all the participants enrolled in the RECOVER
425 Initiative.

426

427 **Author Conflict of Interest Disclosures:** Dr. Mejias reports funding from Janssen, Merck for
428 research support, and Janssen, Merck and Sanofi-Pasteur for Advisory Board participation; Dr.
429 Rao reports prior grant support from GSK and Biofire and is a consultant for Sequiris. Dr.
430 Jhaveri is a consultant for AstraZeneca, Seqirus and Dynavax, and receives an editorial stipend
431 from Elsevier. All other authors have no conflicts of interest to disclose.

432 **Role of funder/sponsor statement:** The funder had no role in the design and conduct of the
433 study; collection, management, analysis, and interpretation of the data; preparation, review, or
434 approval of the manuscript; and decision to submit the manuscript for publication.

435

436

437

References

- 438 1. Davis, H. E., Assaf, G. S., McCorkell, L., Wei, H., Low, R. J., Re'em, Y., Redfield, S.,
439 Austin, J. P., & Akrami, A. (2021). Characterizing long COVID in an international
440 cohort: 7 months of symptoms and their impact. *eClinicalMedicine*, 38, 101019.
441 <https://doi.org/10.1016/j.eclinm.2021.101019>
- 442 2. Katz, B. Z., Shiraishi, Y., Mears, C. J., Binns, H. J., & Taylor, R. (2009). Chronic Fatigue
443 Syndrome After Infectious Mononucleosis in Adolescents. *Pediatrics*, 124(1), 189–193.
444 <https://doi.org/10.1542/peds.2008-1879>
- 445 3. Sellers, S. A., Hagan, R. S., Hayden, F. G., & Fischer, W. A. (2017). The hidden burden of
446 influenza: A review of the extra-pulmonary complications of influenza infection.
447 *Influenza and Other Respiratory Viruses*, 11(5), 372–393.
448 <https://doi.org/10.1111/irv.12470>
- 449 4. Lopez-Leon, S., Wegman-Ostrosky, T., Ayuzo Del Valle, N. C., Perelman, C.,
450 Sepulveda, R., Rebolledo, P. A., Cuapio, A., & Villapol, S. (2022). Long-COVID in
451 children and adolescents: A systematic review and meta-analyses. *Scientific Reports*,
452 12(1), 9950. <https://doi.org/10.1038/s41598-022-13495-5>

- 453 5. Lorman, V., Rao, S., Jhaveri, R., Case, A., Mejias, A., Pajor, N. M., Patel, P., Thacker,
454 D., Bose-Brill, S., Block, J., Hanley, P. C., Prahalad, P., Chen, Y., Forrest, C. B., Bailey,
455 L. C., Lee, G. M., & Razzaghi, H. (2023). Understanding pediatric long COVID using a
456 tree-based scan statistic approach: An EHR-based cohort study from the RECOVER
457 Program. *JAMIA Open*, 6(1), ooad016. <https://doi.org/10.1093/jamiaopen/oad016>
- 458 6. Zheng, Y.-B., Zeng, N., Yuan, K., Tian, S.-S., Yang, Y.-B., Gao, N., Chen, X., Zhang,
459 A.-Y., Kondratiuk, A. L., Shi, P.-P., Zhang, F., Sun, J., Yue, J.-L., Lin, X., Shi, L.,
460 Lalvani, A., Shi, J., Bao, Y.-P., & Lu, L. (2023). Prevalence and risk factor for long
461 COVID in children and adolescents: A meta-analysis and systematic review. *Journal of*
462 *Infection and Public Health*, 16(5), 660–672. <https://doi.org/10.1016/j.jiph.2023.03.005>
- 463 7. Pellegrino, R., Chiappini, E., Licari, A., Galli, L., & Marseglia, G. L. (2022). Prevalence
464 and clinical presentation of long COVID in children: A systematic review. *European*
465 *Journal of Pediatrics*, 181(12), 3995–4009. <https://doi.org/10.1007/s00431-022-04600-x>
- 466 8. Reese J, Blau H, Bergquist T, Loomba JJ, Callahan T, Laraway B, et al. Generalizable
467 Long COVID Subtypes: Findings from the NIH N3C and RECOVER Program. *MedRxiv*
468 *Prepr Serv Health Sci* 2022:2022.05.24.22275398. <https://doi.org/10.1101/2022.05.24.22275398> PMID: 35665012
- 469
- 470 9. Rao, S., Lee, G. M., Razzaghi, H., Lorman, V., Mejias, A., Pajor, N. M., Thacker, D.,
471 Webb, R., Dickinson, K., Bailey, L. C., Jhaveri, R., Christakis, D. A., Bennett, T. D.,
472 Chen, Y., & Forrest, C. B. (2022). Clinical Features and Burden of Postacute Sequelae of
473 SARS-CoV-2 Infection in Children and Adolescents. *JAMA Pediatrics*, 176(10), 1000.
474 <https://doi.org/10.1001/jamapediatrics.2022.2800>

- 475 10. Rao, S., Gross, R. S., Mohandas, S., Stein, C. R., Case, A., Dreyer, B., Pajor, N. M.,
476 Bunnell, H. T., Warburton, D., Berg, E., Overdeest, J. B., Gorelik, M., Milner, J.,
477 Saxena, S., Jhaveri, R., Wood, J. C., Rhee, K. E., Letts, R., Maughan, C., ... Stockwell,
478 M. S. (2024). Postacute Sequelae of SARS-CoV-2 in Children. *Pediatrics*, *153*(3),
479 e2023062570. <https://doi.org/10.1542/peds.2023-062570>
- 480 11. Feldstein, L. R., Rose, E. B., Horwitz, S. M., Collins, J. P., Newhams, M. M., Son, M. B.
481 F., Newburger, J. W., Kleinman, L. C., Heidemann, S. M., Martin, A. A., Singh, A. R.,
482 Li, S., Tarquinio, K. M., Jaggi, P., Oster, M. E., Zackai, S. P., Gillen, J., Ratner, A. J.,
483 Walsh, R. F., ... Randolph, A. G. (2020). Multisystem Inflammatory Syndrome in U.S.
484 Children and Adolescents. *New England Journal of Medicine*, *383*(4), 334–346.
485 <https://doi.org/10.1056/NEJMoa2021680>
- 486 12. Wu, E. Y., & Campbell, M. J. (2021). Cardiac Manifestations of Multisystem
487 Inflammatory Syndrome in Children (MIS-C) Following COVID-19. *Current Cardiology*
488 *Reports*, *23*(11), 168. <https://doi.org/10.1007/s11886-021-01602-3>
- 489 13. Garai, R., Krivácsy, P., Herczeg, V., Kovács, F., Tél, B., Kelemen, J., Máthé, A., Zsáry,
490 E., Takács, J., Veres, D. S., & Szabó, A. J. (2023). Clinical assessment of children with
491 long COVID syndrome. *Pediatric Research*, *93*(6), 1616–1625.
492 <https://doi.org/10.1038/s41390-022-02378-0>
- 493 14. Kikkenborg Berg, S., Palm, P., Nygaard, U., Bundgaard, H., Petersen, M. N. S.,
494 Rosenkilde, S., Thorsted, A. B., Ersbøll, A. K., Thygesen, L. C., Nielsen, S. D., &
495 Vinggaard Christensen, A. (2022). Long COVID symptoms in SARS-CoV-2-positive
496 children aged 0–14 years and matched controls in Denmark (LongCOVIDKidsDK): A

- 497 national, cross-sectional study. *The Lancet Child & Adolescent Health*, 6(9), 614–623.
- 498 [https://doi.org/10.1016/S2352-4642\(22\)00154-7](https://doi.org/10.1016/S2352-4642(22)00154-7)
- 499 15. Fritsche, L. G., Jin, W., Admon, A. J., & Mukherjee, B. (2023). Characterizing and
500 Predicting Post-Acute Sequelae of SARS CoV-2 Infection (PASC) in a Large Academic
501 Medical Center in the US. *Journal of Clinical Medicine*, 12(4), 1328.
502 <https://doi.org/10.3390/jcm12041328>
- 503 16. Pfaff, E. R., Madlock-Brown, C., Baratta, J. M., Bhatia, A., Davis, H., Girvin, A., Hill,
504 E., Kelly, E., Kostka, K., Loomba, J., McMurry, J. A., Wong, R., Bennett, T. D., Moffitt,
505 R., Chute, C. G., Haendel, M., The N3C Consortium, & The RECOVER Consortium.
506 (2023). Coding long COVID: Characterizing a new disease through an ICD-10 lens.
507 *BMC Medicine*, 21(1), 58. <https://doi.org/10.1186/s12916-023-02737-6>
- 508 17. International Classification of Diseases, Tenth Revision (ICD-10), World Health
509 Organization (WHO) 2019/2021. <https://icd.who.int/browse11>. Licensed under Creative
510 Commons Attribution-NoDerivatives 3.0 IGO license (CC BY-ND 3.0 IGO).
- 511 18. Centers for Disease Control and Prevention. Post-covid conditions: Information for
512 healthcare providers. Accessed January 5, 2024. [https://www.cdc.gov/coronavirus/2019-](https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/post-covid-conditions.html)
513 [ncov/hcp/clinical-care/post-covid-conditions.html](https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/post-covid-conditions.html)
- 514 19. Pfaff, E. R., Girvin, A. T., Bennett, T. D., Bhatia, A., Brooks, I. M., Deer, R. R.,
515 Dekermanjian, J. P., Jolley, S. E., Kahn, M. G., Kostka, K., McMurry, J. A., Moffitt, R.,
516 Walden, A., Chute, C. G., Haendel, M. A., Bramante, C., Dorr, D., Morris, M., Parker, A.
517 M., ... Niehaus, E. (2022). Identifying who has long COVID in the USA: A machine

- 518 learning approach using N3C data. *The Lancet Digital Health*, 4(7), e532–e541.
- 519 [https://doi.org/10.1016/S2589-7500\(22\)00048-6](https://doi.org/10.1016/S2589-7500(22)00048-6)
- 520 20. Lorman, V., Razzaghi, H., Song, X., Morse, K., Utidjian, L., Allen, A. J., Rao, S.,
- 521 Rogerson, C., Bennett, T. D., Morizono, H., Eckrich, D., Jhaveri, R., Huang, Y., Ranade,
- 522 D., Pajor, N., Lee, G. M., Forrest, C. B., & Bailey, L. C. (2023). A machine learning-
- 523 based phenotype for long COVID in children: An EHR-based study from the RECOVER
- 524 program. *PLOS ONE*, 18(8), e0289774. <https://doi.org/10.1371/journal.pone.0289774>
- 525 21. RECOVER: Researching COVID to Enhance Recovery. RECOVER: Researching
- 526 COVID to Enhance Recovery. <https://recovercovid.org>. Accessed 4 Jan 2024.
- 527 22. Mejias A, Schuchard J, Rao S, Bennett TD, Jhaveri R, Thacker D, Bailey LC, Christakis
- 528 DA, Pajor NM, Razzaghi H, Forrest CB, Lee GM. Leveraging Serologic Testing to
- 529 Identify Children at Risk for Post-Acute Sequelae of SARS-CoV-2 Infection: An
- 530 Electronic Health Record-Based Cohort Study from the RECOVER Program. *J Pediatr*.
- 531 2023 Jun; 257:113358. doi: 10.1016/j.jpeds.2023.02.005.
- 532 23. PA Harris, R Taylor, R Thielke, J Payne, N Gonzalez, JG. Conde, Research electronic
- 533 data capture (REDCap) – A metadata-driven methodology and workflow process for
- 534 providing translational research informatics support, *J Biomed Inform*. 2009
- 535 *Apr*;42(2):377-81.
- 536 24. Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software,
- 537 PBC, Boston, MA. URL <http://www.posit.co/>.

- 538 25. Simon, T. D., Lawrence, M., Stanford, S., Lyons, D., Woodcox, P., Hood, M., & Chen,
539 A. Y. (2014). Pediatric Medical Complexity Algorithm: A New Method to Stratify
540 Children by Medical Complexity. *Pediatrics*, 133(6).
- 541 26. Wang, L., Davis, P. B., Berger, N., Kaelber, D. C., Volkow, N., & Xu, R. (2023).
542 Association of COVID-19 with respiratory syncytial virus (RSV) infections in children
543 aged 0-5 years in the USA in 2022: a multicenter retrospective cohort study. *Family*
544 *medicine and community health*, 11(4), e002456. [https://doi.org/10.1136/fmch-2023-](https://doi.org/10.1136/fmch-2023-002456)
545 [002456](https://doi.org/10.1136/fmch-2023-002456)
- 546 27. Xie Y, Choi T, Al-Aly Z. Long-term outcomes following hospital admission for COVID-
547 19 versus seasonal influenza: a cohort study. *Lancet Infect Dis*. 2023 Dec 14: S1473-
548 3099(23)00684-9. doi: 10.1016/S1473-3099(23)00684-9.
- 549 28. World Health Organization. (2019). 8D89.2 Postural orthostatic tachycardia syndrome.
550 In *International statistical classification of diseases and related health problems* (11th
551 ed.). <https://icd.who.int/browse/2024-01/mms/en#1533647472>

552

553 **Supporting Information**

554 **Table S1.** Demographics of children with and without Long COVID based on the computable
555 phenotype definition.

556 **Table S2.** Statistics comparing CP and chart review identification of Long COVID presented
557 separately for patients younger and older than 12 years of age.

558 **Table S3.** Start and end dates associated with each era of infection.

559 **Table S4.** Statistics comparing CP and chart review identification of Long COVID presented
560 separately by era of infection.

561 **Table S5.** Statistics comparing CP and chart review identification of Long COVID presented
562 separately by the number of clusters identified by the CP.

563