

1 **Clinical Subphenotypes of Multisystem Inflammatory Syndrome in Children: An EHR-**
2 **based cohort study from the RECOVER program**

3

4 **Authors:** Suchitra Rao, MBBS^{1*}, Naimin Jing, PhD^{2*}, Xiaokang Liu, PhD², Vitaly Lorman, PhD³,
5 Mitchell Maltenfort, PhD³, Julia Schuchard, PhD³, Qiong Wu, PhD², Jiayi Tong, MS², Hanieh
6 Razzaghi, MPH³, Asuncion Mejias, MD⁴, Grace M. Lee, MD⁵; Nathan M. Pajor, MD⁶, Grant S.
7 Schulert, MD⁷, Deepika Thacker, MD⁸, Ravi Jhaveri⁹, MD, Dimitri A. Christakis, MD¹⁰, L. Charles
8 Bailey, MD^{3,11}, Christopher B. Forrest, MD^{3,11}, Yong Chen, PhD² on behalf of the RECOVER
9 consortium

10 *SR and NJ contributed to the manuscript equally

11 **Affiliations:**

12 ¹Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital
13 Colorado, Aurora, CO

14 ²Department of Biostatistics, Epidemiology and Informatics, the Perelman School of Medicine,
15 University of Pennsylvania, Pennsylvania, PA

16 ³Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA

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

19 ⁵Department of Pediatrics (Infectious Diseases), Stanford University School of Medicine,
20 Stanford, CA

21 ⁶Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center and University
22 of Cincinnati College of Medicine, Cincinnati, OH

23 ⁷Division of Rheumatology, Cincinnati Children's Hospital Medical Center and Department of
24 Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH

25 ⁸Division of Cardiology, Nemours Children's Health, Wilmington, DE

26 ⁹Division of Infectious Diseases, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago,
27 IL

28 ¹⁰Center for Child Health, Behavior and Development, Seattle Children's Hospital, Seattle, WA

29 ¹¹Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania,
30 Philadelphia, PA

31 **Corresponding Author:**

32 Suchitra Rao, MBBS, MSCS

33 Associate Professor of Pediatrics (Infectious Diseases/Hospital Medicine/Epidemiology),
34 University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, CO

35 13123 E 16th Ave Box 090

36 Aurora, CO 80045

37 Email: suchitra.rao@childrenscolorado.org

38 **Co-Corresponding Author:**

39 Yong Chen, Ph.D.

40 Professor of Biostatistics

41 Perelman School of Medicine, University of Pennsylvania

42 423 Guardian Dr, Blockley Hall 602, Philadelphia, PA 19104

43 Email: ychen123@penncmedicine.upenn.edu

44 **Abbreviated Title:** MIS-C Subphenotypes

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

48 **Disclaimer:** The views and conclusions contained in this document are those of the authors
49 and should not be interpreted as representing the official policies, either expressed or implied, of
50 the NIH.

51 **Key Words:** Multisystem inflammatory syndrome in children, PEDSnet, post-acute sequelae
52 COVID, latent class analyses

53 **Abbreviations:** LCA-- latent class analyses; MIS-C- multisystem inflammatory syndrome in
54 children; PASC--post-acute sequelae of SARS-CoV-2 infection; COVID-19--coronavirus disease
55 2019; SARS-CoV-2--severe acute respiratory syndrome coronavirus 2; PCR--polymerase chain
56 reaction; EHR--electronic health record; ED--emergency department; UC--urgent care; ICU—
57 intensive care unit; CI--confidence intervals; ICD-10--International Classification of Diseases,
58 version 10; PMCA-- Pediatric Medical Complexity Algorithm

59 **Word Count:** 2492

60

61 **Research in context**

62 **Evidence before this study**

63 We searched PubMed and preprint articles from December 2019, to July 2022, for
64 studies published in English that investigated the clinical subphenotypes of MIS-C using
65 the terms “multi-system inflammatory syndrome in children” or “pediatric inflammatory
66 multisystem syndrome” and “phenotypes”. Most previous research described the
67 symptoms, clinical characteristics and risk factors associated with MIS-C and how these
68 differ from acute COVID-19, Kawasaki Disease and Toxic Shock Syndrome. One
69 single-center study of 63 patients conducted in 2020 divided patients into Kawasaki and
70 non-Kawasaki disease subphenotypes. Another CDC study evaluated 3 subclasses of
71 MIS-C in 570 children, with one class representing the highest number of organ
72 systems, a second class with predominant respiratory system involvement, and a third
73 class with features overlapping with Kawasaki Disease. However, this study evaluated
74 cases from March to July 2020, during the early phase of the pandemic when
75 misclassification of cases as Kawasaki disease or acute COVID-19 may have occurred.
76 Therefore, it is not known from the existing literature whether the presentation of MIS-C
77 has changed with newer variants such as delta and omicron.

78 **Added value of this study**

79 PEDSnet provides one of the largest MIS-C cohorts described so far, providing
80 sufficient power for detailed analyses on MIS-C subphenotypes. Our analyses span the
81 entire length of the pandemic, including the more recent omicron wave, and provide an
82 update on the presentations of MIS-C and its temporal dynamics. We found that
83 children have a spectrum of illness that can be characterized as mild (lower

84 inflammatory markers, fewer organ systems involved), moderate (4-6 organ involvement
85 with clinical overlap with acute COVID-19) and severe (higher inflammatory markers,
86 critically ill, more likely to have cardiac involvement, with hypotension/shock and need
87 for vasopressors).

88 **Implications of all the available evidence**

89 These results provide an update to the subphenotypes of MIS-C including the more
90 recent delta and omicron periods and aid in the understanding of the various
91 presentations of MIS-C. These and other findings provide a useful framework for clinicians
92 in the recognition of MIS-C, identify factors associated with children at risk for increased
93 severity, including the importance of laboratory parameters, for risk stratification, and to
94 facilitate early evaluation, diagnosis and treatment.

95

96 **Abstract:**

97 **Background:** Multi-system inflammatory syndrome in children (MIS-C) represents one of the
98 most severe post-acute sequelae of SARS-CoV-2 infection in children, and there is a critical
99 need to characterize its disease patterns for improved recognition and management. Our
100 objective was to characterize subphenotypes of MIS-C based on presentation, demographics
101 and laboratory parameters.

102 **Methods:** We conducted a retrospective cohort study of children with MIS-C from March 1,
103 2020 - April 30, 2022 and cared for in 8 pediatric medical centers that participate in PEDSnet.
104 We included demographics, symptoms, conditions, laboratory values, medications and
105 outcomes (ICU admission, death), and grouped variables into eight categories according to
106 organ system involvement. We used a heterogeneity-adaptive latent class analysis model to
107 identify three clinically-relevant subphenotypes. We further characterized the sociodemographic
108 and clinical characteristics of each subphenotype, and evaluated their temporal patterns.

109 **Findings:** We identified 1186 children hospitalized with MIS-C. The highest proportion of
110 children (44.4%) were aged between 5-11 years, with a male predominance (61.0%), and non-
111 Hispanic white ethnicity (40.2%). Most (67.8%) children did not have a chronic condition. Class
112 1 represented children with a severe clinical phenotype, with 72.5% admitted to the ICU, higher
113 inflammatory markers, hypotension/shock/dehydration, cardiac involvement, acute kidney injury
114 and respiratory involvement. Class 2 represented a moderate presentation, with 4-6 organ
115 systems involved, and some overlapping features with acute COVID-19. Class 3 represented a
116 mild presentation, with fewer organ systems involved, lower CRP, troponin values and less
117 cardiac involvement. Class 1 initially represented 51.1% of children early in the pandemic,
118 which decreased to 33.9% from the pre-delta period to the omicron period.

119 **Interpretation:** MIS-C has a spectrum of clinical severity, with degree of laboratory
120 abnormalities rather than the number of organ systems involved providing more useful
121 indicators of severity. The proportion of severe/critical MIS-C decreased over time.

122

123 **Introduction**

124 In April 2020, cases of a novel hyperinflammatory disorder associated with SARS-CoV-2
125 infection were first reported in the UK and Italy,^{1,2} which was subsequently termed multisystem
126 inflammatory syndrome in children (MIS-C). The Centers for Disease Control and Prevention
127 (CDC) defines cases as individuals younger than 21 years of age presenting with fever,
128 laboratory evidence of inflammation, and evidence of clinically severe illness requiring
129 hospitalization, with involvement of greater than or equal to 2 organs, with no alternative
130 plausible diagnoses and a positive SARS-CoV-2 test (PCR, serology, antigen) or COVID-19
131 exposure within the preceding 4 weeks.³ Reports indicate that most children are 6-12 years of
132 age, of Black/African American race or Hispanic ethnicity, and obesity was the most common
133 pre-existing condition.⁴⁻⁶ MIS-C is one of the most severe post-acute sequelae of SARS-CoV-2
134 infection in children, and as of June 2022, the CDC has reported 8,639 cases of MIS-C; 70
135 children have died due to associated complications.⁷

136 Children with MIS-C can present with a highly heterogeneous spectrum of clinical features,
137 making the diagnosis challenging. The clinical presentation can overlap with features of
138 Kawasaki disease, acute COVID-19, toxic shock, and macrophage activation syndrome.
139 Common features include fever, gastrointestinal symptoms (including abdominal pain and
140 diarrhea), cardiac complications (myocarditis and coronary artery dilatation), mucocutaneous
141 findings (rash, mucositis) and respiratory symptoms.⁸ Given the wide spectrum of clinical
142 presentations of this novel condition, there is a need to further characterize its disease patterns,
143 which will ultimately improve outcomes. Further elucidation of MIS-C subphenotypes can help
144 identify children at greatest risk of severe outcomes, and determine which subgroups may
145 benefit from more targeted therapies. Therefore, the objectives of this study were to define
146 subphenotypes of MIS-C based on symptoms, diagnoses and laboratory parameters.

147 **Methods**

148 *Data Source*

149 This retrospective cohort study is part of the NIH Researching COVID to Enhance Recovery
150 (RECOVER) Initiative, which seeks to understand, treat, and prevent the post-acute sequelae of
151 SARS-CoV-2 infection (PASC). For more information on RECOVER, visit
152 <https://recovercovid.org/>. The original study cohort has been described elsewhere.⁹ We
153 used electronic health record (EHR) data from PEDSnet member institutions.^{10,11} Participating
154 institutions included Children’s Hospital of Philadelphia, Cincinnati Children’s Hospital Medical
155 Center, Children’s Hospital Colorado, Ann & Robert H. Lurie Children’s Hospital of Chicago,
156 Nationwide Children’s Hospital, Nemours Children’s Health System (a Delaware and Florida
157 health system), Seattle Children’s Hospital, and Stanford Children’s Health. We retrieved EHR
158 data from all healthcare encounters among hospitalized children and adolescents < 21 years of
159 age who were diagnosed with MIS-C from March 1, 2020 to April 30, 2022. Data were extracted
160 from the PEDSnet COVID-19 Database-Version 2022-05-19, which included EHR data with
161 dates of service up to April 30, 2022. Cohort entry was defined as the 1 day before the first
162 hospitalization associated with an MIS-C diagnosis term (index hospitalization). We combined
163 hospitalizations if the discharge date of the first hospitalization was within 24 hours of the
164 subsequent hospitalization admission date. We excluded one patient with missing admission or
165 discharge dates. The Children’s Hospital of Philadelphia’s Institutional Review Board designated
166 this study as not human subjects’ research and waived informed consent.

167 *Variables*

168 We included the following indicator variables: demographics, diagnoses (including symptoms
169 and conditions), medications, laboratory values and clinical outcomes (ICU, mechanical
170 ventilation, death), which were coded dichotomously and later grouped to form variables used in

171 our analysis. The indicator variable was included in the analyses if at least one occurrence was
172 recorded in the EHR from 1 day prior through the entire duration of hospitalization. In addition,
173 age, number of days during hospitalization, number of days in ICU, and laboratory test results
174 were included as continuous or categorical variables.

175 *Statistical Analyses*

176 First, we grouped highly-correlated indicator variables, including diagnoses (symptoms and
177 conditions) and laboratory values, into eight broader categories according to their organ system
178 involvement, which were cardiac, gastrointestinal, hematological, neurologic, renal, respiratory,
179 dermatologic and shock (See **eTable 1** for definitions of each variable). The variable for a
180 particular organ system was defined based on the occurrence of any symptom, condition,
181 laboratory value, or procedure type related to that system. We further divided the cardiac
182 system into three categories: coronary artery abnormalities only, myocardial dysfunction with no
183 coronary artery abnormalities, and coronary artery abnormalities and myocardial dysfunction.

184 Secondly, with the ten manifest variables defined over the eight broad organ systems, we
185 applied a latent class analysis (LCA) model in which an Expectation-Maximization (EM)
186 algorithm was used for estimation.^{12,13} To account for heterogeneity across hospitals yet
187 ensuring unified definitions of subphenotypes, we assumed the same collection of
188 subphenotypes but potentially different mixing proportions of each of the subphenotypes across
189 the health systems. Such generalization of the traditional LCA model is necessary because the
190 prevalence of subphenotypes can be substantially different due to different distributions of age,
191 race/ethnicity, and referral patterns which can be strongly associated with the presentation of
192 MIS-C.

193 For sensitivity analysis, we fitted LCA to the ten manifest variables under different assumed
194 numbers of latent classes, varying from two to six. Log-likelihood, Bayesian Information

195 Criterion (BIC), and clinical adjudication among our clinician work group were incorporated to
196 select the appropriate number of clinical-relevant latent classes.

197 For identification of identified latent classes (subphenotypes), proportions of the ten manifest
198 variables over the eight broad organ systems across latent classes were displayed as heatmaps
199 of prevalence. To further characterize these identified latent classes, distributions of clinical
200 indicator variables, including demographic characteristics, and individual symptoms, conditions,
201 laboratory results, procedures, and medications, were displayed across the latent classes also
202 using heatmaps of prevalence. To characterize the differences in each continuous clinical
203 variable across latent classes, sample mean, standard error, median and interquartile range
204 were calculated and summarized. To reduce the impact of misclassifications of latent classes,
205 instead of using classification based on maximum posterior probability, all calculations of the
206 characterizations were based on weighted prevalence with weights being the posterior
207 probabilities. The results are presented using heatmaps and forest plots.

208 To evaluate potential time-varying effects, we also evaluated the changes of the proportions for
209 each subphenotype over time in 3-month periods. Within each time period, the proportions of
210 subphenotypes were calculated using posterior class membership probabilities. The change in
211 subclass proportions over time was shown as a line graph in Figure 4. In addition, the change in
212 the prevalence of the ten variables for each of the latent classes over time was estimated.

213 To further examine the validity of our LCA approach, we applied the subphenotypes of MIS-C to
214 a cohort of children testing positive for COVID-19 by PCR without an MIS-C diagnosis.⁹ We
215 constructed latent classes for SARS-CoV-2 positive patients, with an evaluation period of -7 to
216 28 days from the test date to evaluate our construct validity. We applied the model with three
217 latent classes to the ten variables as used in the MIS-C analysis and identified three
218 subphenotypes of COVID-19. We compared the characteristics of MIS-C and COVID-19
219 subphenotypes and visualized the distance among them measured by fixation index using

220 multidimensional scaling. Analyses were conducted using R version 4.1.2 (2021-11-01). The
221 package used for LCA was poLCA 1.4.1 (2014-01-10).^{14,15}

222 **Results**

223 Among 1,853,557 children in the PEDSnet cohort who received a COVID-19 diagnosis or
224 vaccination, had a COVID-19 test (PCR, serology or antigen), or demonstrated evidence of
225 respiratory illness or post-acute sequelae of SARS-CoV-2 and had a visit at any of the eight
226 PEDSnet sites between March 2020 and April 2022, we identified 1186 children who were
227 hospitalized with MIS-C (**eFigure 1**). A description of the study cohort is provided in **Table 1**.

228 The highest proportion of children (44%) were aged between 5-11 years, with a male
229 predominance (61%), and non-Hispanic White ethnicity (40.2%). Most (67.8%) children did not
230 have any chronic conditions as defined by the Pediatric Medical Complexity Algorithm (PMCA).
231 ¹⁶ In our MIS-C cohort, 40.5% were admitted to the ICU, 15.2% required mechanical ventilation
232 (invasive and non-invasive), and 11 (0.92%) children died.

233 Latent Class analyses identified three subphenotypes (**Figure 1**), with the patient demographics
234 of each class summarized in **Figure 2 and eFigure 2**. The overall proportions of Class 1, 2, and
235 3 were 40.0%, 24.0%, and 36.0%, respectively. The race and ethnicity of children were similar
236 in each class, with the exception of a higher proportion of children of black/African American
237 race in Class 1. A higher proportion of children 12 years of age and older, and those with a
238 complex chronic condition were identified in Class 1. The proportion of each class by PEDSnet
239 site is shown in **eTable 2**, with some observed heterogeneity. **Figure 3 and eFigure 3** and
240 display specific diagnoses, laboratory results and medications in each group. Class 1
241 represented children with a more severe clinical phenotype, with 72.5% admitted to the ICU,
242 with a longer mean inpatient stay and longer ICU length of stay. Children in Class 1 had a
243 higher prevalence of cardiac involvement, with abnormal troponin levels, hypotension and
244 shock/dehydration, need for vasopressors and fluid resuscitation. This class was more likely to

245 have abnormal laboratory findings including lymphopenia, elevated d-dimer and
246 thrombocytopenia. There was a higher prevalence of acute kidney injury, and respiratory
247 involvement (respiratory failure, pleural effusion, need for ventilation). Children in Class 1 had
248 the highest usage of systemic steroids among all classes. Class 2 represented a group with a
249 moderate presentation, with similar multiple organ system involvement to Class 3, including
250 dermatologic (rashes), gastrointestinal (abdominal pain, nausea and vomiting), hematologic
251 abnormalities (lymphopenia, thrombocytopenia, elevated d-dimer), renal (fluid and electrolyte)
252 and respiratory (cardiorespiratory signs and symptoms), with a heat map pattern similar to acute
253 COVID-19 (**eFigure 4**) and 2-dimensional plot demonstrating close proximity between MIS-C
254 Class 2 and COVID-19 Class 1 (**eFigure 5**). This group had the highest use of non-steroidal
255 anti-inflammatory and antirheumatic medications. Class 3 represented a mild population, with
256 fewer systems involved, and laboratory values demonstrating fluid and electrolyte disturbance,
257 lymphopenia, thrombocytopenia, and elevated d-dimer. The proportions of each subphenotype
258 over time are shown in **Figure 4**. A higher proportion of children had the more severe MIS-C
259 presentation in the earlier phase of the pandemic (Class 1 represented 51.1% of children from
260 March 2020 to August 2020), which decreased over time (Class 1 represented 33.9% of
261 children from January to April, 2022). **eFigure 6** shows the prevalence of organ system
262 involvement for all patients with MIS-C over time, demonstrating significant declines over time in
263 cardiac, hematological, respiratory involvement, and shock.

264 **Discussion**

265 In our multicenter cohort of 1,186 children hospitalized with MIS-C, we identified 3 latent classes
266 with different illness characteristics. Class 1 represented critically ill patients, with a higher
267 proportion of children being of older age, black/African American race, with underlying medical
268 complexity, presenting with hypotension/shock, dehydration, cardiac dysfunction and laboratory
269 abnormalities including lymphopenia, thrombocytopenia, elevated d-dimer and troponin. Class 3

270 represented a younger population with the milder form of MIS-C, and fewer organ systems
271 involved (hematological and renal). Class 2 represented an intermediate population with lower
272 illness severity than Class 1 but greater than Class 3, with 4-5 organ system involvement,
273 including gastrointestinal (nausea, vomiting, abdominal pain, diarrhea), respiratory,
274 hematological and renal, and a lower incidence of cardiac dysfunction and abnormal troponin
275 levels, with a similar pattern to a subset of children with acute COVID-19. We also found a
276 higher proportion of patients with the more severe Class 1 phenotype seen earlier in the
277 pandemic, which decreased over time, with lower proportions of children with respiratory,
278 cardiac and hematological involvement and shock. These findings highlight the varied
279 presentations of children with MIS-C, with wide illness spectrum, variable system involvement,
280 and cardiac involvement mostly confined to myocardial dysfunction rather than coronary artery
281 involvement. Our findings suggest that the degree of laboratory abnormalities rather than the
282 number of organ systems involved may be more useful indicators of illness severity, and may
283 reflect differences in the degree of immune responses elicited.

284 Our multi-site study included children from diverse geographical regions across the United
285 States, and captured diagnoses, laboratory values, medications and procedures. Another
286 strength of our study is the use of LCA, which can be especially useful for identifying subgroups
287 of individuals within a syndrome or condition who could benefit from a common intervention
288 based on their shared characteristics. LCA is also well-suited to describing different
289 manifestations of the novel syndrome of MIS-C. It divides patients into groups that might have
290 been previously unrecognized, based on shared characteristics, allowing for an unbiased
291 determination of phenotypic manifestations.^{13, 17,18} Importantly, our LCA properly accounts for
292 between-site heterogeneity in patients by allowing different mixing proportions of
293 subphenotypes, avoiding bias in estimated patient-level subphenotypes. Further, after
294 identification of subphenotypes using our LCA, we also characterized the differences in

295 variables in demographic, clinical conditions, and medications and labs, that were not used as
296 manifest variables in our LCA. These results showed important differences in these variables
297 across identified subphenotypes, which clinically validates our LCA results.

298 Our study is subject to several limitations. Given that inclusion of patients into our cohort was
299 based on a physician's diagnosis, we may have inadvertently included cases of suspected MIS-
300 C who ultimately had other diagnoses, or children whose illnesses overlapped with Kawasaki
301 disease. This risk for misclassification is low given that these diagnoses are assigned at
302 discharge. For example, we examined Kawasaki disease diagnoses in our analyses and the
303 prevalence was low across the three latent classes, suggesting that these patients were
304 accurately classified. Next, our data captures hospitalizations at a PEDSnet site, and data from
305 earlier hospitalizations at other sites was not available. More subtle symptoms from MIS-C may
306 not have been captured in EHR data sources. Finally, the number of latent classes in LCA may
307 be difficult to determine, which is a well-known challenge in statistical inference, and may lead
308 to deceptive interpretations that the set of latent classes identified in an analysis represents the
309 actual types of individuals in the population. To account for this, we have applied multiple model
310 selection criteria in selecting the number of subphenotypes. Further, we have conducted
311 sensitivity analyses on LCA using different numbers of subphenotypes. The final selected LCA
312 model is a parsimonious model that describes distinct subphenotypes, which fits with clinical
313 observations and existing literature.¹⁹

314 Our unsupervised learning methods using LCA identified a severe MIS-C subgroup with a
315 greater extent of left ventricular dysfunction leading to shock, and need for
316 vasopressors/inotropes. Abnormal cardiac strain has been shown to be associated with
317 increased odds of vasoactive support requirement, duration of vasoactive support, and ICU
318 length of stay.²⁰ An association has also been shown between the degree of inflammation and
319 cardiac injury severity, in particular CRP, troponin and NT-proBNP,²¹ as supported by our data.

320 Our data identified that those in the critically ill group were older and more likely to have an
321 underlying medical condition. In the literature, children affected by MIS-C tend to have fewer
322 underlying medical conditions, with obesity a common diagnosis if chronic conditions were
323 present.²² Older age has been associated with worse outcomes from MIS-C,²³ but the
324 underlying immunologic mechanisms behind this observation have not yet been elucidated.

325 We identified a latent class subphenotype with 4-5 organ system involvement (Class 2), with
326 overlapping features with children with acute COVID-19. A similar subphenotype has been
327 described in earlier studies of MIS-C.¹⁹ These and other studies indicate the clinical overlap in
328 some of the presentations of MIS-C and acute COVID-19. Further, the presence of abnormal
329 strain on echocardiogram has also been associated with greater number of organ systems
330 involved,²⁴ but our findings suggest that the number of organ systems involved does not
331 necessarily correlate with the degree of cardiac dysfunction and illness severity. Our third latent
332 class had a milder presentation, with shorter length of stay and a lower incidence of cardiac
333 involvement. This milder subgroup is an important cohort that warrants further exploration, as
334 they may not require as aggressive treatment with immunomodulators to achieve the same
335 positive outcomes.

336 We evaluated the proportions of each subphenotype over time, and observed a trend towards
337 decreasing frequency of the severe MIS-C subphenotype. Earlier reports of MIS-C have
338 indicated high levels of ICU support for patients with MIS-C. In a 2020 systematic review,
339 Ahmed et al. reported on 39 observational studies of MIS-C patients encompassing the period
340 January 1st, 2020 to July 25th, 2020. Out of 662 patients, 71% required admission to the
341 intensive care unit (ICU).²⁵ Further exploration of our cohort identified decreasing respiratory
342 system involvement, cardiac involvement and shock, indicating decreased need for critical care
343 interventions. The reasons for this decreasing illness severity over time are likely multifactorial.

344 As clinicians' experience with this new condition matures, earlier identification and treatment

345 may lead to improved outcomes and decreased need for critical care. SARS-CoV-2 immunity in
346 children has been increasing through natural infection and vaccination, rendering less children
347 susceptible to first-time encounters with SARS-CoV-2 infection triggering MIS-C. Another
348 hypothesis is that newer SARS-CoV-2 strains are less likely to stimulate the immune response
349 which causes multi-system inflammation.

350 In conclusion, we identified a set of subgroups of MIS-C with varying illness severity,
351 sociodemographic characteristics and organ system involvement using latent class analyses.
352 These findings highlight the need for clinicians to recognize the varied clinical spectrum of
353 presentations of MIS-C, and the importance of laboratory parameters to facilitate early
354 evaluation, diagnosis and treatment.

355 **Declaration of Interests:** Dr. Mejias reports funding from Janssen, Merck for research
356 support, and Janssen, Merck and Sanofi-Pasteur for Advisory Board participation; Dr. Rao
357 reports prior grant support from GSK and Biofire. Dr. Chen receives consulting support from
358 GSK. Dr. Jhaveri is a consultant for AstraZeneca, Seqirus and Dynavax, and receives an
359 editorial stipend from Elsevier. All other authors have no conflicts of interest to disclose.

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

363 **Author contribution statement:**

364 Drs Chen, Jing and Liu had full access to all of the data in the study and take responsibility for
365 the integrity of the data and the accuracy of the data analysis.

366 *Study concept and design:* Rao, Chen, Jing, Liu, Forrest, Maltenfort, Lorman, Schuchard

367 *Acquisition, analysis and interpretation of data:* Rao, Jing, Liu, Lorman, Maltenfort, Schuchard,

368 Wu, Tong, Razzaghi, Forrest, Bailey

369 *Drafting of the manuscript:* Rao, Jing, Chen

370 *Critical revision of the manuscript for important intellectual content:* all authors

371 *Statistical analyses:* Jing, Liu, Chen, Wu, Tong

372 *Obtained funding:* Lee, Bailey, Forrest

373 *Administrative, technical or material support:* Rao, Chen, Bailey, Lee, Chen, Forrest

374 *Study supervision:* Rao, Chen, Forrest, Bailey, Forrest

375

376 **Data sharing statement:** The data is not publicly available due to privacy concerns. The

377 individual de-identified participant data will not be shared. The data that support the findings of

378 this study may be available through request and DUA process to the corresponding authors.

379

380 **Table 1. Description of Study Cohort**

Characteristic		Number of patients (%) (N = 1186)
Age at MIS-C diagnosis, years	<1	30 (2.5%)
	1-4	259 (21.8%)
	5-11	526 (44.4%)
	12-15	244 (20.6%)
	16-20	127 (10.7%)
Sex	Female	463 (39.0%)
	Male	723 (61.0%)
Race/Ethnicity	Hispanic	253 (21.3%)
	Non-Hispanic White	477 (40.2%)
	Non-Hispanic Black/African-American	291 (24.5%)
	Non-Hispanic Asian/Pacific Islander	44 (3.7%)
	Other/Unknown	69 (5.8%)
	Multiple	52 (4.4%)
Institution	A	193 (16.3%)
	B	246 (20.8%)
	C	67 (5.6%)
	D	139 (11.7%)
	E	189 (15.9%)
	F	42 (3.5%)
	G	197 (16.6%)
	H	113 (9.5%)
PMCA index^a	Non-complex chronic	806 (67.8%)
	Chronic	64 (5.4%)
	Complex-chronic	318 (26.8%)
Obesity		134 (11.3%)
Admitted to ICU		480 (40.5%)
Died		11 (0.9%)
Received mechanical ventilation^b		180 (15.2%)

381

382 **Footnote:** ^aPediatric Medical Complexity Algorithm (PMCA) classified children as complex-
 383 chronic, chronic, or non-complex-chronic using an ordinal scale

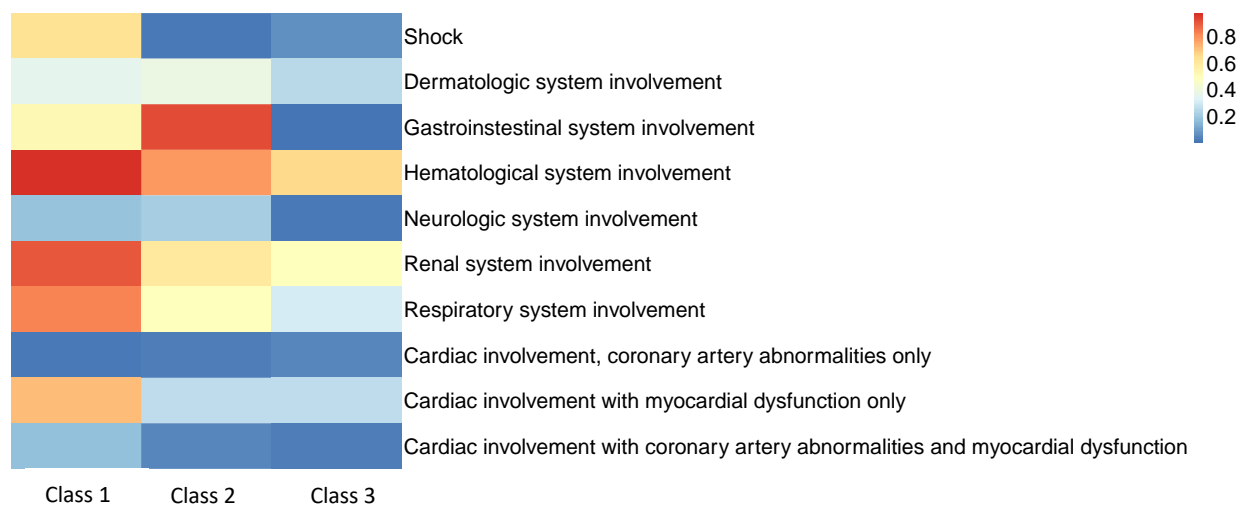
384 ^binvasive or non-invasive

385 **Abbreviations:** MIS-C- multisystem inflammatory syndrome in children; ICU- intensive care unit

386

387 **Figure 1. Heatmap of eight organ system variables used in latent class analyses with**
388 **three classes.**

389 **Figure legend:** The proportions of Class 1, 2, and 3 are 40.0%, 24.0%, and 36.0%,
390 respectively. Each column represents a latent class, and each row represents a manifest
391 variable. The color of the boxes represents the prevalence of variables. The legend on the top
392 right shows the scale of the colors. Red represents prevalence close to 100% and blue
393 represents prevalence close to 0%.

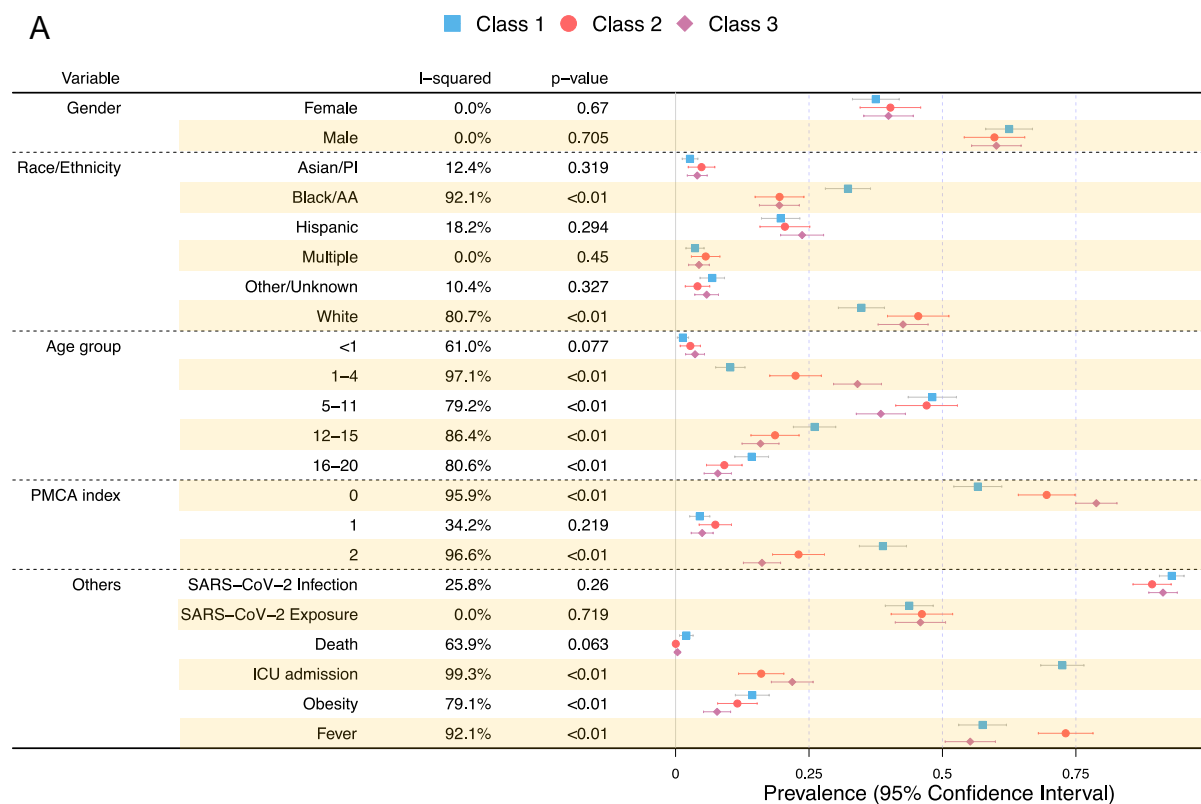


394

395 **Figure 2. Patient demographic characteristics in three latent classes.**

396 **Figure legend:** Each line plot indicates the estimated prevalence and associated 95%
397 confidence interval of each variable in each class. Class 1 is represented as a blue square,
398 Class 2 as a red circle, and Class 3 as a purple diamond. The I-squared statistic measures the
399 heterogeneity among the latent classes. P-values are obtained through Cochran's Q test, where
400 the null hypothesis is no between-class heterogeneity. A large I-squared and a small p-squared
401 indicate a large between-class heterogeneity. **A)** I-squared, p-value, estimated prevalence, and

402 its 95% confidence interval for discrete variables. **B)** I-squared, p-value, estimated mean,
 403 standard deviation, median and interquartile range for continuous variables.



404 **B**

	Mean (SD)			Median (IQR)			I-squared	p-value
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3		
Age in years	10.2 (4.5)	8.5 (4.7)	7.3 (4.7)	10.1 (6.9, 13.6)	8.2 (4.5, 12.0)	6.3 (3.0, 11.2)	97.6%	<0.01
Inpatient days	8.6 (13.7)	5.7 (8.7)	4.7 (6.9)	6 (5, 9)	5 (4, 6)	4 (3, 5.2)	93.3%	<0.01
Days in ICU	3.6 (5.8)	0.7 (3.0)	0.8 (2.1)	3 (0, 4)	0 (0, 0)	0 (0, 0)	98.0%	<0.01
Highest CRP (mg/dL)	20.3 (10.9)	14.1 (10.0)	12.3 (9.4)	20.0 (11.1, 26.6)	13.4 (6.0, 21.6)	9 (5.1, 19.2)	78.4%	<0.01
Highest troponin (ng/ml)	1.3 (8.4)	0.1 (0.7)	0.1 (0.8)	0.01 (0, 0.05)	0.14 (0.02, 0.36)	0.01 (0, 0.05)	98.6%	<0.01

405

406 **Footnote:** SARS-CoV-2 infection includes positive PCR, antigen or serology; SARS-CoV-2
407 exposure indicates that patient had a history of exposure to a contact with known SARS-CoV-2
408 based on ICD-10 codes; age group is presented in years; Pediatric Medical Complexity
409 Algorithm (PMCA) classified children as complex-chronic (2), chronic (1), or non-complex-
410 chronic (0).

411 **Abbreviations:** SD--standard deviation; IQR—interquartile range; PMCA-- Pediatric Medical
412 Complexity Algorithm; PI- Pacific Islander; AA- African American; SARS-CoV-2--severe acute
413 respiratory syndrome coronavirus; ICU- intensive care unit; CRP- C-reactive protein

414

415 **Figure 3. Heatmap of patient characteristics of conditions, laboratory results and**
416 **medication in three latent classes.**

417 **Figure legend:** Each column represents a latent class, and each row represents a variable. The
418 color of the boxes represents the estimated prevalence of variables. The legend on the top right
419 shows the scale of the colors. Red represents prevalence close to 100% and blue represents
420 prevalence close to 0%. The conditions and laboratory results are ordered based on organ
421 systems.

422

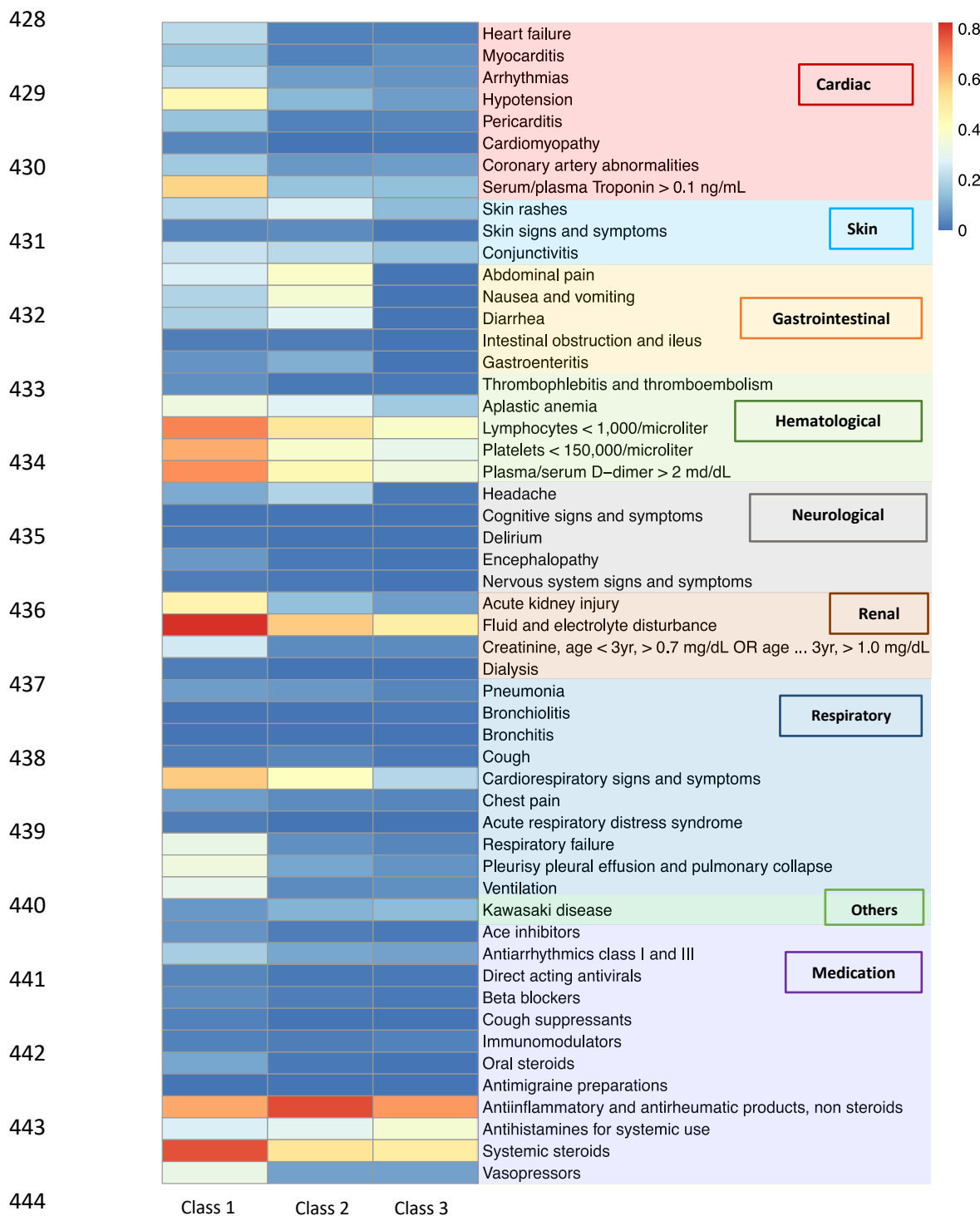
423

424

425

426

427



445
446
447

448 **Figure 4. Proportions of three MIS-C classes over time**

449 **Figure legend:** Each line represents the proportion of the three classes over 3-month periods.

450 The proportions of the classes are calculated using the average of the posterior class

451 membership probabilities of the patients who are diagnosed in the given time window. The

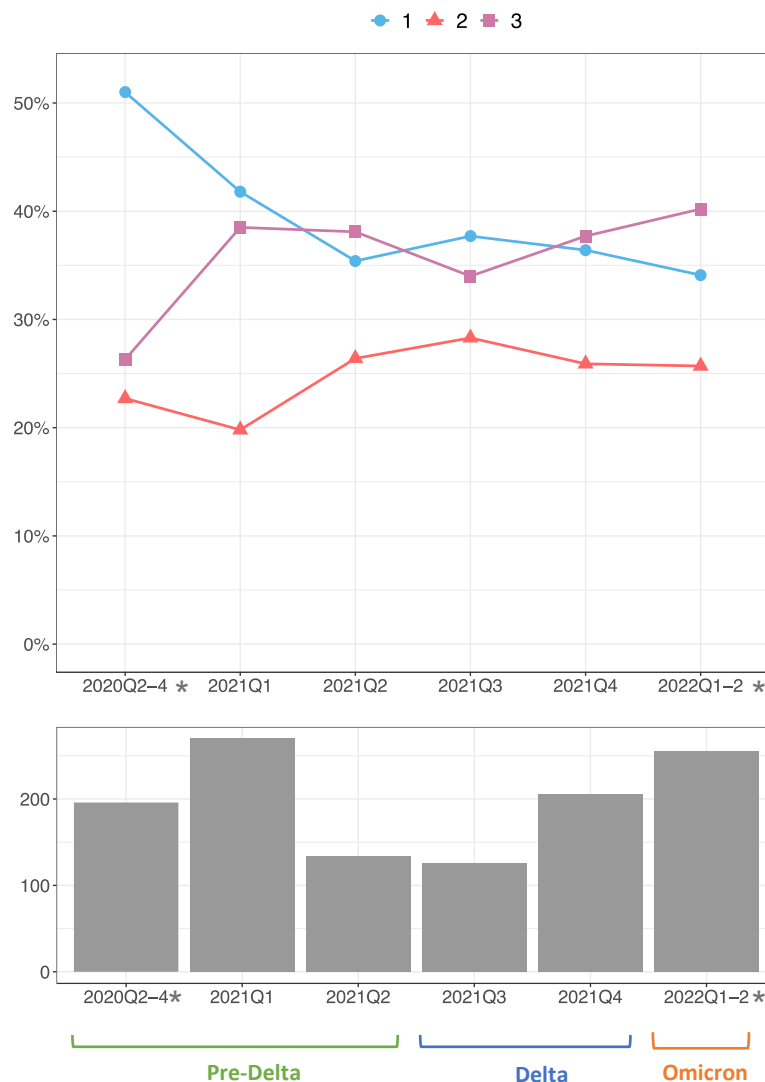
452 calendar time is segmented quarterly. Time intervals with small sample sizes are combined and

453 labeled by a star. **A)** Proportions of the three subphenotypes across time. Each line represents

454 a subphenotype. **B)** Histogram of the number of patients in each time window, including timing

455 of predominant SARS-CoV-2 variant activity.

456



457 References

- 458 1. Verdoni L, Mazza A, Gervasoni A, et al. An outbreak of severe Kawasaki-like disease at the Italian
459 epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet*. Jun 6
460 2020;395(10239):1771-1778. doi:10.1016/S0140-6736(20)31103-X
- 461 2. Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory
462 shock in children during COVID-19 pandemic. *Lancet*. May 23 2020;395(10237):1607-1608.
463 doi:10.1016/S0140-6736(20)31094-1
- 464 3. Centers for Disease Control and Prevention. Multisystem Inflammatory Syndrome in Children
465 (MIS-C) Associated with Coronavirus Disease 2019 (COVID-19). Accessed May 20, 2022,
466 <https://emergency.cdc.gov/han/2020/han00432.asp>
- 467 4. Kaushik A, Gupta S, Sood M, Sharma S, Verma S. A Systematic Review of Multisystem
468 Inflammatory Syndrome in Children Associated With SARS-CoV-2 Infection. *Pediatr Infect Dis J*. Nov
469 2020;39(11):e340-e346. doi:10.1097/INF.0000000000002888
- 470 5. Dufort EM, Koumans EH, Chow EJ, et al. Multisystem Inflammatory Syndrome in Children in New
471 York State. *N Engl J Med*. Jul 23 2020;383(4):347-358. doi:10.1056/NEJMoa2021756
- 472 6. Patel JM. Multisystem Inflammatory Syndrome in Children (MIS-C). *Curr Allergy Asthma Rep*.
473 May 2022;22(5):53-60. doi:10.1007/s11882-022-01031-4
- 474 7. Center for Disease Control and Prevention. COVID Data Tracker. Health Department-Reported
475 Cases of Multisystem Inflammatory Syndrome in Children (MIS-C) in the United States. Accessed June
476 20, 2022, <https://www.cdc.gov/mis-c/cases>
- 477 8. Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem Inflammatory Syndrome in U.S. Children
478 and Adolescents. *N Engl J Med*. Jul 23 2020;383(4):334-346. doi:10.1056/NEJMoa2021680
- 479 9. Rao S LG, Razzaghi H, Lorman V, Mejias A, Pajor NM, Thacker D, Webb R, Dickinson K, Bailey LC,
480 Jhaveri R, Christakis DA, Bennett TD, Chen Y, Forrest CB. Clinical features and burden of post-acute
481 sequelae of SARS-CoV-2 infection in children and adolescents: an exploratory EHR-based cohort study
482 from the RECOVER program. *JAMA Pediatrics*. 2022;in press
- 483 10. Forrest CB, Margolis PA, Bailey LC, et al. PEDSnet: a National Pediatric Learning Health System. *J*
484 *Am Med Inform Assoc*. Jul-Aug 2014;21(4):602-6. doi:10.1136/amiainl-2014-002743
- 485 11. Forrest CB, Margolis P, Seid M, Colletti RB. PEDSnet: how a prototype pediatric learning health
486 system is being expanded into a national network. *Health Aff (Millwood)*. Jul 2014;33(7):1171-7.
487 doi:10.1377/hlthaff.2014.0127
- 488 12. Melton B, Liang KY, Pulver AE. Extended latent class approach to the study of familial/sporadic
489 forms of a disease: its application to the study of the heterogeneity of schizophrenia. *Genet Epidemiol*.
490 1994;11(4):311-27. doi:10.1002/gepi.1370110402
- 491 13. Bandeem-roche K MD, Zeger SL, Rathouz PJ. Latent Variable Regression for Multiple
492 Discrete Outcomes. *Journal of the American Statistical Association*. 1997;92(440):1375-1386.
- 493 14. Linzer DA LJ. polCA: an R Package for Polytomous Variable Latent Class Analysis. *Journal of*
494 *Statistical Software*. 2011;42(10):1-29.
- 495 15. polCA: Polytomous Variable Latent Class Analysis." R package version 1.4. .
496 <http://dlinzer.github.com/polCA>.
- 497 16. Simon TD, Cawthon ML, Popalisky J, Mangione-Smith R, Center of Excellence on Quality of Care
498 Measures for Children with Complex N. Development and Validation of the Pediatric Medical Complexity
499 Algorithm (PMCA) Version 2.0. *Hosp Pediatr*. Jul 2017;7(7):373-377. doi:10.1542/hpeds.2016-0173
- 500 17. !!! INVALID CITATION !!! 15-18;
- 501 18. Linzer DA LJ. polCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of*

- 502 *Statistical Software*. 2011;41(10):1-29.
- 503 19. Godfred-Cato S, Bryant B, Leung J, et al. COVID-19-Associated Multisystem Inflammatory
504 Syndrome in Children - United States, March-July 2020. *MMWR Morb Mortal Wkly Rep*. Aug 14
505 2020;69(32):1074-1080. doi:10.15585/mmwr.mm6932e2
- 506 20. Basu S, Kim EJ, Sharron MP, et al. Strain Echocardiography and Myocardial Dysfunction in
507 Critically Ill Children With Multisystem Inflammatory Syndrome Unrecognized by Conventional
508 Echocardiography: A Retrospective Cohort Analysis. *Pediatr Crit Care Med*. Mar 1 2022;23(3):e145-e152.
509 doi:10.1097/PCC.0000000000002850
- 510 21. Farooqi KM, Chan A, Weller RJ, et al. Longitudinal Outcomes for Multisystem Inflammatory
511 Syndrome in Children. *Pediatrics*. Aug 2021;148(2)doi:10.1542/peds.2021-051155
- 512 22. Chow EJ, Englund JA. Severe Acute Respiratory Syndrome Coronavirus 2 Infections in Children.
513 *Infect Dis Clin North Am*. Jun 2022;36(2):435-479. doi:10.1016/j.idc.2022.01.005
- 514 23. Merckx J, Cooke S, El Tal T, et al. Predictors of severe illness in children with multisystem
515 inflammatory syndrome after SARS-CoV-2 infection: a multicentre cohort study. *CMAJ*. Apr 11
516 2022;194(14):E513-E523. doi:10.1503/cmaj.210873
- 517 24. Fremed MA, Farooqi KM. Longitudinal Outcomes and Monitoring of Patients With Multisystem
518 Inflammatory Syndrome in Children. *Front Pediatr*. 2022;10:820229. doi:10.3389/fped.2022.820229
- 519 25. Ahmed M, Advani S, Moreira A, et al. Multisystem inflammatory syndrome in children: A
520 systematic review. *EClinicalMedicine*. Sep 2020;26:100527. doi:10.1016/j.eclinm.2020.100527

521