Risk Variants in the Exomes of Children With Critical Illness

Joshua E. Motelow, MD, PhD; Natalie C. Lippa, MS, CGC; Joseph Hostyk, MA; Evin Feldman, MD; Matthew Nelligan, MD; Zhong Ren, MS; Anna Alkelai, PhD; Joshua D. Milner, MD; Ali G. Gharavi, MD; Yingying Tang, MD; David B. Goldstein, PhD; Steven G. Kernie, MD

Abstract

IMPORTANCE Diagnostic genetic testing can lead to changes in management in the pediatric intensive care unit. Genetic risk in children with critical illness but nondiagnostic exome sequencing (ES) has not been explored.

OBJECTIVE To assess the association between loss-of-function (LOF) variants and pediatric critical illness.

DESIGN, SETTING, AND PARTICIPANTS This genetic association study examined ES first screened for causative variants among 267 children at the Morgan Stanley Children's Hospital of New York-Presbyterian, of whom 22 were otherwise healthy with viral respiratory failure; 18 deceased children with bronchiolitis from the Office of the Chief Medical Examiner of New York City, of whom 14 were previously healthy; and 9990 controls from the Institute for Genomic Medicine at Columbia University Irving Medical Center. The ES data were generated between January 1, 2015, and December 31, 2020, and analyzed between January 1, 2017, and September 2, 2022.

EXPOSURE Critical illness.

MAIN OUTCOMES AND MEASURES Odds ratios and \(P\) values for genes and gene-sets enriched for rare LOF variants and the loss-of-function observed/expected upper bound fraction (LOEUF) score at which cases have a significant enrichment.

RESULTS This study included 285 children with critical illness (median [range] age, 4.1 [0-18.9] years; 148 [52%] male) and 9990 controls. A total of 228 children (80%) did not receive a genetic diagnosis. After quality control (QC), 231 children harbored excess rare LOF variants in genes with a LOEUF score of 0.680 or less (intolerant genes) (\(P = 1 \times 10^{-5}\)). After QC, 176 children without a diagnosis harbored excess ultrarare LOF variants in intolerant genes but only in those without a known disease association (odds ratio, 1.8; 95% CI, 1.3-2.5). After QC, 25 children with viral respiratory failure harbored excess ultrarare LOF variants in intolerant genes but only in those without a known disease association (odds ratio, 2.8; 95% CI, 1.1-6.6). A total of 114 undiagnosed children were enriched for de novo LOF variants in genes without a known disease association (observed, 14; expected, 6.8; enrichment, 2.05).

CONCLUSIONS AND RELEVANCE In this genetic association study, excess LOF variants were observed among critically ill children despite nondiagnostic ES. Variants lay in genes without a known disease association, suggesting future investigation may connect phenotypes to causative genes.


Key Points

Question Do the exomes of critically ill children without a genetic diagnosis carry risk factors?

Findings This genetic association study of 267 children with critical illness and 18 deceased children found a statistically significant burden of predicted loss-of-function variants in cases compared with 9990 controls among genes intolerant to loss-of-function variation.

Meaning Children with critical illness may carry loss-of-function risk variants for their clinical presentation regardless of whether a causative diagnostic variant is identified.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.
Introduction

Children with monogenic or chromosomal abnormalities account for 16% of pediatric admissions and have increased mortality risk in pediatric intensive care units (PICUs).1-4 Whole exome sequencing (ES) and genome sequencing (GS) achieve diagnoses in 24% to 52% of children with critical illness and can lead to changes in management.5-13 Among patients with nondiagnostic ES (ndES), an explanatory variant may be unreported because it was (1) undetected by ES, (2) detected by ES but not determined to be causative, or (3) detected by ES in a gene of unknown significance, among other reasons.14,15

Bronchiolitis is a viral infection of the lower respiratory tract.16 Nearly every child will be exposed to a virus that causes bronchiolitis.17 Of children admitted to the hospital for bronchiolitis, approximately 9% require ventilatory support, and death occurs in less than 0.1%.18,19 Although risk factors for hospitalization have been identified, including prematurity, most hospitalized children are otherwise healthy.16,19-21 Genetic risk factors have been proposed, but a definitive answer is elusive.22-26 We hypothesized that children with critical illness and ndES carry deleterious variants in genes without a known disease association. These variants are risk factors for the child’s illness but do not yield a diagnosis because the genes are not currently associated with disease. If such an association existed, larger association studies could uncover causative genes for pediatric critical illness.

Methods

Study Population

Children with critical illness were enrolled through 2 arms: (1) critically ill children at New York-Presbyterian Morgan Stanley Children’s Hospital/Columbia University Irving Medical Center (MSCH/CUIMC; MSCH cohort, n = 267) and (2) deceased children from the Office of the Chief Medical Examiner (OCME) of New York City (OCME cohort, n = 18). Critical illness in the MSCH cohort was defined by admission to the PICU or respiratory failure from a virus requiring noninvasive positive pressure. The OCME data were collected from January 1, 2003, to December 31, 2016, using a keyword search for viral bronchiolitis in the immediate cause of death. Of 24 cases identified, 18 had adequate samples for ES (Table 1; eTable 1 in the Supplement). We defined a combined cohort by merging the MSCH and OCME cohorts (n = 285) (Table 1). A viral respiratory failure cohort was created by combining healthy probands from the MSCH cohort with respiratory failure secondary to a virus (n = 22) (eTable 2 in the Supplement) and (2) probands from the OCME cohort with no phenotype data suggestive of chronic illness or an alternative cause before death (n = 14) (eTable 1 in the Supplement). Children admitted for trauma or who were 19 years or older were excluded.

Controls were individuals with data at the Institute for Genomic Medicine (IGM) enrolled as controls (n = 3358) or healthy family members (n = 6632) across multiple studies. The Columbia University Institutional Review Board approved protocols, and written informed consent was provided for using DNA in genetic research at the IGM at CUIMC.27 Individuals enrolled at the MSCH were referred by their CUIMC clinician. Written informed consent was provided by the legal guardian for participants younger than 18 years or by the individual when 18 years or older. This report follows the guidelines for case-control genetic association studies as outlined by the Strengthening the Reporting of Genetic Association Studies (STREGA) reporting guideline.

Next-Generation Sequencing, Variant Calling, and Quality Control

Sequencing was performed at or transferred to the IGM. Case data are from ES. Controls underwent ES (n = 9353) or GS (n = 637). Established protocols have been used at the IGM to generate ES or GS for more than 120 000 individuals (eMethods in the Supplement).28-30 Variants were annotated with an in-house analysis tool for annotated variants (ATAV) platform.31
Diagnostic Analysis

The diagnostic analysis of the MSCH cohort is a subset of that reported by Lippa et al. Briefly, trio (ie, ES available from proband and both biological parents) or proband-only ES analyses were completed. Candidate genotypes were reviewed with the research and clinical teams to reach consensus. Variants deemed causative were confirmed in a Clinical Laboratory Improvement Amendments–certified laboratory before return to patients and clinicians. Only single-nucleotide variants and insertion/deletions were analyzed. No other variants were assessed in the research pipeline. Individuals may have also undergone genetic evaluation by their clinical team.

Probands harboring genetic findings that explained all or part of their phenotypes were deemed resolved (n = 57) (Table 1). Genetic findings were drawn from the research pipeline and medical record review. Six of the 57 probands were resolved outside this research pipeline. The remainder were considered the unresolved cohort (n = 228). The ES from these probands were considered nondiagnostic ES (ndES). Participants recruited from the OCME underwent a modified diagnostic pipeline because of limited phenotype information (eTable 1 in the Supplement). No explanatory variants were identified, and all were considered unresolved.

Statistical Analysis

Clustering

We matched cases and controls by geographic ancestry using clustering as previously described because it is important to correct for the underlying rate of variation in samples of different geographic ancestry in case-control experimental designs (eFigure 1 in the Supplement). Principal component analysis for dimensionality reduction was performed on a set of predefined variants to capture population structure. Clusters were identified using the Louvain method of

Table 1. Demographic Characteristics of 285 Children With Critical Illness

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%)</th>
<th>Age, median (IQR) [range], y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cohort</td>
<td>285 (100)</td>
<td>4.1 (0.9-11.6) [0.0-18.9]</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>137 (48)</td>
<td>4.8 (1.0-12.7) [0.0-18.9]</td>
</tr>
<tr>
<td>Male</td>
<td>148 (52)</td>
<td>3.4 (0.9-10.6) [0.1-18.4]</td>
</tr>
<tr>
<td>Enrollment location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinic</td>
<td>131 (46)</td>
<td>5.2 (2.1-12.4) [0.0-18.6]</td>
</tr>
<tr>
<td>PICU*</td>
<td>136 (48)</td>
<td>3.6 (0.7-11.9) [0.0-18.9]</td>
</tr>
<tr>
<td>OCME</td>
<td>18 (6)</td>
<td>0.3 (0.1-0.9) [0.0-1.8]</td>
</tr>
<tr>
<td>Diagnostic status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolved</td>
<td>46 (16)</td>
<td>4.3 (0.7-11.5) [0.1-18.0]</td>
</tr>
<tr>
<td>Partially resolved</td>
<td>11 (4)</td>
<td>5.1 (2.2-9.0) [0.2-15.3]</td>
</tr>
<tr>
<td>Unresolved</td>
<td>228 (80)</td>
<td>3.7 (1.0-11.7) [0.0-18.9]</td>
</tr>
<tr>
<td>Viral respiratory failure cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCH</td>
<td>22 (61)</td>
<td>1.0 (0.2-1.7) [0.0-9.6]</td>
</tr>
<tr>
<td>OCME</td>
<td>14 (39)</td>
<td>0.2 (0.1-0.4) [0.0-1.8]</td>
</tr>
<tr>
<td>Geographic ancestry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>87 (31)</td>
<td>NR</td>
</tr>
<tr>
<td>East Asian</td>
<td>12 (4)</td>
<td>NR</td>
</tr>
<tr>
<td>European</td>
<td>42 (15)</td>
<td>NR</td>
</tr>
<tr>
<td>Latino</td>
<td>92 (32)</td>
<td>NR</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>1 (0)</td>
<td>NR</td>
</tr>
<tr>
<td>South Asian</td>
<td>14 (5)</td>
<td>NR</td>
</tr>
<tr>
<td>Admixed</td>
<td>37 (13)</td>
<td>NR</td>
</tr>
<tr>
<td>Trios</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolved</td>
<td>46</td>
<td>4.1 (0.7-10.0) [0.1-17.6]</td>
</tr>
<tr>
<td>Unresolved</td>
<td>114</td>
<td>3.6 (1.0-9.9) [0.0-18.9]</td>
</tr>
</tbody>
</table>

Abbreviations: MSCH, NewYork-Presbyterian Morgan Stanley Children's Hospital; NR, not reported; OCME, Office of the Chief Medical Examiner; PICU, pediatric intensive care unit.

* Includes 1 proband with a PICU admission at an outside hospital and 2 children enrolled from the general pediatric ward but still meeting criteria for respiratory failure based on respiratory failure that required noninvasive positive pressure ventilation (eMethods in the Supplement).
community detection using the first 6 principal components as input, which reflected geographic ancestry.32,36

**Gene-Based Collapsing**
Collapsing refers to an analysis in which variants satisfying specific criteria (qualifying variants [QVs]) are considered equivalent and a statistical association is determined between harboring a QV in a gene or gene set and case-control status. The QV specification determines the collapsing model (eTable 3 in the Supplement).32,34,37,38 For each gene and individual within each cluster, an indicator variable (1/0 states) was assigned if a QV in the gene was present (state 1) or not (state 0) to create a gene × individual matrix for each cluster. From the collapsing matrices of each cluster, we extracted the number of cases and controls with and without a QV per gene and tested for an enrichment in cases or controls.32,34,39-41

**LOF Analysis**
We assessed the optimal gene intolerance threshold associating case-control status and QV carrier status (Figure 1).15 In the rare loss-of-function (LOF) model (eTable 3 in the Supplement) in the combined cohort, 8702 genes with 1860 unique loss-of-function observed/expected upper bound fraction (LOEUF) scores harbored QVs.42 Genes were ordered based on LOEUF scores from most to least intolerant to LOF variation, and gene sets were created for genes less than or equal to each LOEUF threshold. For each gene set, we calculated a P value using the Cochran-Mantel-Haenszel test that signified the statistical association of case-control status and LOF carrier status. Empirical P values were determined by 100 000 permutations of the Cochran-Mantel-Haenszel P values. We used a Bonferroni correction for the 1860 comparisons to assess significance. A 2-sided P < 2.7 × 10⁻⁵ was considered to be statistically significant.

**Figure 1. Loss-of-Function (LOF) Variants in Children With Critical Illness**
We assessed excess LOF variants in our combined case-control cohort (231 cases and 5322 controls) in gene sets defined by loss-of-function observed/expected upper bound fraction (LOEUF) scores. Gene sets to the left of the plot are more intolerant to LOF variation compared with subsets of gene sets to the right of the plot. For each gene set, enrichment was evaluated and a P value was obtained. Empirical P values (y-axis) were determined by 100 000 permutations. The horizontal orange line represents significance using Bonferroni correction. The minimum empirically derived P value, which indicates the most significant enrichment across all LOEUF thresholds, occurs at a LOEUF of 0.680 or less (orange dot). The 1 to 10 labels indicate the intolerance decile based on LOEUF scores, with 1 representing the 10% most intolerant genes and 10 representing the 10% least intolerant genes.
Gene Set Collapsing

We tested associations of case-control status with QV status in these gene sets. Genes with a confirmed phenotype in Online Mendelian Inheritance in Man (4145 genes), Development Disorder Genotype-Phenotype Database (877 genes), and The Human Gene module of SFARI Gene (1898 genes) were considered to be associated with a disease. Combining these yielded 4662 unique disease-associated genes (eMethods in the Supplement). Of these, 2274 harbored a variant in a nonsynonymous model. Of the genes without a known association (n = 13,624), 6478 harbored a variant in a nonsynonymous model. To assess potential associations among the viral respiratory failure cohort, we included gene lists for primary immunodeficiencies, primary ciliary dyskinesia, asthma, and genes associated with susceptibility to life-threatening viral illness (eMethods and eTable 4 in the Supplement). We used a false discovery rate–adjusted P value for multiple comparisons. We performed 22 Cochran-Mantel-Haenszel or Fisher exact tests to determine odds ratios for gene set enrichment testing and defined a significant enrichment at a false discovery rate–adjusted P < .05.

De Novo Mutation Calling, Filtering, and Analysis

We obtained de novo variants for probands with both biological parents enrolled. To test the null hypothesis that our cohort contained an expected number of de novo variants in genes without a known disease association, we used denovolyzeR, version 0.2.0. denovolyzeR is an R package that compares the rate of de novo variants in a data set to an expected rate to assess enrichment.

Results

Diagnostic Outcomes and Phenotypes

This analysis included 285 children (148 [52%] male and 137 [48%] female; 87 [31%] African, 12 [4%] East Asian, 42 [15%] European, 92 [32%] Latino, 1 [0%] Middle Eastern, 14 [5%] South Asian, and 37 [13%] admixed geographic ancestry) with critical illness (Table 1) enrolled from the MSCH/CUIMC or OCME who were 18 years or younger at enrollment (median [range] age, 4.1 [0-18.9] years). Fifty-seven of 285 (20%) achieved a full or partial diagnosis (eTable 5 in the Supplement). Enrolled phenotypes were heterogeneous (eMethods and eTable 6 in the Supplement) but typical of PICU admissions.

Clustering and Collapsing

A total of 231 cases and 5322 controls in the combined cohort, 176 cases and 5180 controls in the unresolved cohort, and 25 cases and 2973 controls in the viral respiratory failure cohort were analyzed after quality control and ancestry matching (eTable 7 in the Supplement) for cluster-based, rare variant collapsing association analyses. For each cohort, 2 models were tested with LOF variants (eTable 3 in the Supplement). The ultrarare LOF model included only variants absent from public data sets, whereas the flex LOF model allowed variants with a minor allele frequency less than 0.1%. No model showed a single gene association with study-wide significance (eFigures 2-10 and eTables 8-16 in the Supplement). The included clusters encompassed samples of European, Latino, and African geographic descent (eFigure 1 in the Supplement). The ultrarare synonymous variant association analysis functioned as a control, and the low genomic-inflation factor indicated adequate ancestry substructure matching (eFigures 2-4 and eTables 8-10 in the Supplement).

LOF Burden in Children With Critical Illness

We assessed whether children with critical illness carried excess LOF variants and whether this association was mediated by genic intolerance (Figure 1). We identified 8702 unique genes that harbored a rare LOF variant in either a case or control (minor allele frequency <0.1%, flex LOF model) (eFigure 8 and eTables 3 and 14 in the Supplement), which comprised 1860 unique LOEUF scores. The LOEUF score quantifies a gene’s intolerance to LOF variation. We created 1860 gene sets
composed of genes less than or equal to the unique LOEUF scores and ordered the gene sets from most to least intolerant. We assessed the association of children with critical illness and carrier status in each gene set. The most statistically significant association between case-control status and LOF variants occurred at a LOEUF threshold of 0.680 ($P = 1.0 \times 10^{-3}$) by permutation.

We asked 3 questions: Are the variants present in public databases? Are the genes that harbor the variants associated with diseases? Is the association driven by children with a genetic diagnosis? We considered genes with a LOEUF score of 0.680 or less (intolerant genes) (Figure 1). In the combined cohort (Figure 2A; eTable 17 in the Supplement), the association between LOF variants and children with critical illness existed in genes with (OR, 1.5; 95% CI, 1.1-2.0; adjusted $P = .58$) and without (OR, 1.6; 95% CI, 1.1-2.1; adjusted $P = 2.9 \times 10^{-3}$) known disease associations. The association was significant for variants absent from public databases (ultrarare) in genes with (OR, 1.7; 95% CI, 1.2-2.2; adjusted $P = 3.8 \times 10^{-3}$) known disease associations. Carrying variants that were rare (minor allele frequency <0.1%) but still present in public data sets was not associated with case-control status.

We removed children who carried variants that partially or fully explained their phenotypes to assess probands with ndES (unresolved cohort) (Figure 2B; eTable 18 in the Supplement). The association between LOF variants and children with critical illness in genes with (OR, 1.5; 95% CI, 1.1-2.0; adjusted $P = .58$) and without (OR, 1.6; 95% CI, 1.2-2.2; adjusted $P = 3.8 \times 10^{-3}$) known disease associations. The association remained significant (OR, 1.8; 95% CI, 1.3-2.5; adjusted $P = 2.3 \times 10^{-3}$) known disease associations. Carrying variants that were rare (minor allele frequency <0.1%) but still present in public data sets was not associated with case-control status.

We considered genes with a LOEUF score of 0.680 or less (intolerant genes) (Figure 1). In the combined cohort (Figure 2A; eTable 17 in the Supplement), the association between LOF variants and children with critical illness existed in genes with (OR, 1.5; 95% CI, 1.1-2.0; adjusted $P = .58$) and without (OR, 1.6; 95% CI, 1.1-2.1; adjusted $P = 2.9 \times 10^{-3}$) known disease associations. The association was significant for variants absent from public databases (ultrarare) in genes with (OR, 1.7; 95% CI, 1.2-2.2; adjusted $P = 3.8 \times 10^{-3}$) known disease associations. Carrying variants that were rare (minor allele frequency <0.1%) but still present in public data sets was not associated with case-control status.

We assessed whether harboring ultrarare LOF variants in intolerant genes was a risk factor for viral respiratory failure (Figure 3; eTable 19 in the Supplement). No association was found between case-control designation and carrier status of LOF variants in genes with a disease association (OR, 1.0; $P > .99$) known disease associations. The association in genes without a known disease association remained significant (OR, 1.8; 95% CI, 1.3-2.5; adjusted $P = 2.3 \times 10^{-3}$). As in the combined cohort, this association extended only to variants absent from public databases (OR, 1.8; 95% CI, 1.3-2.5; adjusted $P = 2.9 \times 10^{-3}$).

**Association of Viral Respiratory Failure With Genetic Risk**

We assessed whether harboring ultrarare LOF variants in intolerant genes was a risk factor for viral respiratory failure (Figure 3; eTable 19 in the Supplement). No association was found between case-control designation and carrier status of LOF variants in genes with a disease association (OR, 1.0; $P > .99$) known disease associations. The association in genes without a known disease association remained significant (OR, 1.8; 95% CI, 1.3-2.5; adjusted $P = 2.3 \times 10^{-3}$). As in the combined cohort, this association extended only to variants absent from public databases (OR, 1.8; 95% CI, 1.3-2.5; adjusted $P = 2.9 \times 10^{-3}$).

**Figure 2. Gene Set Enrichment Analysis Juxtaposing the Combined Cohort Against the Unresolved Cohort**

A] Combined cohort (231 cases and 5322 controls)

<table>
<thead>
<tr>
<th>Model</th>
<th>Genes, No.</th>
<th>Case participants, %</th>
<th>Control participants, %</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intolerant genes</td>
<td>2143</td>
<td>65.8</td>
<td>52.7</td>
<td>1.80 (1.35-2.41)</td>
<td>$7.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>Disease association</td>
<td>752</td>
<td>32.0</td>
<td>24.7</td>
<td>1.48 (1.10-1.99)</td>
<td>.02</td>
</tr>
<tr>
<td>Disease association UR</td>
<td>614</td>
<td>23.8</td>
<td>15.9</td>
<td>1.72 (1.23-2.37)</td>
<td>$3.8 \times 10^{-3}$</td>
</tr>
<tr>
<td>Disease association rare</td>
<td>333</td>
<td>10.0</td>
<td>10.6</td>
<td>0.94 (0.58-1.47)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>No disease association</td>
<td>1391</td>
<td>48.1</td>
<td>37.4</td>
<td>1.60 (1.21-2.10)</td>
<td>$2.9 \times 10^{-3}$</td>
</tr>
<tr>
<td>No disease association UR</td>
<td>1074</td>
<td>33.3</td>
<td>24.5</td>
<td>1.61 (1.20-2.16)</td>
<td>$3.8 \times 10^{-3}$</td>
</tr>
<tr>
<td>No disease association rare</td>
<td>609</td>
<td>20.3</td>
<td>17.1</td>
<td>1.24 (0.87-1.74)</td>
<td>.33</td>
</tr>
</tbody>
</table>

B] Unresolved cohort (176 cases and 5180 controls)

<table>
<thead>
<tr>
<th>Model</th>
<th>Genes, No.</th>
<th>Case participants, %</th>
<th>Control participants, %</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intolerant genes</td>
<td>2068</td>
<td>64.2</td>
<td>52.3</td>
<td>1.73 (1.25-2.41)</td>
<td>$2.9 \times 10^{-3}$</td>
</tr>
<tr>
<td>Disease association</td>
<td>720</td>
<td>26.7</td>
<td>24.6</td>
<td>1.16 (0.81-1.65)</td>
<td>.58</td>
</tr>
<tr>
<td>Disease association UR</td>
<td>586</td>
<td>17.6</td>
<td>15.9</td>
<td>1.20 (0.78-1.80)</td>
<td>.57</td>
</tr>
<tr>
<td>Disease association rare</td>
<td>318</td>
<td>10.2</td>
<td>10.5</td>
<td>0.98 (0.56-1.61)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>No disease association</td>
<td>1348</td>
<td>50.0</td>
<td>37.0</td>
<td>1.79 (1.31-2.45)</td>
<td>$2.3 \times 10^{-3}$</td>
</tr>
<tr>
<td>No disease association UR</td>
<td>1034</td>
<td>34.7</td>
<td>24.1</td>
<td>1.80 (1.29-2.50)</td>
<td>$2.9 \times 10^{-3}$</td>
</tr>
<tr>
<td>No disease association rare</td>
<td>588</td>
<td>21.6</td>
<td>17.0</td>
<td>1.35 (0.91-1.96)</td>
<td>.26</td>
</tr>
</tbody>
</table>

Variants were divided into those in genes with a disease (disease association) and genes without a known disease association (no disease association) and ultrarare (UR) (absent from external control data sets) and rare (minor allele frequency <0.1% but present in external control data sets). Only genes with a loss-of-function observed/expected upper bound fraction (LOEUF) score of 0.680 or less (intolerant genes) were included. The genes column indicates the number of genes in the gene set with a qualifying variant in either a case or control participant. Pooled odds ratio (ORs), 95% CIs, and false discovery rate-corrected $P$ values were generated from the exact 2-sided Cochran-Mantel-Haenszel test.
95% CI, 0.3-2.6; adjusted \( P > .99 \)). For genes without a known disease association, the association with ultrarare variants was significant (OR, 2.8; 95% CI, 1.1-6.6; adjusted \( P = .04 \)), but the association with rare variants present in public databases was not (OR, 1.9; 95% CI, 0.7-4.9; adjusted \( P = .29 \)). No significant association was found among genes implicated in primary immunodeficiencies, asthma, primary ciliary dyskinesia, or prior studies of viral respiratory failure (eTable 20 in the Supplement).47,55-62

The viral respiratory failure cohort sample size was inadequate to assess gene-specific associations (eFigures 7 and 10 and eTables 13 and 16 in the Supplement). We highlighted 9 candidate intolerant genes without a known disease association harboring ultrarare LOF variants in cases but not controls (eTable 21 in the Supplement). \( \text{CTTN (OMIM 164765)} \) is implicated in lung disease among patients of African descent,63 \( \text{RSBN1L (HGNC 24765)} \) is upregulated in COVID-19,64 and \( \text{RASGRP4 (OMIM 607320)} \) regulates dendritic and mast cell function.65,66

**De Novo Enrichment in Children With Critical Illness**

Among the 57 resolved or partially resolved probands, 46 were trios, and among the 227 unresolved probands, 114 were trios (Table 1). We assessed enrichment of de novo variants in genes without a known disease association (Table 2; eTable 8 in the Supplement).48,50 De novo LOF variants were enriched in unresolved probands (observed, 14; expected, 6.8; enrichment, 2.05; \( P = .01 \)) but not in resolved probands (observed, 3; expected, 2.8; enrichment, 1.09; \( P = .52 \)). These data support the hypothesis that despite ndES probands in the PICU likely carry LOF risk variants.

**Figure 3. Children With Viral Respiratory Failure Harbor Ultrarare Loss-of-Function Genetic Risk Factors**

<table>
<thead>
<tr>
<th>Model</th>
<th>Genes, No.</th>
<th>Case participants, %</th>
<th>Control participants, %</th>
<th>OR (95% CI)</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease association</td>
<td>494</td>
<td>24.0</td>
<td>24.2</td>
<td>0.99 (0.32-2.59)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>No disease association</td>
<td>828</td>
<td>56.0</td>
<td>35.6</td>
<td>2.31 (0.97-5.63)</td>
<td>.08</td>
</tr>
<tr>
<td>No disease association UR</td>
<td>576</td>
<td>44.0</td>
<td>22.2</td>
<td>2.76 (1.13-6.57)</td>
<td>.04</td>
</tr>
<tr>
<td>No disease association rare</td>
<td>356</td>
<td>28.0</td>
<td>16.8</td>
<td>1.92 (0.68-4.86)</td>
<td>.29</td>
</tr>
<tr>
<td>Synonymous</td>
<td></td>
<td>9</td>
<td>3.0</td>
<td></td>
<td>.58</td>
</tr>
<tr>
<td>Missense</td>
<td>11</td>
<td>5.0</td>
<td>5.0</td>
<td>0.54</td>
<td>.99</td>
</tr>
<tr>
<td>LOF</td>
<td>3</td>
<td>2.0</td>
<td>1.0</td>
<td>1.09</td>
<td>.52</td>
</tr>
<tr>
<td>Missense and LOF</td>
<td>14</td>
<td>2.3</td>
<td>1.0</td>
<td>0.60</td>
<td>.98</td>
</tr>
<tr>
<td>All</td>
<td>23</td>
<td>3.1</td>
<td>1.0</td>
<td>0.74</td>
<td>.95</td>
</tr>
<tr>
<td>Unresolved cohort (n = 114)</td>
<td></td>
<td>12</td>
<td>23.0</td>
<td>0.52</td>
<td>.99</td>
</tr>
<tr>
<td>Synonymous</td>
<td>47</td>
<td>5.0</td>
<td>5.0</td>
<td>0.93</td>
<td>.72</td>
</tr>
<tr>
<td>Missense</td>
<td>14</td>
<td>2.0</td>
<td>1.0</td>
<td>2.05</td>
<td>.01</td>
</tr>
<tr>
<td>LOF</td>
<td>61</td>
<td>7.6</td>
<td>1.0</td>
<td>1.06</td>
<td>.35</td>
</tr>
<tr>
<td>All</td>
<td>73</td>
<td>7.7</td>
<td>1.0</td>
<td>0.94</td>
<td>.71</td>
</tr>
</tbody>
</table>

Abbreviation: LOF, loss of function.

\( \text{Table 2. Enrichment of De Novo Variants in Children With Critical Illness in Genes Without a Known Disease Association} \)

Data include 25 cases compared with 2973 controls. Qualifying variants were divided into those in genes with a disease (disease association) and genes without a known disease association (no disease association) and ultrarare (UR) (absent from external control data sets) and rare (minor allele frequency <0.1% but present in external control data sets). Only genes with a loss-of-function observed/expected upper bound fraction (LOEUF) score of 0.680 or less (intolerant genes) were included. The genes column indicates the number of genes in the gene set with a qualifying variant in either a case or control participant. Odds ratios (ORs), 95% CIs, and false discovery rate-corrected \( P \) values were generated with the Fisher exact test.
Discussion

In this study, the first to our knowledge of exomes of critically ill children without a genetic diagnosis, we found that ndES may identify disease-risk variants in genes without a known disease association. The diagnostic rate of our cohort was 20% (Table 1). Carrying LOF variants in genes with a LOEUF score of 0.680 or less (intolerant genes) was significantly associated with critical illness (Figure 1). In the combined cohort, intolerant genes with and without a known disease association harbored excess LOF variants (Figure 2A). Among children without a genetic diagnosis, only associations with genes without a known disease association remained (Figure 2B). Children with viral respiratory failure harbored excess LOF variants in intolerant genes without a known disease association (Figure 3). Among undiagnosed children with critical illness, we found more de novo LOF variants than expected in genes without a known disease association (Table 2).

Our diagnostic rate was lower than that of other PICU cohorts, but our inclusion criteria allowed enrollment whether or not the proband was suspected to have a monogenic condition. Prior studies enrolled patients with a high pretest probability of diagnosis. We hypothesized that children with critical illness, regardless of genetic diagnosis, harbored excess LOF variants among genes intolerant to LOF variation but did not know the degree of intolerance. We used an unbiased approach to finding the optimal intolerance threshold. The most significant association occurred at a LOEUF threshold of 0.680 or less, which approximately represents the most intolerant tertile of genes. A statistically significant threshold identifying genes in the most intolerant tertile of the genome supports the hypothesis that critically ill children are more likely to harbor LOF variants regardless of genetic diagnosis. The LOEUF value in the PICU population can be compared with other cohorts. Stillbirth cases undergoing a similar analysis had a peak threshold of 0.239. Stillbirth is a more severe phenotype; therefore, we might expect the LOEUF threshold to capture more intolerant genes than when children survived the neonatal period. The connection among intolerance, disease onset, and exposure (eg, viral infection) should be further explored.

Excess LOF variants among critically ill children with ndES in genes without a known disease association implies that these variants are risk factors for their phenotype, but no diagnosis is made because the gene-disease association is not understood. Comparing the rate of carrier status of ultrarare LOF variants in intolerant genes in the unresolved cohort (34.7% cases vs 24.1% controls) suggests that approximately 10% of these cases will ultimately receive a genetic diagnosis after gene-disease association discovery. Among children with viral respiratory failure, a finding of excess LOF variants (Figure 3) indicates that larger sample sizes could uncover gene-disease associations. Associations with ultrarare variants support the hypothesis that rare diseases (eg, pediatric critical illness) are driven by variants absent in healthy data sets.

Limitations

This study has some limitations. Case sample sizes are small, which biases the findings toward spurious associations. The diagnostic assessment was not held to the regulatory standards of a clinical laboratory, included only single-nucleotide variants and insertion/deletions, included additional genetic testing inconsistently, and analyzed the OCME phenotype with restricted phenotype data. Thus, inclusion of potentially resolved probands in the unresolved cohort may have depressed associations in the unresolved cohort. The enrollment in the PICU was not comprehensive, so the diagnostic rate should be interpreted with caution. Given the limited phenotype data, including the OCME probands in the viral respiratory failure cohort, may bias associations toward genes of multisystem disease rather than isolated viral respiratory failure. Early respiratory viral infection is linked to wheezing and asthma. Given the challenges of diagnosing asthma in young children, viral respiratory failure associations may be confounded with genes for asthma susceptibility. For the control population, we could not exclude a history of critical illness, which would bias associations toward controls but only marginally given the phenotype's rarity. The preponderance of probands of African geographic descent from the OCME (eTable 1 in the...
Supplement) is concerning given high rates of mortality among Black infants. With limited information available, accounting for nongenetic confounds was challenging. Although generalizability may be questioned because this is a single-center study, the cohort was typical of PICUs (eTable 6 in the Supplement), suggesting the genetic findings are likely found broadly.

Conclusions

This genetic association study found that critically ill children harbored excess LOF variants. Among children without a genetic diagnosis, these variants were ultrarare and in genes without a known disease association. Associations remained true among a subset with viral respiratory failure despite a small sample size. Among children without a genetic diagnosis, we found excess de novo LOF variants in genes without a known disease association. Findings from this study indicate that critically ill children harbor ultrarare risk variants detected by ES even in the absence of genetic diagnoses.
Funding/Support: This work was supported by grant TL1TR001875 from the National Center for Advancing Translational Science (Dr Motelow), grant RO1NS095803 from the National Institute of Neurological Disorders and Stroke (Dr Kernie), and the Thrasher Research Fund (Dr Motelow).

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: Katherine Schlosser Metitiri, MD, MA, Division of Critical Care and Hospital Medicine, Department of Pediatrics, Columbia University Irving Medical Center, NewYork-Presbyterian Morgan Stanley Children's Hospital, New York, New York, assisted in determining pediatric intensive care unit admission status. She was not compensated for her work.

REFERENCES


nonsense mutation.

58 and genetic theories.

57 J Exp Med

60 diagnosis and management practices.

61 Genet

60

55 2008.09.017

56 Clin Exp Immunol

64 2022.02.002

63

58 NEJM199501193320301

62 Associates. Asthma and wheezing in the first six years of life.

64 N Engl J Med


SUPPLEMENT.
eTable 1. Probands recruited from the Office of the Chief Medical Examiner of New York City (OCME)
eTable 2. Summary of MSCH Respiratory Failure Cohort
eTable 3. Collapsing Model Qualifying Variant Definition
eTable 4. Genes Associated With Viral Respiratory Failure
eTable 5. Summary of Diagnostic Variants
eTable 6. Top Human Phenotype Ontology Terms
eTable 7. Samples Removed During Data Cleaning
eTable 8. Top 10 Genes in Ultra-Rare Synonymous Model in Combined Cohort
eTable 9. Top 10 Genes in Ultra-Rare Synonymous Model in Unresolved Cohort
eTable 10. Top 10 Genes in Ultra-Rare Synonymous Model in Respiratory Failure Cohort
eTable 11. Top 10 Genes in Ultra-Rare Loss-of-Function Model in Combined Cohort
eTable 12. Top 10 Genes in Ultra-Rare Loss-of-Function Model in Unresolved Cohort
eTable 13. Top 10 Genes in Ultra-Rare Loss-of-Function Model in Respiratory Failure Cohort
eTable 14. Top 10 Genes in Flex Loss-of-Function Model in Combined Cohort
eTable 15. Top 10 Genes in Flex Loss-of-Function Model in Unresolved Cohort
eTable 16. Top 10 Genes in Flex Loss-of-Function Model in Respiratory Failure Cohort
eTable 17. Data for Gene-Set Enrichment Analysis for Combined Cohort
eTable 18. Loss-of-Function Burden in Children With Critical Illness but Without a Causative Genetic Diagnosis
eTable 19. Loss-of-Function Burden in Children With Viral Respiratory Failure in Genes With and Without a Known Disease Association
eTable 20. Loss-of-Function Burden in Children With Viral Respiratory Failure in Immunodeficiency and Primary Ciliary Dyskinesia Gene Lists
eTable 21. Candidate Genes Proposed to Harbor Risk Variants Associated With Risk for Viral Respiratory Failure
eTable 22. Capture Kits Used in Whole Exome Sequencing Generation
eFigure 1. Geographic Ancestry and Clustering for Combined Cohort
eFigure 2. Ultra-Rare Synonymous Model Quantile-Quantile Plot for Combined Cohort
eFigure 3. Ultra-Rare Synonymous Model Quantile-Quantile Plot for Unresolved Cohort
eFigure 4. Ultra-Rare Synonymous Model Quantile-Quantile Plot for Respiratory Failure Cohort
eFigure 5. Ultra-Rare Loss-of-Function Model Quantile-Quantile Plot for Combined Cohort
eFigure 6. Ultra-Rare Loss-of-Function Model Quantile-Quantile Plot for Unresolved Cohort
eFigure 7. Ultra-Rare Loss-of-Function Model Quantile-Quantile Plot for Respiratory Failure Cohort
eFigure 8. Flex Loss-of-Function Model Quantile-Quantile Plot for Combined Cohort
eFigure 9. Flex Loss-of-Function Model Quantile-Quantile Plot for Unresolved Cohort
eFigure 10. Flex Loss-of-Function Model Quantile-Quantile Plot for Respiratory Failure Cohort
emMethods. Supplemental Methods
eReferences