

# Identifying and predicting severe bronchiolitis profiles at high risk for developing asthma: Analysis of three prospective cohorts

Michimasa Fujioji,<sup>a,\*</sup> Orianne Dumas,<sup>b</sup> Kohei Hasegawa,<sup>a</sup> Tuomas Jartti,<sup>c,d,e</sup> and Carlos A. Camargo<sup>a</sup>

<sup>a</sup>Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, 125 Nashua Street, Suite 920, Boston, MA 02114-1101, USA

<sup>b</sup>Équipe d'Épidémiologie Respiratoire Intégrative, Université Paris-Saclay, UVSQ, Université Paris-Sud, Inserm, CESP, Villejuif 94807, France

<sup>c</sup>PEDEGO Research Unit, Medical Research Center, University of Oulu, Oulu, Finland

<sup>d</sup>Department of Pediatrics and Adolescent Medicine, Oulu University Hospital, Oulu, Finland

<sup>e</sup>Department of Pediatrics and Adolescent Medicine, University of Turku and Turku University Hospital, Turku, Finland

## Summary

**Background** Bronchiolitis is the leading cause of infants hospitalization in the U.S. and Europe. Additionally, bronchiolitis is a major risk factor for the development of childhood asthma. Growing evidence suggests heterogeneity within bronchiolitis. We sought to identify distinct, reproducible bronchiolitis subgroups (profiles) and to develop a decision rule accurately predicting the profile at the highest risk for developing asthma.

**Methods** In three multicenter prospective cohorts of infants (age < 12 months) hospitalized for bronchiolitis in the U.S. and Finland (combined  $n = 3081$ ) in 2007–2014, we identified clinically distinct bronchiolitis profiles by using latent class analysis. We examined the association of the profiles with the risk for developing asthma by age 6–7 years. By performing recursive partitioning analyses, we developed a decision rule predicting the profile at highest risk for asthma, and measured its predictive performance in two separate cohorts.

**Findings** We identified four bronchiolitis profiles (profiles A–D). Profile A ( $n = 388$ ; 13%) was characterized by a history of breathing problems/eczema and non-respiratory syncytial virus (non-RSV) infection. In contrast, profile B ( $n = 1064$ ; 34%) resembled classic RSV-induced bronchiolitis. Profile C ( $n = 993$ ; 32%) was comprised of the most severely ill group. Profile D ( $n = 636$ ; 21%) was the least-ill group. Profile A infants had a significantly higher risk for asthma, compared to profile B infants (38% vs. 23%, adjusted odds ratio [adjOR] 2.57, 95% confidence interval [CI] 1.63–4.06). The derived 4-predictor (RSV infection, history of breathing problems, history of eczema, and parental history of asthma) decision rule strongly predicted profile A—*e.g.*, area under the curve [AUC] of 0.98 (95% CI 0.97–0.99), sensitivity of 1.00 (95% CI 0.96–1.00), and specificity of 0.90 (95% CI 0.89–0.93) in a validation cohort.

**Interpretation** In three prospective cohorts of infants with bronchiolitis, we identified clinically distinct profiles and their longitudinal relationship with asthma risk. We also derived and validated an accurate prediction rule to determine the profile at highest risk. The current results should advance research into the development of profile-specific preventive strategies for asthma.

**Copyright** © 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Keywords:** asthma; bronchiolitis; latent class analysis; prediction; phenotypes; virus

## Introduction

Bronchiolitis is the leading cause of infant hospitalizations in the U.S., contributing to approximately 110,000 hospitalizations annually.<sup>1</sup> In addition to the large acute morbidity, 30% of infants hospitalized with

bronchiolitis (“severe bronchiolitis”) will subsequently develop childhood asthma.<sup>2–4</sup> However, there is no effective approach to identify the subgroup of infants at highest risk for asthma—during an important period of airway development.

While bronchiolitis has been conventionally viewed as a single disease condition with a similar mechanism,<sup>5,6</sup> growing evidence supports the heterogeneity of bronchiolitis. Indeed, recent studies have

\*Corresponding author.

E-mail address: [mfujioji@mgh.harvard.edu](mailto:mfujioji@mgh.harvard.edu) (M. Fujioji).

### Research in context

#### *Evidence before this study*

Bronchiolitis is the leading cause of infant hospitalizations in the U.S and Europe. Viral bronchiolitis during infancy is a major risk factor for the development of childhood asthma. As the prevention of asthma likely depends on early (*e.g.*, infancy) identification of high-risk children, our still limited ability has hindered efforts to prevent childhood asthma in this high-risk population. We searched PubMed for studies published from inception to August 2021, with the search terms “asthma”, or “bronchiolitis”, or “latent class analysis”, or “prediction”, or “phenotype”, or “endotype”, or “respiratory syncytial virus”, or “rhinovirus” with no language restrictions. No study has developed a decision rule that accurately predicts clinical profiles of bronchiolitis at a higher risk for asthma by using routinely-available clinical characteristics.

#### *Added value of this study*

By applying a clustering approach to data from three multicenter prospective cohorts of 3081 infants with severe bronchiolitis, we identified four clinically-distinct, reproducible profiles. These profiles were associated with differential risks for developing asthma. We also developed a simple decision rule predicting the profile that has the highest risk for developing asthma. This decision tree showed high accuracy (area under the curve = 0.96) and was validated. Implications of all the available evidence.

#### *Implications of all the available evidence*

Base for the early identification of high-risk patients during early infancy. Our data should advance research into the development of profile-specific preventive strategies for asthma in this large population of children.

reported clinical phenotypes (or “profiles”)<sup>7–9</sup> with differential risks of subsequently developing recurrent wheeze<sup>10</sup> and asthma.<sup>11</sup> Additionally, our team has identified several molecular phenotypes (or endotypes) within respiratory syncytial virus (RSV)<sup>12</sup> and rhinovirus bronchiolitis.<sup>13</sup> Studies have also reported between-virus differences in the host response and molecular characteristics (*e.g.*, transcriptomic profiles.<sup>14,15</sup>) among infants with bronchiolitis. While these data suggest the presence of distinct bronchiolitis subtypes (particularly, the subtype at higher risk for asthma development), their discoveries cannot be directly applied to the clinical setting. No study has developed a decision rule based on routinely-available clinical characteristics that accurately predicts the high risk subgroup. As the prevention of asthma likely depends on early (*e.g.*, infancy)

identification of high-risk children, our still limited ability has hindered efforts to prevent childhood asthma.

To address this knowledge gap, we analyzed data from three multicenter prospective cohort studies of infants with severe bronchiolitis to address the following aims: (1) to identify distinct, reproducible profiles of severe bronchiolitis, (2) to investigate their relationship with the subsequent risk for developing asthma, and (3) to develop and validate a decision rule that accurately predicts the profile at the highest risk for asthma.

## Methods

### Study design, setting, and participants

The current study combines data from three multicenter, multi-year prospective cohorts of infants with severe bronchiolitis. These studies were performed as part of the Multicenter Airway Research Collaboration (MARC) by using a similar protocol<sup>10,16,17</sup> with two performed in the U.S. (MARC-30 USA, MARC-35) and the other in Finland (MARC-30 Finland). Details of the design, setting, participants, and data collection of these studies have been reported previously<sup>10,16,17</sup> and are summarized in Table E1.

In brief, the first study—called the 30th Multicenter Airway Research Collaboration (MARC-30 USA)—was performed at 16 sites across 12 U.S. states during three consecutive bronchiolitis seasons in 2007–2010 (Table E1).<sup>16</sup> MARC-30 USA initially enrolled children (aged <2 years) first hospitalized with an attending physician diagnosis of bronchiolitis. The diagnosis of bronchiolitis was made according to the American Academy of Pediatrics bronchiolitis guidelines, defined as an acute respiratory illness with a combination of rhinitis, cough, tachypnea, wheezing, crackles, or chest retractions.<sup>6</sup> The second study—MARC-35—was performed at 17 sites across 14 U.S. states in 2011–2014.<sup>10</sup> Using a similar standardized protocol as the MARC-30 study, MARC-35 enrolled infants (aged <1 year) first hospitalized with bronchiolitis. Lastly, the third study—MARC-30 Finland—was performed at three sites across Finland in 2008–2010.<sup>11,16,17</sup> With the use of the standardized protocol, MARC-30 Finland enrolled children (aged <2 years) first hospitalized with bronchiolitis.

In these studies, all patients were treated at the discretion of the treating physicians. The institutional review board at each participating hospital approved the study with a written informed consent obtained from the parent or guardian. To harmonize the three cohorts, which had different ages of enrollment (aged <1 vs <2 years), the current study focused on the 3,081 infants (aged <1 year) with severe bronchiolitis (Fig. E1).

### Data collection and quality assurance

Investigators conducted a structured interview that assessed patients' demographic characteristics, medical and family history, and details of the bronchiolitis course during the hospitalization using a standardized protocol.<sup>17</sup> Additionally, nasopharyngeal specimens were collected within 24 h of hospitalization, and tested for respiratory viruses (*e.g.*, RSV, rhinovirus) using real-time polymerase chain reaction (PCR) at Baylor College of Medicine (Houston, TX, USA).<sup>17</sup> To ascertain longitudinal outcomes after the index hospitalization for bronchiolitis, in MARC-35 and MARC-30 Finland, interviewers began interviewing parents/guardians by telephone in addition to medical record review by trained physicians (Table E1). In MARC-30 Finland, the participants were also linked to a national social insurance database (Social Insurance Institution of Finland database).<sup>18</sup> All data were reviewed at the EMNet Coordinating Center at Massachusetts General Hospital (Boston, MA, USA), and site investigators were queried about missing data and discrepancies identified by manual data checks.

### Clinical outcome

The clinical outcome of interest was the development of asthma by age 6–7 years. In MARC-35, asthma was defined using a commonly-used epidemiologic definition<sup>19</sup>): physician-diagnosis of asthma, with either asthma medication use (*e.g.*, albuterol inhaler, inhaled corticosteroids) or asthma-related symptoms (*e.g.*, wheezing, nocturnal cough) in the preceding year. In MARC-30 Finland, asthma was ascertained by linking to the national social database and defined as a repeated purchase of asthma control medication in the preceding year.<sup>18</sup>

### Statistical analysis

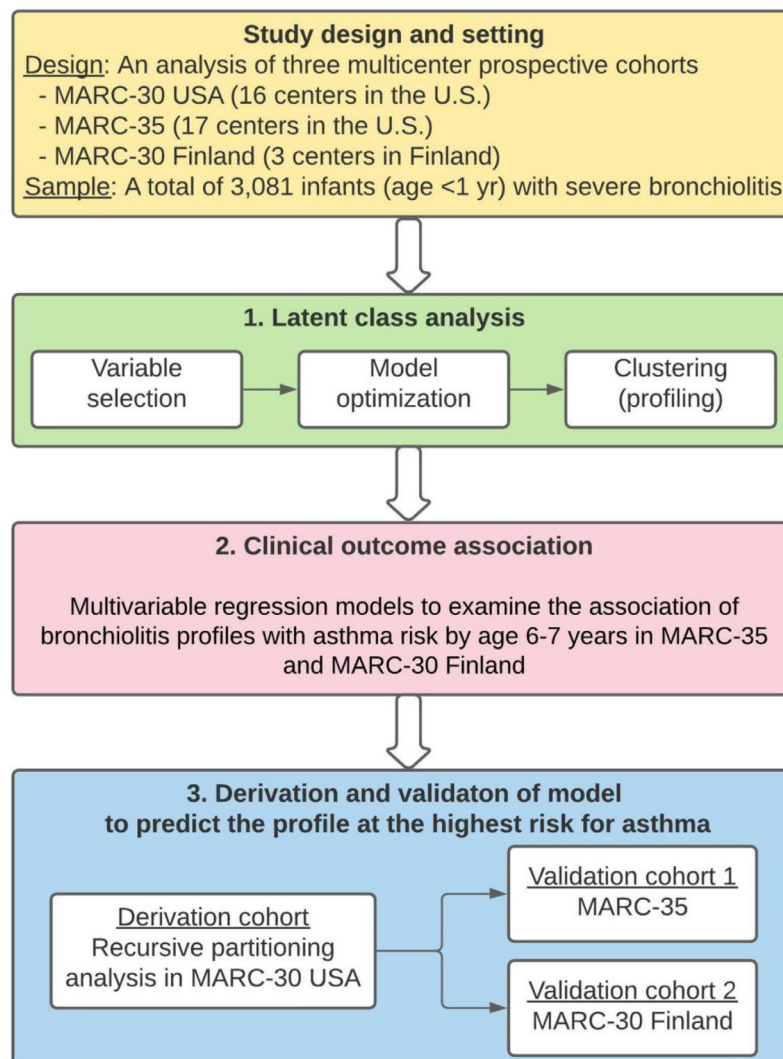
The aims of the current study are to (1) identify bronchiolitis profiles among infants with severe bronchiolitis (*i.e.*, clustering), (2) relate them to the risk of asthma development, and (3) develop and validate a decision rule that predicts the profile at the highest risk for asthma (*i.e.*, prediction). The analytic workflow is summarized in Figure 1.

First, to identify bronchiolitis profiles, we performed a latent class analysis using the data of medical history, clinical course, and viral etiology in the three cohorts combined by using R *poLCA* package.<sup>20</sup> Latent class analysis is a clustering method that identifies more homogeneous subgroups (profiles) of patients from a large set of clinical characteristics.<sup>10,11,16</sup> We selected nine candidate variables (history of breathing problems, history of eczema, wheeze, cough, and inadequate oral intake, and retractions at presentation, hospital length-of-stay, RSV, and rhinovirus) based on previous

studies<sup>10,11,16</sup> and availability of data across three cohorts. Breathing problems were defined as a child having cough that wakes him/her at night and/or causes emesis, or when the child has wheezing or shortness of breath without a cough.<sup>10,21</sup> In the LCA, we examined models with 2–6 classes and determined an optimal model by using the lowest Bayesian information criterion and the mean class membership probability. The final model assigned each infant to the class (profile) for which he/she had the highest membership probability.

Second, to examine the longitudinal association of the identified profiles with the asthma risk (in MARC-35 and MARC-30 Finland), we constructed multivariable logistic regression models adjusting for potential confounders (age at enrollment, sex, and race/ethnicity) in each cohort separately. Results from the two cohorts were then meta-analyzed using fixed-effect models under the assumption that the true effect size would be similar in both studies. In the sensitivity analysis, we repeated the analysis in each of these cohorts separately. We also constructed models adjusting for eight additional covariates (delivery mode, gestational age, breastfeeding, daycare attendance, number of household siblings, maternal smoking during pregnancy, exposure to environmental tobacco smoke, and household income) and accounting for patient clustering at the hospital level with the use of a generalized estimating equation. Additionally, we also performed a meta-analysis in MARC-35 and MARC-30 Finland combined using random-effect models. Both multivariable logistic regression models were performed with complete case analysis.

Third, to develop a decision rule that accurately predicts the profile with the highest asthma risk, we selected predictors based on the variables used in the LCA model, clinical plausibility, and availability in most prehospitalization settings (*e.g.*, parental history of asthma). Then, in the derivation cohort (MARC-30 USA), we conducted a binary recursive partitioning analysis using these candidate predictors with R *rpart* package.<sup>22</sup> We chose the final prediction tree before applying the results to two separate validation cohorts (MARC-35 and MARC-30 Finland) separately. In both derivation and validation cohorts, we estimated the prediction performance by using the area under the receiver-operating-characteristic curve (AUC) and confusion matrix results (*i.e.*, sensitivity, specificity, positive predictive value, and negative predictive value), with corresponding 95% confidence intervals. We also performed another analysis to examine the incremental benefit of predictive ability by including rhinovirus detection by PCR. To compare the AUC between the prediction rules, we used Delong's test. We analyzed the data using R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).



**Figure 1.** Analytic workflow.

The analytic cohort consists of a total of 3,081 infants (age < 1 year) hospitalized with bronchiolitis (severe bronchiolitis) from three multicenter prospective cohort studies: MARC-30 USA, MARC-35, and MARC-30 Finland. Details of the study design, setting, participants, and methods of data collection are summarized in Table E1.; 1 To identify bronchiolitis profiles, we performed a latent class analysis (LCA) using the data of medical history, clinical course, and viral etiology in the three cohorts combined. We selected nine variables based on previous studies and data availability across three cohorts. Then, we optimized the LCA model with the use of the Bayesian information criterion and the mean class membership probability, and identified four mutually-exclusive bronchiolitis profiles.. 1 To examine the longitudinal relationship of the identified bronchiolitis profiles with the risk for developing asthma by age 6-7 years in MARC-35 and MARC-30 Finland, we constructed multivariable logistic regression models adjusting for potential confounders. As infants with a profile B clinically resembled “classic” bronchiolitis and had the largest profile size, this profile served as the reference group. 3 We developed a rule (or a decision tree) to accurately predict the profile at the highest risk for asthma (*i.e.*, profile A) through a binary recursive partitioning analysis in the derivation cohort (MARC-30 USA). Then, we examined the predictive performance in two validation cohorts (MARC-35 and MARC-30 Finland) separately.

Abbreviation: MARC, Multicenter Airway Research Collaboration.

#### Role of the funding source

The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding organizations were not involved in the

collection, management, or analysis the data; preparation or approval of the manuscript; or decision to submit the manuscript for publication. MF, OD, KH, TJ, and CAC had access to the raw data. MF had final responsibility for the decision to submit for publication.

| Variables used in latent class analysis   | All<br>n = 3081 | Profiles           |                     |                    |                    | P-value |
|---|-----------------|--------------------|---------------------|--------------------|--------------------|---------|
|   |                 | A<br>n = 388 (13%) | B<br>n = 1064 (34%) | C<br>n = 993 (32%) | D<br>n = 636 (21%) |         |
| History of breathing problems             | 20%             | 54%                | 16%                 | 19%                | 7%                 | <0.001  |
| History of eczema                         | 14%             | 27%                | 19%                 | 7%                 | 9%                 | <0.001  |
| Wheeze at ED presentation                 | 63%             | 76%                | 100%                | 57%                | 0%                 | <0.001  |
| Cough at ED presentation                  | 87%             | 85%                | 93%                 | 88%                | 77%                | <0.001  |
| Inadequate oral intake at ED presentation | 51%             | 38%                | 28%                 | 90%                | 37%                | <0.001  |
| Retractions at ED presentation            |                 |                    |                     |                    |                    | <0.001  |
| None                                      | 22%             | 18%                | 18%                 | 5%                 | 55%                |         |
| Mild                                      | 45%             | 44%                | 57%                 | 32%                | 45%                |         |
| Severe                                    | 33%             | 38%                | 25%                 | 62%                | 0%                 |         |
| Hospital length-of-stay                   |                 |                    |                     |                    |                    | <0.001  |
| <3 days                                   | 57%             | 74%                | 74%                 | 23%                | 73%                |         |
| 3-6 days                                  | 33%             | 20%                | 26%                 | 53%                | 22%                |         |
| ≥7 days                                   | 10%             | 6%                 | 0%                  | 24%                | 6%                 |         |
| RSV                                       | 75%             | <1%                | 86%                 | 92%                | 75%                | <0.001  |
| Rhinovirus                                | 23%             | 74%                | 15%                 | 19%                | 13%                | <0.001  |

**Table 1: Description of infants with severe bronchiolitis according to profiles (A-D) identified through latent class analysis in three cohort studies.**

Abbreviations: ED, emergency department; RSV, respiratory syncytial virus.

## Results

Of 3538 children (aged <2 years) enrolled into one of the cohorts, 3081 infants (aged <1 year)—1896 in MARC-30 USA, 921 in MARC-35, and 264 in MARC-30 Finland—were eligible for the current analysis (Fig. E1). Overall, the median age was 3.3 months (IQR 1.6–6.2 months) and 41% were female. Additionally, 74% had bronchiolitis with RSV and 23% with rhinovirus; 12% of children had RSV/rhinovirus coinfection.

### Latent class analysis to profile severe bronchiolitis

To identify clinically distinct profiles of infant bronchiolitis, we performed a latent class analysis using the clinical and virology data in the three cohorts combined. The Bayesian information criterion found 4- and 5-class models had an optimal fit (Fig. E2). Of these, the mean class membership probability was highest in the 4-class model (75% for all classes), indicating that participants were assigned to profiles with a fairly high probability with the four bronchiolitis profiles called A, B, C, and D.<sup>10,16,23</sup>

The description of four distinct bronchiolitis profiles is summarized in Table 1. For example, profile A infants (13%) were primarily characterized by a high proportion of previous breathing problems and eczema, low prevalence of RSV, and high prevalence of rhinovirus infection. In contrast, profile B infants (34%) had a low proportion of previous breathing problems, high proportion of wheeze at presentation, high prevalence of RSV, and low prevalence of rhinovirus infection. In many respects, profile B resembled “classic” bronchiolitis.<sup>10,16</sup> Profile C (32%) was the most severely-ill group with a high proportion of inadequate oral intake and severe retraction at presentation and prolonged hospital length-of-stay. By contrast, profile D (21%) was the least-ill group. The distribution of other clinical characteristics between the four profiles is summarized in Table 2. Infants with profile A were older and had a high proportion of parental history of asthma, while none of them had solo-RSV infection. In contrast, infants with profile C were younger with highest acute severity (e.g., mechanical ventilation use). These bronchiolitis profiles identified in the three-cohort pooled analysis were consistent with the previously-reported cohort-specific profiles.<sup>10,16</sup> For example, profile A had a moderate-to-high concordance with the profile with similar clinical characteristics (66% in MARC-30 USA, 91% in MARC-35, 87% in MARC-30 Finland; Table E2).

### Severe bronchiolitis profiles and subsequent development of asthma

The overall risk of asthma development by age 6–7 years was 23% (Table E3). The bronchiolitis profiles had differential risks for developing asthma by age 6–7 years. For example, in the multivariable model, profile A infants had a significantly higher risk compared

to profile B infants who most resembled classic RSV bronchiolitis (38% vs. 23%, adjOR 2.57, 95%CI 1.63–4.06;  $P_{\text{heterogeneity}}=0.82$ ; Figure 2) in MARC-35 and MARC-30 Finland combined. In contrast, those with profile C or D did not have significantly differential risks. In the sensitivity analysis examining each cohort separately, the results were consistent. For example, profile A infants had a significantly higher risk of asthma in MARC-35 (adjOR 2.65, 95%CI 1.59–4.41). In MARC-30 Finland with a limited statistical power, despite the consistent point estimate, there was no significant association with the risk (adjOR 2.32, 95%CI 0.85–6.35). Additionally, with the adjustment for additional confounders and the use of the random-effect model in MARC-35 and MARC-30 Finland combined, the results did not materially change (Tables E4–7).

### Predicting the profile at the highest asthma risk

To develop a decision rule that accurately predicts the profile at the highest risk for developing asthma (i.e., profile A), we performed the recursive partitioning analysis in the derivation cohort (MARC-30 USA). The final rule (decision tree) retained four predictors—RSV infection, history of breathing problems, history of eczema, and parental history of asthma—that together identified infants with profile A with different probabilities (Figure 3A). In the 4-predictor model, for example, the absence of RSV infection and the presence of breathing problem history identified a group with an 82% probability of profile A. In the first validation cohort (MARC-35), the prediction rule achieved a high discriminatory ability with an AUC of 0.98 (95%CI 0.97–0.99; Figure 3B). Similarly, the rule demonstrated a high sensitivity (1.00; 95%CI 0.96–1.00) and specificity (0.90; 95%CI 0.89–0.93; Figure 3C) to identify the high-risk profile. Despite the relatively low proportion of profile A in the cohort (10%; Table 2), the positive predictive value was 0.55 (95%CI 0.48–0.63). Likewise, in the second validation cohort (MARC-30 Finland), the performance of the prediction rule was similarly high (Figure 3B,C).

To determine whether we could further improve the prediction ability, we repeated the recursive partitioning analysis by including another predictor—rhinovirus (Figure 4). In the 5-predictor model, for example, the absence of RSV infection but the presence of rhinovirus and breathing problem history identified a group with a 100% probability of profile A. In the first validation cohort, this prediction rule also achieved an improved discriminatory ability with an AUC of 0.99 (95%CI 0.99–0.99;  $P < 0.001$ ; Figure 4B,C) with high sensitivity (1.00; 95%CI 0.96–1.00) and specificity (0.98; 95%CI 0.97–0.99; Figure 4C). Additionally, the positive predictive value improved to 0.85 (95%CI 0.77–0.91). Likewise, in the second validation cohort

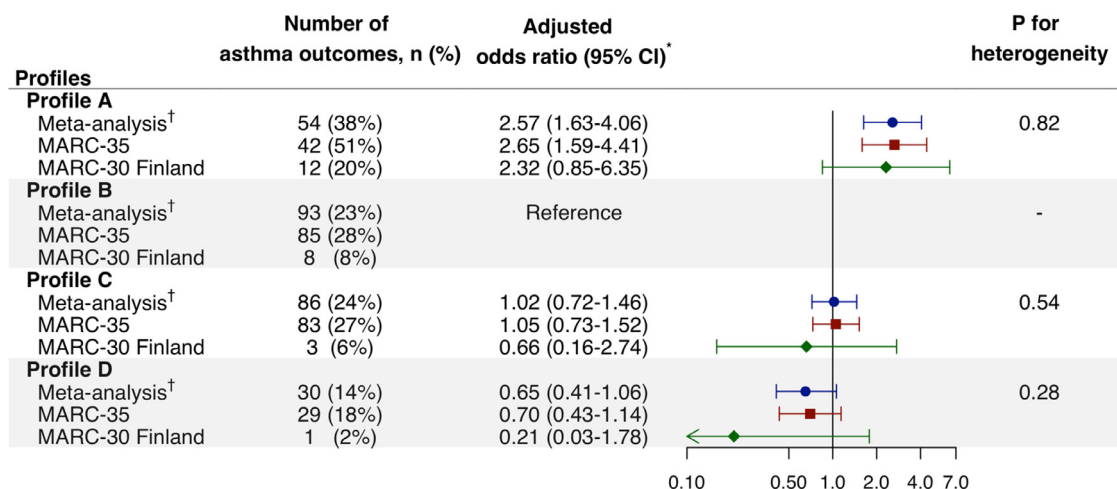
| Patient characteristics*                | All<br>n = 3081 | Profiles           |                     |                    |                    | P-value |
|---|-----------------|--------------------|---------------------|--------------------|--------------------|---------|
|   |                 | A<br>n = 388 (13%) | B<br>n = 1064 (34%) | C<br>n = 993 (32%) | D<br>n = 636 (21%) |         |
| Age (months)                            |                 |                    |                     |                    |                    | <0.001  |
| <2                                      | 31%             | 13%                | 22%                 | 37%                | 49%                |         |
| 2-5.9                                   | 42%             | 35%                | 48%                 | 43%                | 36%                |         |
| 6-11.9                                  | 26%             | 52%                | 29%                 | 20%                | 15%                |         |
| Girls                                   | 41%             | 35%                | 36%                 | 43%                | 48%                | <0.001  |
| Race/ethnicity                          |                 |                    |                     |                    |                    | <0.001  |
| Non-Hispanic white                      | 45%             | 43%                | 44%                 | 45%                | 48%                |         |
| Non-Hispanic black                      | 20%             | 25%                | 25%                 | 16%                | 17%                |         |
| Hispanic                                | 30%             | 30%                | 27%                 | 33%                | 30%                |         |
| Other                                   | 5%              | 3%                 | 4%                  | 5%                 | 5%                 |         |
| Prematurity (32-37 weeks)               | 21%             | 27%                | 20%                 | 22%                | 19%                | <0.001  |
| Breastfeeding                           | 58%             | 59%                | 59%                 | 55%                | 60%                | 0.19    |
| Daycare use                             | 18%             | 18%                | 24%                 | 16%                | 12%                | <0.001  |
| Number of siblings at home              |                 |                    |                     |                    |                    | 0.27    |
| 0                                       | 22%             | 22%                | 24%                 | 19%                | 22%                |         |
| 1                                       | 36%             | 35%                | 35%                 | 38%                | 37%                |         |
| ≥2                                      | 42%             | 43%                | 42%                 | 43%                | 41%                |         |
| Maternal smoking during pregnancy       | 16%             | 18%                | 16%                 | 15%                | 14%                | 0.31    |
| Exposure to environmental tobacco smoke | 13%             | 12%                | 13%                 | 14%                | 11%                | 0.39    |
| Parental history of asthma              | 32%             | 35%                | 34%                 | 29%                | 30%                | 0.046   |
| Viral etiology                          |                 |                    |                     |                    |                    |         |
| RSV only                                | 53%             | 0%                 | 60%                 | 62%                | 61%                | <0.001  |
| Rhinovirus only                         | 7%              | 47%                | 0%                  | 1%                 | 4%                 | <0.001  |
| RSV and rhinovirus                      | 12%             | <1%                | 15%                 | 18%                | 7%                 | <0.001  |
| RSV and other non-rhinovirus            | 9%              | 0%                 | 12%                 | 12%                | 7%                 | <0.001  |
| Rhinovirus and other non-RSV            | 4%              | 26%                | 0%                  | 1%                 | 2%                 | <0.001  |
| Non-RSV and non-rhinovirus              | 14%             | 26%                | 14%                 | 7%                 | 19%                | <0.001  |
| Use of mechanical ventilation           | 7%              | 4%                 | 1%                  | 16%                | 3%                 | <0.001  |
| Admission to intensive care unit        | 16%             | 14%                | 4%                  | 34%                | 9%                 | <0.001  |
| Cohort†                                 |                 |                    |                     |                    |                    | <0.001  |
| MARC-30 USA                             | -               | 12%                | 34%                 | 32%                | 22%                |         |
| MARC-35                                 | -               | 10%                | 35%                 | 36%                | 17%                |         |
| MARC-30 Finland                         | -               | 23%                | 37%                 | 21%                | 19%                |         |

**Table 2: Characteristics of infants with severe bronchiolitis according to profiles (A–D) in three cohort studies.**

Abbreviations: ED, emergency department; MARC, Multicenter Airway Research Collaboration; RSV, respiratory syncytial virus. Percentages may not equal 100 because of rounding and missingness.

\* The variables used in the latent class analysis are shown in [Table 1](#).

† Row percentages are presented only for the cohort variable.



**Figure 2.** Multivariable associations of severe bronchiolitis profiles with risk for developing asthma by age 6.7 years, according to cohorts. \* Multivariable-adjusted logistic regression models adjusting for age at baseline, sex, and race/ethnicity in MARC-35 and age at baseline and sex in MARC-30 Finland. No missing data. Arrow indicates that the 95% CI of the odds ratio exceeds the lower limit of the x-axis. Abbreviations: CI, confidence interval; MARC, Multicenter Airway Research Collaboration.

(MARC-30 Finland), the performance of the prediction rule was similarly high (Figure 4B,C).

## Discussion

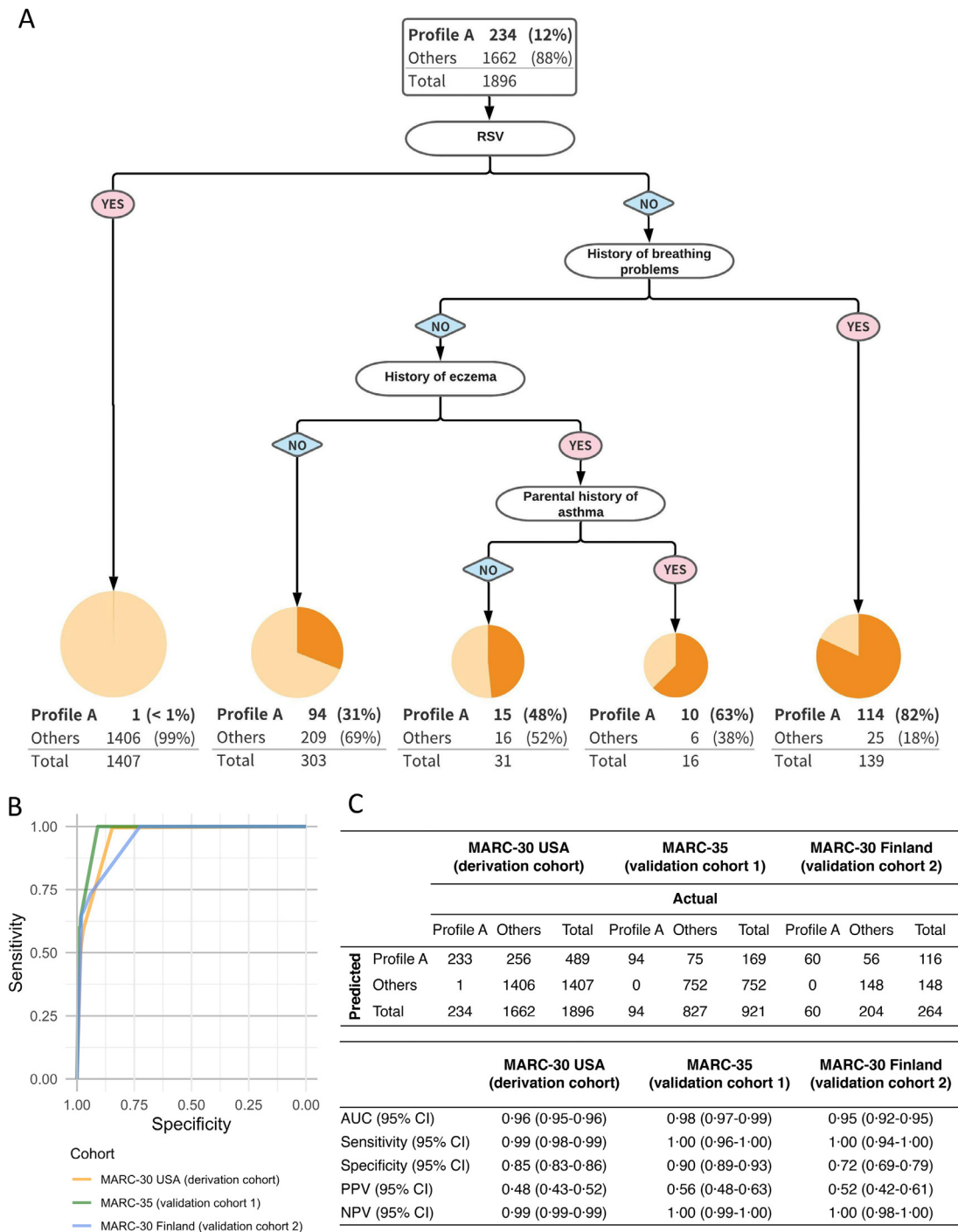
By applying a clustering approach to clinical and virology data from three multicenter prospective cohorts of U.S. and Finnish infants hospitalized for bronchiolitis (combined  $n = 3081$ ), we identified four clinically-distinct, reproducible profiles. We also found that these profiles were longitudinally associated with differential risks for developing asthma by age 6-7 years. In particular, infants with profile A—who were characterized by a history of previous breathing problems and eczema, as well as and non-RSV infection—had the highest risk for asthma development. We also derived and validated a simple decision rule that accurately predicts this high-risk profile with high sensitivity and specificity. To the best of our knowledge, this is the first study that has developed a prediction rule that identifies a specific subgroup of bronchiolitis infants who are at high risk for developing asthma.

Bronchiolitis has been conventionally thought of a single disease entity with similar mechanisms. Indeed, current national guidelines for diagnosis and management of bronchiolitis are based on this assumption.<sup>6,24</sup> However, in agreement with the results of the current study, a growing body of evidence supports the complexity of bronchiolitis, as reflected by heterogeneous clinical characteristics.<sup>7-9</sup> Additionally, research has shown between-virus (RSV and rhinovirus) differences in the transcriptome,<sup>14,15</sup> small non-coding RNAs,<sup>25</sup> cytokines,<sup>26</sup> metabolome,<sup>27-29</sup> and upper airway microbiome<sup>30-32</sup> in infants with bronchiolitis. Furthermore,

our team recently identified molecular phenotypes (endotypes) of severe bronchiolitis with differential risks of subsequent wheezing illness and asthma.<sup>12,13</sup> For example, an integrated omics analysis (using the nasopharyngeal transcriptome, metabolome, and microbiome data) of 221 infants with RSV bronchiolitis identified four biologically-distinct endotypes with differential risks for developing asthma by age five years.<sup>12</sup> These earlier studies collectively support the heterogeneity of bronchiolitis and suggest underlying mechanisms of the bronchiolitis-asthma link. The current study—with the use of data from three prospective cohorts—corroborates earlier findings, and extends them not only by identifying a clinically-distinct reproducible profile at highest risk for asthma but also by developing a decision rule that accurately predicts the high-risk profile.

The mechanisms underlying the observed profiles—particularly profile A characterized by a low prevalence of RSV, high prevalence of rhinovirus, and highest asthma risk—warrant further clarification. Of various non-RSV agents, rhinovirus is the most common causative virus contributing to 20–30% of severe bronchiolitis.<sup>33</sup> Epidemiological studies have also reported that young children with rhinovirus (and its specific genotypes, rhinovirus C) infection are at higher risk of developing asthma.<sup>2,4,19</sup> Additionally, an integrated omics analysis—with the use of virus, nasopharyngeal microbiota, cytokine, and metabolome data—of 122 infants with rhinovirus bronchiolitis has demonstrated that a specific endotype had a higher rate of recurrent wheezing by age three years.<sup>13</sup> Concordant with the profile A observed in the current study, this high-risk endotype (called virus<sup>RV-C</sup>microbiome<sup>Moraxella</sup>T<sub>2</sub><sup>high</sup> endotype)



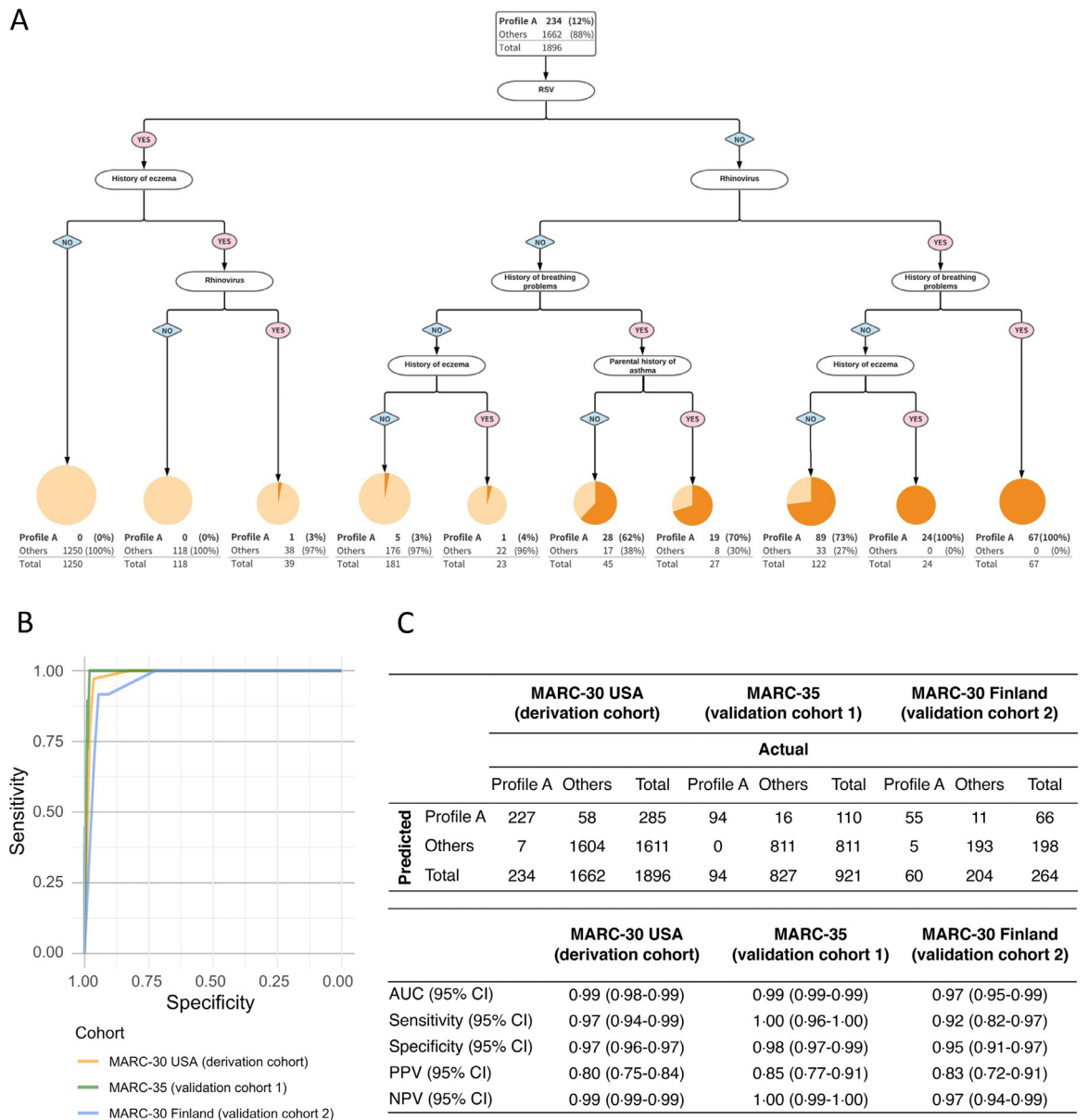


**Figure 3.** Development and validation of 4-predictor model predicting profile at the highest risk for developing asthma. A, Recursive partitioning analysis.

Only the derivation cohort (MARC-30 USA) is shown in the tree. The area of the circle is proportional to the log<sub>2</sub>-transformed size of each branch. The darker orange part in the circle represents the proportion of profile A in each branch. A Receiver-operating-characteristic curves.

To maximize the sum of sensitivity and specificity for all the possible values of the cut-off point, the Youden index method was applied for the following analysis. B Prediction performance

Overall classification counts and characteristics are shown for both the derivation (MARC-30 USA) and validation (MARC-35 and MARC-30 Finland, separately) cohorts.



**Figure 4.** Development and validation of 5-predictor model predicting profile at the highest risk for developing asthma, A. Recursive partitioning analysis, Only the derivation cohort (MARC-30 USA) is shown in the tree. The area of the circle is proportional to the log<sub>2</sub>-transformed size of each branch. The darker orange part in the circle represents the proportion of profile A in each branch. B. Receiver-operating-characteristic curves, To maximize the sum of sensitivity and specificity for all the possible values of the cut-off point, the Youden index method was applied for the following analysis. C. Prediction performance, Overall classification counts and characteristics are shown for both the derivation (MARC-30 USA) and validation (MARC-35 and MARC-30 Finland, separately) cohorts.

Abbreviations: AUC, area under the receiver-operating-characteristic curve; CI, confidence interval; MARC, Multicenter Airway Research Collaboration; NPV, negative predictive value; PPV, positive predictive value; RSV, respiratory syncytial virus.

was characterized by atopy, rhinovirus infection, and increased type 2 inflammation response (including higher levels of T<sub>H</sub>2 [IL-4, IL-5, and IL-13] and epithelial-derived [e.g., IL-33, TSLP] cytokines), as well as inflammatory lipids (e.g., arachidonic acid, hydroxyoctadecadienoic acids, sphingolipids). While profile A had the highest asthma risk in the current study, we also note that other profiles (e.g., profiles B and C) also had a >20% risk of developing asthma. These findings are consistent with the literature showing severe RSV infection in early life leads to decreased lung function and increased lung morbidity.<sup>34,35</sup> Notwithstanding the complexity of bronchiolitis mechanisms, the identification of reproducible profiles across the cohorts is an important finding that facilitates research into identifying profile-specific prevention strategies for asthma. There are also potential explanations for the high prediction ability of our prediction rule in the two validation cohorts. For example, the recursive partitioning analysis incorporates interactions between predictors and nonlinear associations with the bronchiolitis profile of interest. Additionally, our rule was parsimonious (*i.e.*, the use of only four predictors), and hence was able to mitigate overfitting. Moreover, the use of harmonized multicenter data with rigorous quality assurance may have contributed to low bias and variance of the prediction rule development. Nevertheless, we also note that the prediction performance was imperfect. This might be explained, at least in part, by the limited set of predictors in the decision rule (the exclusion of molecular data), the subjectivity of some data elements (e.g., oral intake), and difference in provider's practice patterns. Nonetheless, despite the ongoing challenges of accurately predicting high-risk infants, our observations support cautious optimism that a subgroup of bronchiolitis patients who are likely to develop asthma can be identified before asthma inception, which provides a critical window for preventive interventions.

The current study has potential limitations. Firstly, there were missing data in the current study, which may have led to selection bias in the analysis. However, infants with missing data accounted for only a small proportion. Even if the missing mechanism was not random, it is likely that removing the small number of infants with missingness from the regression models had little impact on our inferences.<sup>36,37</sup> Secondly, the data may be subject to measurement bias. However, these MARC studies were conducted using a standardized protocol.<sup>16,17</sup> Additionally, it is possible that asthma diagnosis is misclassified and that some children are going to develop asthma at a later age. To address this point, study participants are currently being followed up to age nine years in MARC-35. Our inferences may be subject to self-reporting and recall biases despite the rigorous use of structured interview and standardized protocols. Fourthly, there were small differences in the inclusion and exclusion criteria between the cohorts.

Yet, the bronchiolitis profile-asthma risk association was similar both in MARC-35 and MARC-30 Finland. Fifthly, some of the profiles were relatively inconsistent with the previous studies,<sup>10,11,16</sup> while profile A (the primary focus) remained consistent. The potential explanations are that we limited the analysis to infants (age <1 year while a previous study used children aged <2 years) and that the study derived the profiles by using the variables that are consistently available across the cohorts. Sixthly, the profiles were determined by using the data of a single time-point (*i.e.*, hospitalization for bronchiolitis). While longitudinal clinical and environmental (e.g., allergen exposures, climate) data and their potential interaction with the bronchiolitis profiles would also be informative, the study goal was to identify profiles of severe bronchiolitis at early infancy. Additionally, even with single-time point data, we successfully identified clinically distinct profiles that are longitudinally related to asthma risk. Seventhly, the MARC-30 USA cohort—the largest contributor for deriving the profiles—had no longitudinal follow-up. Eighthly, the current prediction models were developed by using some of the same variables that derived the profiles in the same cohorts. Therefore, the capability of these models may have been overestimated, and its examination warrants exertional validation. Finally, while the study samples comprised racially/ethnically- and geographically-diverse infants hospitalized with bronchiolitis, our data may not be generalizable to children with mild-to-moderate bronchiolitis.

In summary, by applying a clustering approach to data from three multicenter prospective cohorts of 3,081 infants with severe bronchiolitis, we identified four clinically-distinct, reproducible profiles. These profiles were associated with differential risks for developing asthma. For example, profile A—characterized by a high proportion of previous breathing problems, eczema, and rhinovirus infection—accounted for 13% of the cohorts and had the highest probability of developing asthma with a 2.5-fold increased odds compared to the “classic” bronchiolitis profile. We also developed and validated a simple decision rule accurately predicting the profile (profile A) that has the highest risk for developing asthma. For clinicians, our observations offer an evidence base for the early identification of high-risk patients during a critical period of airway development—early infancy. For researchers, our data should advance research into the development of profile-specific preventive strategies for asthma (e.g., modification of host response,<sup>2</sup> prophylaxis for severe viral infection<sup>38</sup>) in this large population of children.

#### Contributors

M.F. carried out the main statistical analysis, drafted the initial manuscript, and approved the final manuscript as submitted. O. D. carried out the main statistical

analysis, critically reviewed and revised the initial manuscript, and approved the final manuscript as submitted. K.H. conceptualized the study, obtained funding, supervised the statistical analysis, reviewed and revised the initial manuscript, and approved the final manuscript as submitted.

T. J., critically reviewed, revised the initial manuscript, and approved the final manuscript. C.A.C. conceptualized and designed the study, obtained funding, collected the data, supervised the conduct of study and the analysis, critically reviewed and revised the initial manuscript, and approved the final manuscript as submitted.

#### Funding

This study was supported by grants from the National Institutes of Health (Bethesda, MD): U01 AI-067693, U01 AI-087881, R01 AI-127507, R01 AI-134940, and R01 AI-137091, and UG3/UH3 OD-023253.

#### Data sharing statement

The data that support the findings of this study are available on request from the corresponding author through controlled access to be compliant with the informed consent forms of the MARC studies. The data are not publicly available due to privacy and ethical restrictions.

#### Declaration of interests

Kohei Hasegawa and Carlos A. Camargo, Jr. reports NIH research grants to Massachusetts General Hospital (U01 AI-067693, U01 AI-087881, R01 AI-127507, R01 AI-134940, and R01 AI-137091, and UG3/UH3 OD-023253). The other authors have no financial relationships relevant to this article to disclose.

#### Acknowledgments

We thank the MARC-35 study hospitals and research personnel for their ongoing dedication to bronchiolitis and asthma research, and Ashley F. Sullivan, MS, MPH and Janice A. Espinola, MPH (Massachusetts General Hospital, Boston, MA) for their many contributions to the MARC-35 study.

#### Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2021.101257.

#### References

- 1 Fujiogi M, Goto T, Yasunaga H, et al. Trends in bronchiolitis hospitalizations in the United States: 2000-2016. *Pediatrics* 2019; 144.
- 2 Hasegawa K, Dumas O, Hartert TV, et al. Advancing our understanding of infant bronchiolitis through phenotyping and endotyping: clinical and molecular approaches. *Expert Rev Respir Med* 2017;10:891-9.
- 3 Régnier SA, Huels J. Association between respiratory syncytial virus hospitalizations in infants and respiratory sequelae: systematic review and meta-analysis. *Pediatr Infect Dis J* 2013;32:820-6.
- 4 Liu L, Pan Y, Zhu Y, et al. Association between rhinovirus wheezing illness and the development of childhood asthma: a meta-analysis. *BMJ Open* 2017;7:1-9.
- 5 Rodríguez-Martínez CE, Castro-Rodríguez JA, Nino G, Midulla F. The impact of viral bronchiolitis phenotyping: Is it time to consider phenotype-specific responses to individualize pharmacological management? *Paediatr Respir Rev* 2020;34:53-8.
- 6 Ralston SL, Lieberthal AS, Meissner HC, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics* 2014;134:e1474-502.
- 7 Ferrante G, Fondacaro C, Cilluffo G, Dones P, Cardella F, Corsello G. Identification of bronchiolitis profiles in Italian children through the application of latent class analysis. *Ital J Pediatr* 2020;46:1-8.
- 8 Niu H, Chang AB, Oguoma VM, Wang Z, McCallum GB. Latent class analysis to identify clinical profiles among indigenous infants with bronchiolitis. *Pediatr Pulmonol* 2020;55:3096-103.
- 9 Arroyo M, Salka K, Perez GF, et al. Phenotypic sub-setting of the first episode of severe viral respiratory infection based on clinical assessment and underlying airway disease: a pilot study. *Front Pediatr* 2020;8:1-7.
- 10 Dumas O, Hasegawa K, Mansbach JM, Sullivan AF, Piedra PA, Camargo CA. Severe bronchiolitis profiles and risk of recurrent wheeze by age 3 years. *J Allergy Clin Immunol* 2019;143:1371-9. e7.
- 11 Dumas O, Erkkola R, Bergroth E, Hasegawa K, Mansbach JM, Piedra PA, et al. Severe bronchiolitis profiles and risk of asthma development in Finnish children. *J Allergy Clin Immunol* 2021. <https://doi.org/10.1016/j.jaci.2021.08.035>.
- 12 Raita Y, Pérez-Losada M, Freishtat RJ, et al. Integrated omics endotyping of infants with respiratory syncytial virus bronchiolitis and risk of childhood asthma. *Nat Commun* 2021;12(1):3601. <https://doi.org/10.1038/s41467-021-23859-6>.
- 13 Raita Y, Camargo CA, Bochkov YA, et al. Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma. *J Allergy Clin Immunol* 2021;147(6):2108-17.
- 14 De Steenhuijsen Piters WAA, Heinonen S, Hasrat R, et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am J Respir Crit Care Med* 2016;194:1104-15.
- 15 Fujiogi M, Camargo Jr CA, Bernot JP, et al. In infants with severe bronchiolitis: dual-transcriptomic profiling of nasopharyngeal microbiome and host response. *Pediatr Res* 2020;88(2):144-6.
- 16 Dumas O, Mansbach JM, Jartti T, et al. A clustering approach to identify severe bronchiolitis profiles in children. *Thorax* 2016;71:712-8.
- 17 Hasegawa K, Jartti T, Mansbach JM, et al. Respiratory syncytial virus genomic load and disease severity among children hospitalized with bronchiolitis: multicenter cohort studies in the United States and Finland. *J Infect Dis* 2015;211:1550-9.
- 18 Kela: Chronic asthma and closely related chronic obstructive pulmonary disease. <https://www.kela.fi/laake203> (accessed June 22, 2021).
- 19 Hasegawa K, Mansbach JM, Bochkov YA, et al. Association of rhinovirus C bronchiolitis and immunoglobulin E sensitization during infancy with development of recurrent wheeze. *JAMA Pediatr* 2019;173:544-52.
- 20 Drew A, Linzer JBL. polCA: an R package for polytomous variable latent class analysis. *J Stat Softw* 2013;42:1-29.
- 21 Hasegawa K, Stewart CJ, Celedón JC, Mansbach JM, Tierney C, Camargo CA. Circulating 25-hydroxyvitamin D, nasopharyngeal airway metabolome, and bronchiolitis severity. *Allergy* 2018;73:1135-40.
- 22 T. Therneau, B. Atkinson rpart: recursive partitioning and regression trees. R package version 4.1-15. 2019. <https://cran.r-project.org/package=rpart>.
- 23 Spurk D, Hirschi A, Wang M, Valero D, Kauffeld S. Latent profile analysis: a review and "how to" guide of its application within vocational behavior research. *J Vocat Behav* 2020;120:103445.
- 24 National Collaborating Centre for Women's and Children's Health (UK). Bronchiolitis: diagnosis and management of bronchiolitis in

- children. London: national institute for health and care excellence (UK). 2015; published online June.
- 25 Hasegawa K, Hoptay CE, Epstein S, et al. RSV bronchiolitis versus rhinovirus: difference in nasal airway microRNA profiles and NFκB signaling. *Pediatr Res* 2018;**83**:606–14.
  - 26 Hasegawa K, Hoptay CE, Harmon B, et al. Association of type 2 cytokines in severe rhinovirus bronchiolitis during infancy with risk of developing asthma: a multicenter prospective study. *Allergy* 2019;**74**:1374–7.
  - 27 Fujiogi M, Camargo CAJ, Raita Y, et al. Association of rhinovirus species with nasopharyngeal metabolome in bronchiolitis infants: a multicenter study. *Allergy* 2020;**75**:2379–83.
  - 28 Fujiogi M, Camargo CA, Raita Y, et al. Respiratory viruses are associated with serum metabolome among infants hospitalized for bronchiolitis: a multicenter study. *Pediatr Allergy Immunol* 2020: 1–12.
  - 29 Stewart CJ, Mansbach JM, Piedra PA, Toivonen L, Camargo CA, Hasegawa K. Association of respiratory viruses with serum metabolome in infants with severe bronchiolitis. *Pediatr Allergy Immunol* 2019;**30**:848–51.
  - 30 Toivonen L, Camargo CA, Gern JE, et al. Association between rhinovirus species and nasopharyngeal microbiota in infants with severe bronchiolitis. *J Allergy Clin Immunol* 2019;**143**:1925–8. e7.
  - 31 Mansbach JM, Hasegawa K, Henke DM, et al. Respiratory syncytial virus and rhinovirus severe bronchiolitis are associated with distinct nasopharyngeal microbiota. *J Allergy Clin Immunol* 2016;**137**:1909–13. e4.
  - 32 Rosas-Salazar C, Shilts MH, Tovchigrechko A, et al. Differences in the nasopharyngeal microbiome during acute respiratory tract infection with human rhinovirus and respiratory syncytial virus in infancy. *J Infect Dis* 2016;**214**:1924–8.
  - 33 Hasegawa K, Mansbach JM, Camargo CA. Infectious pathogens and bronchiolitis outcomes. *Expert Rev Anti Infect Ther* 2014;**12**:817–28.
  - 34 Zomer-Kooijker K, Van Der Ent CK, Ermers MJJ, Uiterwaal C, Rovers MM, Bont LJ. Increased risk of wheeze and decreased lung function after respiratory syncytial virus infection. *PLoS One* 2014;**9**:3–8.
  - 35 Bont L, Van Aalderen WMC, Kimpen JLL. Long-term consequences of respiratory syncytial virus (RSV) bronchiolitis. *Paediatr Respir Rev* 2000;**1**:221–7.
  - 36 Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;**8**:3–15.
  - 37 Dong Y, Peng CYJ. Principled missing data methods for researchers. *Springerplus* 2013;**2**:1–17.
  - 38 Basnet S, Palmenberg AC, Gern JE. Rhinoviruses and their receptors. *Chest* 2019;**155**:1018–25.