

# Identifying Asthma Exacerbation-Related Emergency Department Visit Using Electronic Medical Record and Claims Data

Agnes S. Sundaresan<sup>1,2</sup> Gargi Schneider<sup>3</sup> Joy Reynolds<sup>4</sup> H. Lester Kirchner<sup>5,6,7</sup>

<sup>1</sup>Department of Epidemiology and Health Services Research, Geisinger Health System, Danville, Pennsylvania, United States

<sup>2</sup>Medicine Institute, Geisinger Health System, Danville, Pennsylvania, United States

<sup>3</sup>MedPeds Program, Geisinger Medical Center, Danville, Pennsylvania, United States

<sup>4</sup>Lewis Katz School of Medicine at Temple University, Temple University, Philadelphia, Pennsylvania, United States

<sup>5</sup>Department of Biomedical and Translational Informatics, Geisinger Health System, Danville, Pennsylvania, United States

<sup>6</sup>Department of Clinical Sciences, Geisinger Commonwealth School of Medicine, Geisinger Health System, Scranton, Pennsylvania, United States

<sup>7</sup>Department of Pediatrics, Global and Immigrant Health Section, Baylor College of Medicine, Houston, Texas, United States

**Address for correspondence** Agnes S. Sundaresan, MD, MPH, Department of Epidemiology and Health Services Research, Geisinger Health System, 100 North Academy Avenue, Danville, PA 17822-4400, United States (e-mail: ashundaresan@geisinger.edu).

Appl Clin Inform 2018;9:528–540.

## Abstract

**Background** Asthma exacerbation leading to emergency department (ED) visit is prevalent, an indicator of poor control of asthma, and is a potentially preventable clinical outcome.

**Objective** We propose to utilize multiple data elements available in electronic medical records (EMRs) and claims database to create separate algorithms with high validity for clinical and research purposes to identify asthma exacerbation-related ED visit among the general population.

**Methods** We performed a retrospective study with inclusion criteria of patients aged 4 to 40 years, a visit to Geisinger ED from January 1, 2006, to October 28, 2013, with asthma on their problem list. Different electronic data elements including chief complaints, vitals, season, smoking, medication use, and discharge diagnoses were obtained to create the algorithm. A stratified random sample was generated to select the charts for review. Chart review was performed to classify patients with asthma-related ED visit, that is, the gold standard. Two reviewers performed the chart review and validation was done on a small subset.

**Results** There were 966 eligible ED visits in the EMR sample and 731 in the claims sample. Agreement between reviewers was 95.45% and kappa statistic was 0.91. Mean age of the EMR sample was 22 years, and mostly white (93%). Multiple models conventionally used in studies were evaluated and the final model chosen included principal diagnosis, bronchodilator, and steroid use for both algorithms, chief complaints for EMR, and secondary diagnosis for claims. Area under the curve was 0.93

## Keywords

- ▶ system improvement
- ▶ disease management
- ▶ clinical research informatics
- ▶ allergy and immunology
- ▶ clinical data management

received  
December 7, 2017  
accepted after revision  
May 28, 2018

DOI <https://doi.org/10.1055/s-0038-1666994>  
ISSN 1869-0327.

© 2018 Georg Thieme Verlag KG  
Stuttgart · New York

License terms



(95% confidence interval: 0.91–0.94) and 0.94 (0.93–0.96), respectively, for EMR and claims data, with positive predictive value of > 94%. The algorithms are visually presented using nomograms.

**Conclusion** We were able to develop two separate algorithms for EMR and claims to identify asthma exacerbation-related ED visit with excellent diagnostic ability and varying discrimination threshold for clinical and research purposes.

## Background and Significance

Asthma exacerbation, characterized by an acute worsening of asthma symptoms, is quite prevalent, with 1.6 million emergency department (ED) visits a year in the United States.<sup>1,2</sup> These are associated with significant morbidity in terms of abrupt decline in physical function, lung function, and health-related quality of life, including missed school and work days, increased health care utilization, and mortality.<sup>3–8</sup> Multiple predictors for future risk of asthma exacerbation have been found but recent exacerbation history, especially leading to ED visit or hospitalization, is the strongest predictor of future asthma exacerbations.<sup>9–11</sup> Evidence suggests that currently available therapy prevents exacerbation during treatment.<sup>12,13</sup> Therefore, asthma exacerbation leading to ED visit is a potentially preventable clinical outcome and is of great interest both to health services research as an indicator of quality of care to the patient, and to epidemiologic studies.

It has been previously shown that less than a fifth of the asthma population was responsible for more than four-fifths of the total direct costs emphasizing the need for better management of a high-risk cohort.<sup>14</sup> For the purposes of quality improvement, we need performance measures that are easily identified using clear *administrative or clinical (electronic medical record [EMR]) criteria*, actionable based on evidence-based guidelines without ambiguity for most patients, and reliable. Commonly used Healthcare Effectiveness Data and Information Set asthma quality metrics pertain to controller medication filling and controller-to-total asthma medication ratio.<sup>15</sup> However, the use and response to asthma medication is variable unlike other chronic diseases and these process measures are inadequate to identify the risk of health care utilization.<sup>16</sup> Asthma exacerbation leading to ED visit is a good performance metric in the asthma population as it is the outcome that needs to be identified, managed, and avoided for the previously mentioned reasons. As performance measures are definitive standards of care against which the care provided is judged, extensive care needs to be taken in defining such a measure.<sup>17</sup>

Identification of asthma-related ED visits in studies has been conventionally done by International Classification of Diseases, Ninth Revision (ICD-9) discharge codes, chief complaints, and medication use in EMR and claims data.<sup>18–29</sup> A systematic review of validated methods to capture acute bronchospasm using administrative or claims data showed that there are only two studies with validation characteristics among the 38 studies using an algorithm.<sup>30</sup> An EMR algorithm

based on chief complaints implemented in the Southern U.S. showed sensitivity of 45%, specificity of 92%, positive predictive value (PPV) of 79%, and negative predictive value (NPV) of 70%.<sup>31</sup> A study based on claims data showed that claims failed to identify 29% of encounters with asthma diagnoses and 45% of nebulization procedures administered during encounters; 30% of documented asthma prescriptions were not associated with filed claims and vice versa.<sup>32</sup> Another EMR study performed in the Midwest since the systematic review, based on Bayesian network which used electronic information available at the time of ED visit like age, respiratory rate, chief complaint, oxygen saturation, and acuity level, and historical data like past medical history, medications, and billing codes, yielded a PPV of 69.9%.<sup>33</sup> Severe chronic lower respiratory disease exacerbations which combined asthma and other respiratory diseases as measured by discharge diagnosis code had a predictive value of 85 to 95% based on the threshold for gold standard.<sup>34</sup>

In spite of the lack of data to correctly identify ED visit related to asthma exacerbation, the above conventional methods are used by health systems to evaluate their performance using EMR data, by health plans to track utilization using claims data, and by researchers to assess outcome in epidemiologic studies using either data.<sup>35–39</sup> Routine operation of health care systems provides tremendous electronic information at patient level, but this needs to be harnessed and adapted for research and administrative purposes. Developing algorithms using clinical and administrative data elements help to identify various health outcomes of interest. Given the need for the use of asthma exacerbation-related ED visit as a performance measure for systems-based practice and for epidemiologic studies based on health care databases, there is an urgent need to develop an algorithm with adequate performance metrics, based on any available data source—EMR or claims.

## Objective

We propose to utilize multiple data elements available in the EMR and claims database to create separate algorithms with high validity for clinical and research purposes to identify asthma exacerbation-related ED visit among the general population.

## Methods

### Study Overview

This is a retrospective study of 1,000 ED visits randomly selected out of all eligible visits, for patients in a single health

system in Pennsylvania with asthma, with a subset of these visits also having claims data. Chart review was performed to classify ED visits into those related to asthma exacerbation or otherwise, which was the gold standard. Different data elements, like demographics, chief complaints, vital signs, medications, and discharge diagnoses as available in different databases, were obtained to create separate asthma exacerbation-related ED visit algorithms in EMR and claims by comparing against the gold standard.

### Study Population and Subject Selection

Geisinger is a large integrated health system which serves more than 400,000 primary care patients across a 45-county area in Pennsylvania, United States, with this population being representative of the general population of central and northeastern Pennsylvania.<sup>40</sup> It utilizes EMR software by Epic Systems Corporation. Geisinger Health Plan (GHP) is a full-service regional plan covering approximately 500,000 people.<sup>41</sup> Our inclusion criteria were ED visit to a Geisinger ED from January 1, 2006, to October 28, 2013, of patients who were of age 4 to 40 years with asthma in the problem list (ICD-9 code 493.xx). Exclusion criteria were chronic obstructive pulmonary disease (COPD) (ICD-9 codes 491, 492, 496) or coronary artery disease/heart failure (ICD-9 codes 410–414, 428) or cystic fibrosis (ICD-9 code 277.xx) or bronchiectasis (ICD-9 code 494) in the problem list or ED visits resulting in hospital admission. These were comorbidities which can present with lower respiratory symptoms, making it difficult to attribute an exacerbation to one specific disease condition.

Given that not all patients who visit the Geisinger ED have GHP insurance, we enriched our study sample for ED visits with GHP insurance, providing claims data. We also enriched for ED visits with classifiers of interest, as asthma patients visit ED for asthma-related and nonasthma-related concerns. A total of 4,708 eligible ED visits with both EMR and claims data were identified and we aimed for a total of 1,000 ED visits. The visits were stratified into 15 strata according to the presence or absence of classifiers of interest (primary diagnosis, secondary diagnosis, respiratory complaint, bronchodilator use) in isolation or in combination. First, from these strata, random sample of visits of predetermined proportion (as determined by the size of the stratum in relation to other strata and the probable value of that stratum in predicting asthma exacerbation), or all visits, if the strata were small, were taken for a total  $n = 780$  ED visits. An additional 220 ED visits were randomly identified from similar strata in the eligible 20,379 ED visits in the Geisinger population without claims data, for a total of 1,000 ED visits. This stratified sampling methodology was done to increase the probability of presence of classifiers of interest in the study data set.

### Gold Standard Generation

The research team of allergist, informaticist, and primary care physicians developed a structured chart review abstraction tool by consensus decision-making, which evaluated both the exclusion criteria and the diagnosis of asthma

exacerbation in the ED visit. The abstraction tool included various data elements from problem list, past medical history, history of present illness, physical exam, assessment and plan in the ED notes, medication administration, and progress notes, and was used to conclude the presence or absence of asthma exacerbation during ED visit. Two reviewers were trained on chart abstraction and performed chart reviews. A subset of charts was randomly selected from the study sample for validation of the exacerbation classification by another independent reviewer. The asthma exacerbation status in the ED visit by chart review was considered as the gold standard in algorithm development.

### Definitions of Predictor Variables

Principal and secondary diagnoses of asthma exacerbation were identified by clinically coded ICD-9 code of 493.xx as listed in the encounter or claims. If an ED visit had asthma as both principal and secondary diagnosis, it was only counted for principal diagnosis. Respiratory complaints were identified by the terms “upper respiratory infection,” “cough,” “flu,” “congestion,” “wheezing,” “respiratory distress,” “short of breath,” “asthma,” “asthma attack,” “chest discomfort,” “chest pressure,” “pneumonia,” “chest pain,” “bronchitis,” “hyperventilating,” “hemoptysis,” “cold symptoms,” “airway obstruction,” or “chest tightness” listed in the chief complaint. Bronchodilator use in the EMR data was identified by any acute bronchodilator use (via metered dose inhaler or nebulizer) and systemic steroids by medication class of corticosteroids. Nebulization codes used in the claims data were “aerosol or vapor inhalations” and “airway inhalation treatment.” All the above variables were coded as binary variables. Vital signs used for analysis were the initial readings at the time of the ED visit.

### Statistical Analysis

Separate analyses were done for EMR and claims data elements. The EMR data set included the ED visit subset with claims data, but the data elements for analysis were from EMR. We compared the predictor variables for those with asthma exacerbation and those without in the ED visit using *t*-test for continuous variables and chi-square test for categorical variables. We performed unadjusted and adjusted survey logistic regression analyses to identify predictors of asthma exacerbation in an ED visit. Continuous variables were modeled using both linear and quadratic terms to assess for nonlinearity after centering, and the quadratic term was dropped if there was no quadratic association. Model building was done separately for EMR and claims data, by adding covariates (age, sex, race, season, insurance, respiratory rate, pulse rate, temperature, body mass index [BMI], smoking, principal and secondary diagnoses of asthma, respiratory complaint, bronchodilator use, and corticosteroid use) to the model and looking for conditional significance. The final base model included respiratory complaints listed in chief complaints, principal diagnosis, any short-acting bronchodilator, and steroid (oral or intravenous) use in the ED for EMR data; and principal and secondary diagnoses, nebulization, and steroid (oral or intravenous) use in the ED for claims data.

For analysis using survey methodology, some strata were combined to prevent single sampling unit as recommended.<sup>42</sup> These resulted in 15 strata for EMR data and 11 strata for claims data during analysis (→ **Supplementary Tables S1 and S2**, available in the online version). Weight for each stratum was calculated based on the principle of inverse probability weighting (weight = 1/proportion of patients in the study population for that particular stratum).<sup>40,43,44</sup> Extreme weight was predetermined to be truncated to the next highest weight or to a preset weight of 10.<sup>45,46</sup> Based on this, one extreme weight for EMR data was truncated to 10.02, and to 10 for claims data for regression analysis. However, untruncated weights were used for prevalence estimates. Finite population correction was used as large proportions were sampled. We calculated Cohen's kappa statistic to evaluate the interrater agreement on the chart review conclusion and for assessing agreement between EMR and claims-based measurement of predictor variables.

We were interested in assessing the performance of currently used algorithm elements, both in isolation and combination. We evaluated the performance of 5 models in EMR and 4 models in claims. Model performance discrimination was evaluated using the area under the curve (AUC), which was calculated by nonparametric trapezoidal approximation to the estimated false-positive and true-positive rate points and compared with each other to choose the final predictive model.<sup>47</sup> AUC interpretation for the usefulness of the classifier was considered with the following cut-offs: > 0.9: high accuracy; 0.7 to 0.9: moderate accuracy; and 0.5 to 0.7: low accuracy.<sup>48,49</sup> Nomograms were constructed for the final predictive models to provide visual approach for calculating probabilities of ED exacerbation.<sup>50</sup> We compared EMR models to the claims models for the sample with overlapping data. Models were redone for EMR algorithm using this restricted sample for the sake of comparability and AUC was compared.

We performed sensitivity analyses by using different definitions of the predictor variables. For EMR, we utilized only nebulizer administration in the ED instead of any short-acting bronchodilator. For claims, we utilized additional claims submitted on the same day as ED visit claims, which provided separate principal and secondary diagnoses and medications. We combined both the same-day claim files and ED visit claim files and performed the analysis. A *p*-value of ≤ 0.05 was considered statistically significant. All analyses were done in the Stata statistical software, version 14.2 (StataCorp LP, Texas, United States).

## Results

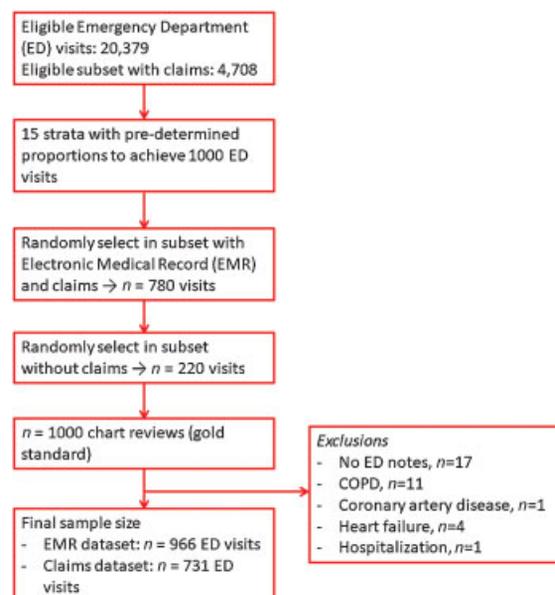
### Validation of Chart Review

An independent reviewer evaluated 66 charts previously reviewed by the initial reviewers (29 for one and 37 for another). Agreement between reviewers was 95.45% (63/66) and the kappa statistic was 0.91. The discrepancies were in missing cardiomyopathy which was an exclusion criterion by one; missing asthma exacerbation in the setting of a fall; and allergic reaction to nuts considered as asthma exacerbation, by another.

### Characteristics of the Study Sample

There were 966 eligible ED visits in the total population and 731 in the GHP population (→ **Fig. 1**) after those excluded by chart review (no ED notes, *n* = 17; COPD, *n* = 11; coronary artery disease, *n* = 1; heart failure, *n* = 4; hospitalization, *n* = 1). The prevalence of asthma exacerbation in the source population meeting the inclusion criteria, calculated using the EMR data, was 10.84% (95% confidence interval: 10.47–11.22). Mean age of the EMR sample was 22 years, mostly white (93%), never smoked (56%), and female (64%) (→ **Table 1**). There were seasonal variations in the ED visits studied, with the most common season being the fall, followed by spring. Overall claims sample was similar to the EMR sample in age, sex, race, season, BMI, and smoking status.

The asthma exacerbation group in the EMR sample compared with the no asthma exacerbation group was younger, less likely to have the ED visit in the summer, and more likely to be a never smoker. The distribution of those with asthma exacerbation was slightly different for EMR and claims sample, with those in claims being younger, female, with more spring visits, and less current smokers compared with EMR. Asthma principal diagnosis was more likely in the asthma exacerbation EMR sample while secondary diagnosis was more likely in the no asthma exacerbation EMR sample. While 6% of EMR sample with asthma principal diagnosis did not have asthma exacerbation, less than 1% of claims sample with principal diagnosis did not have asthma exacerbation. In the EMR sample, most of those with asthma exacerbation had a respiratory complaint (99%), while three-quarters of them (74%) had bronchodilator use and half of them had corticosteroid use (53%) and principal diagnosis (54%). The initial respiratory and pulse rates and temperature were higher in the asthma exacerbation group (*p* < 0.05 for all). In the claims sample, three-quarters of the asthma exacerbation group had nebulization, while two-thirds had principal diagnosis and half of them had corticosteroid use.



**Fig. 1** Study flowchart.

**Table 1** Selected demographic variables by asthma exacerbation status for EMR and claims data

Characteristics	Total (n = 966)	Asthma exacerbation (n = 497)	No asthma exacerbation (n = 469)	p-Value <sup>a</sup>
<b>EMR data</b>				
Age at ED visit (mean, SD in y)	22.04 (10.56)	21.08 (10.97)	23.05 (10.03)	0.004
Age at ED visit (n, column %)				
≤ 18 y	374 (38.72)	211 (42.45)	163 (34.75)	0.014
> 18 y	592 (61.28)	286 (57.55)	306 (65.25)	
Male (n, %)	351 (36.34)	195 (39.24)	156 (33.26)	0.05
White (n, %)	898 (92.96)	461 (92.76)	437 (93.18)	0.80
Season (n, column %)				
Summer	203 (21.01)	82 (16.50)	121 (25.80)	0.003
Fall	334 (34.58)	176 (35.41)	158 (33.69)	
Winter	172 (17.81)	101 (20.32)	71 (15.14)	
Spring	257 (26.60)	138 (27.77)	119 (25.37)	
Has GHP insurance at ED visit (n, %)	751 (77.74)	371 (74.65)	380 (81.02)	0.02
Asthma principal diagnosis (n, %)	296 (30.64)	267 (53.72)	29 (6.18)	< 0.001
Asthma secondary diagnosis (n, %)	326 (33.75)	116 (23.34)	210 (44.78)	< 0.001
Respiratory complaint (n, %)	803 (83.13)	490 (98.59)	313 (66.74)	< 0.001
Bronchodilator use (n, %)	459 (47.52)	366 (73.64)	93 (19.83)	< 0.001
Corticosteroids (n, %)	298 (30.85)	262 (52.72)	36 (7.68)	< 0.001
Bronchodilator use or corticosteroids (n, %)	560 (57.97)	440 (88.53)	120 (25.59)	< 0.001
Initial respiratory rate at ED visit <sup>b</sup> (mean, SD, /min)	19.13 (3.60)	20.05 (3.82)	18.14 (3.06)	< 0.001
Initial pulse rate at ED visit <sup>c</sup> (mean, SD, /min)	97.99 (20.69)	102.33 (21.07)	93.30 (19.21)	< 0.001
Initial temperature at ED visit <sup>d</sup> (mean, /min)	98.38 (0.99)	98.46 (1.07)	98.31 (0.89)	0.02
BMI <sup>e</sup> (kg/m <sup>2</sup> ), mean (SD)	28.28 (10.05)	28.23 (10.08)	28.33 (10.02)	0.88
Smoking <sup>f</sup> (n, column %)				
Never smoker	520 (55.97)	291 (60.75)	229 (50.89)	0.009
Past smoker	150 (16.15)	67 (13.99)	83 (18.44)	
Current smoker	259 (27.88)	121 (25.26)	138 (30.67)	
<b>Claims data</b>				
	<b>Total (n = 731)</b>	<b>Asthma exacerbation (n = 367)</b>	<b>No asthma exacerbation (n = 364)</b>	<b>p-Value<sup>a</sup></b>
Age at ED visit (mean, SD in y)	22.24 (10.46)	20.85 (10.77)	23.65 (9.96)	< 0.001
Age at ED visit (n, column %)				
≤ 18 y	283 (38.71)	162 (44.14)	121 (33.24)	0.002
> 18 y	448 (61.29)	205 (55.86)	243 (66.76)	
Male (n, %)	260 (35.57)	138 (37.60)	122 (33.52)	0.25
White (n, %)	682 (93.30)	343 (93.46)	339 (93.13)	0.86
Season (n, column %)				
Summer	161 (22.02)	61 (16.62)	100 (27.47)	0.002
Fall	246 (33.65)	123 (33.51)	123 (33.79)	
Winter	116 (15.87)	64 (17.44)	52 (14.29)	
Spring	208 (28.45)	119 (32.43)	89 (24.45)	
Asthma principal diagnosis (n, %)	229 (31.33)	226 (61.58)	3 (0.82)	< 0.001

**Table 1** (Continued)

Characteristics	Total (n = 966)	Asthma exacerbation (n = 497)	No asthma exacerbation (n = 469)	p-Value <sup>a</sup>
Asthma secondary diagnosis (n, %)	108 (14.77)	62 (16.89)	46 (12.64)	0.11
Nebulization (n, %)	315 (43.09)	273 (74.39)	42 (11.54)	< 0.001
Corticosteroids (n, %)	213 (29.14)	185 (50.41)	28 (7.69)	< 0.001
Nebulization or steroids (n, %)	372 (50.89)	310 (84.47)	62 (17.03)	< 0.001
BMI <sup>g</sup> (kg/m <sup>2</sup> ), mean (SD)	28.27 (9.75)	28.12 (9.88)	28.42 (9.64)	0.69
Smoking <sup>h</sup> (n, column %)				
Never smoker	411 (57.89)	229 (63.97)	182 (51.70)	0.003
Past smoker	116 (16.34)	54 (15.08)	62 (17.61)	
Current smoker	183 (25.77)	75 (20.95)	108 (30.68)	

Abbreviations: BMI, body mass index; ED, emergency department; EMR, electronic medical record; GHP, Geisinger Health Plan; SD, standard deviation.

<sup>a</sup>t-Test for continuous variables and chi-square test for categorical variables comparing asthma exacerbation and no asthma exacerbation.

<sup>b</sup>Sixteen missing values.

<sup>c</sup>Twenty-two missing values.

<sup>d</sup>Fourteen missing values.

<sup>e</sup>Ninety-eight missing values.

<sup>f</sup>Thirty-seven missing values.

<sup>g</sup>Seventy-one missing values.

<sup>h</sup>Twenty-one missing values.

### Concordance between EMR and Claims Data

We evaluated the concordance in documentation between EMR and claims data among the overlapping 731 ED visits (→ **Supplementary Table S3**, available in the online version). There was 96% concordance for principal diagnosis (perfect agreement using  $\kappa$  statistic); 38% ( $n = 10$ ) of those with principal diagnosis present in EMR but missed by claims had asthma exacerbation while 100% ( $n = 6$ ) for vice versa. There was 77% concordance for secondary diagnosis (moderate agreement); 30% ( $n = 50$ ) of those with secondary diagnosis in EMR and missed by claims had asthma exacerbation while 100% ( $n = 4$ ) for vice versa. There was 84% concordance for nebulization (substantial agreement); 74% ( $n = 50$ ) of those with nebulization use in EMR but missed by claims had asthma exacerbation while 92% ( $n = 44$ ) for vice versa. There was 91% concordance for corticosteroids (substantial agreement); 87% ( $n = 39$ ) of those with steroids use in EMR but missed by claims had asthma exacerbation while 75% ( $n = 15$ ) for vice versa.

### Unadjusted and Adjusted Associations of Predictors

All unadjusted predictors evaluated that showed an association with asthma exacerbation-related ED visit in both claims and EMR data are shown in → **Table 2**, with most of them showing a positive association. In the adjusted analysis, principal diagnosis had the highest odds ratio (OR) among the different predictors in EMR and claims data, followed by bronchodilator use. Among the vital signs, the odds of asthma exacerbation-related ED visit increased by 12% for

every breath/minute increase in the initial respiratory rate (OR = 1.12 [1.05, 1.19]).

### Receiver Operating Characteristic Curves

#### EMR Data

We evaluated the AUC for five receiver operating characteristic (ROC) curves with different combinations of commonly used predictors of asthma exacerbation-related ED visit in the EMR, and the final base model chosen was respiratory complaint, primary diagnosis, any short-acting bronchodilator use, and oral or intravenous steroids (→ **Fig. 2**). The model with initial respiratory rate in addition to the base model had statistically significantly better performance than the base model ( $p = 0.04$ ) with ROC for the former being 0.93 (0.92, 0.95) versus 0.93 (0.91, 0.94) for the latter. At a probability cut-off of 0.55 for the former model, sensitivity was 83.98%, specificity was 82.93%, PPV was 84.15%, and NPV was 82.75%. However, given that the gains in AUC were minimal when compared with the difficulty of obtaining initial vital sign for routine implementation purposes when an algorithm needs to be used in a large scale, we decided to use the simpler latter model as the final model. Different probability cut-offs were evaluated for the final model to optimize the sensitivity and specificity (→ **Table 3**). Probability cut-off of 0.5 had the highest sensitivity of 95.57% and 0.7 had the highest PPV of 94.15%. The regression equation for the final model was: probability of asthma exacerbation =

$$\frac{e^{-4.66 + 1.81 (\text{respiratory complaint}) + 3.37 (\text{primary diagnosis}) + 3.16 (\text{short-acting bronchodilator}) + 1.80 (\text{steroids})}}{1 + e^{-4.66 + 1.81 (\text{respiratory complaint}) + 3.37 (\text{primary diagnosis}) + 3.16 (\text{short-acting bronchodilator}) + 1.80 (\text{steroids})}}$$

**Table 2** Unadjusted and adjusted associations<sup>a</sup> by survey logistic regression<sup>b</sup> of selected predictor variables with asthma exacerbation-related ED visit

Characteristics–EMR data	Unadjusted association OR (95% CI)	Adjusted association OR (95% CI)
Asthma principal diagnosis	46.19 (31.61, 67.50)	29.16 (14.97, 56.81)
Asthma secondary diagnosis	0.70 (0.58, 0.83)	0.85 (0.53, 1.38)
Respiratory complaint	68.37 (34.92, 133.87)	6.09 (2.30, 16.14)
Bronchodilator use	41.05 (30.73, 54.84)	23.67 (12.99, 43.11)
Corticosteroids	21.24 (13.92, 32.40)	6.06 (3.34, 11.01)
Initial respiratory rate	1.29 (1.22, 1.37)	1.12 (1.05, 1.19)
Initial pulse rate	1.03 (1.02, 1.04)	1.01 (1.00, 1.02)
Initial temperature	1.22 (1.05, 1.41)	1.24 (0.98, 1.57)
Smoking		
Never smoker	Reference	Reference
Past smoker	0.66 (0.44, 0.98)	0.57 (0.26, 1.23)
Current smoker	0.56 (0.41, 0.77)	0.66 (0.41, 1.05)
Characteristics–Claims data	Unadjusted association OR (95% CI)	Adjusted association OR (95% CI)
Asthma principal diagnosis	401.88 (228.34, 707.31)	305.56 (158.11, 590.53)
Asthma secondary diagnosis	1.86 (1.47, 2.35)	6.63 (4.53, 9.70)
Nebulization	45.92 (38.55, 54.69)	32.88 (25.39, 42.58)
Corticosteroids	19.75 (15.16, 25.73)	4.39 (2.98, 6.46)

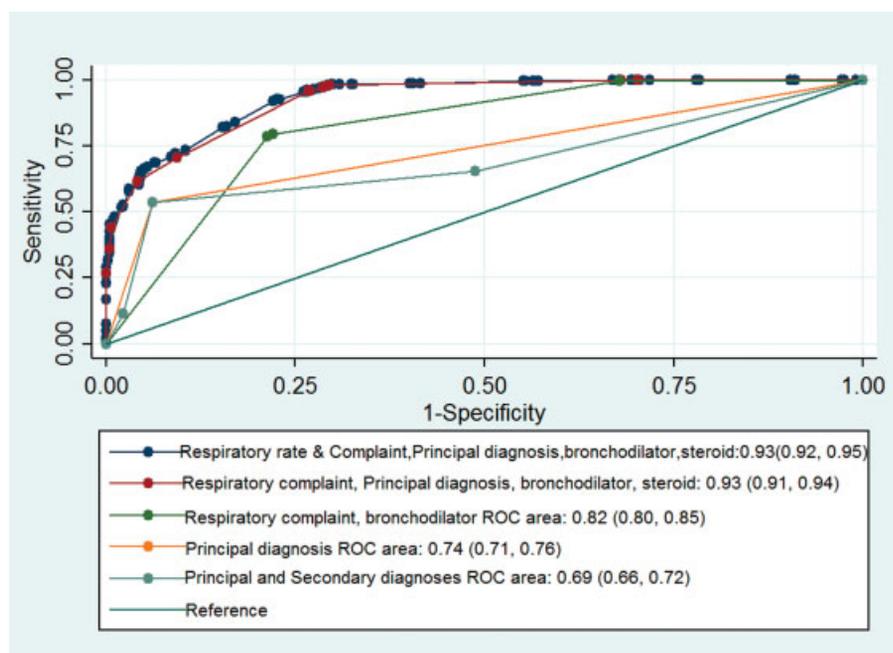
Abbreviations: CI, confidence interval; ED, emergency department; EMR, electronic medical record; OR, odds ratio.

<sup>a</sup>Adjusted associations are for each variable in the final model or added to the final model. Final model variables were respiratory complaints (like cough, congestion, wheezing, short of breath, etc.), principal diagnosis, any short-acting bronchodilator, and steroid (oral or intravenous) use in the ED for EMR data; and principal and secondary diagnoses, nebulization, and steroid (oral or intravenous) use in the ED for claims data.

<sup>b</sup>Survey logistic regression was used for the analysis; weighted based on sampling proportions; accounted for stratified survey sampling.

A nomogram was constructed to calculate the predicted probability of asthma exacerbation-related ED visit (→ Fig. 3). To interpret the nomogram, first the score for each variable in the model should be calculated by lining

up the value for the variable with the score scale in the upper part of the nomogram, and then all the scores should be summed. The resulting total score should be matched to the probability of asthma exacerbation in the lower part of

**Fig. 2** Receiver operating characteristic (ROC) curves for electronic medical record data.

**Table 3** Sensitivity, specificity, positive predictive value, and negative predictive value for the final models in EMR for different asthma exacerbation probability cut-offs

Classification function	Probability cut-offs for EMR data final model		
	0.5	0.6	0.7
Sensitivity	95.57%	70.62%	61.57%
Specificity	73.56%	90.62%	95.95%
Positive predictive value	79.30%	88.86%	94.15%
Negative predictive value	94.01%	74.43%	70.20%

Abbreviation: EMR, electronic medical record.

An extended model with centered age, squared centered age, race, and season as nonordinal variable in addition to the base model was the best model ( $p = 0.046$ ), with ROC for final model being 0.94 (0.93, 0.96) versus 0.95 (0.93, 0.96) for the extended model. Again, for the sake of simplicity for large-scale implementation, we did not choose the extended model for the final model. Different probability cut-offs were evaluated for the final model, with 0.5 being the most sensitive model at 91% and 0.6 having the highest PPV at 95% (→Table 4). Nomogram was constructed to calculate the predicted probability of asthma exacerbation-related ED visit (→Fig. 5).

The regression equation for the final model was: probability of asthma exacerbation =

$$\frac{e^{-3.27 + 5.72 (\text{principal diagnosis}) + 1.89 (\text{secondary diagnosis}) + 3.49 (\text{nebulization}) + 1.48 (\text{steroids})}}{1 + e^{-3.27 + 5.72 (\text{principal diagnosis}) + 1.89 (\text{secondary diagnosis}) + 3.49 (\text{nebulization}) + 1.48 (\text{steroids})}}$$

the nomogram.<sup>51</sup> Binary probability cut-off to indicate asthma exacerbation can be chosen for each study based on the desired test statistics of sensitivity, specificity, and predictive values as in →Table 3. Each patient’s probability of asthma exacerbation obtained from the nomogram can finally be converted to “yes or no” for asthma exacerbation based on the predetermined probability cut-off.

**Claims Data**

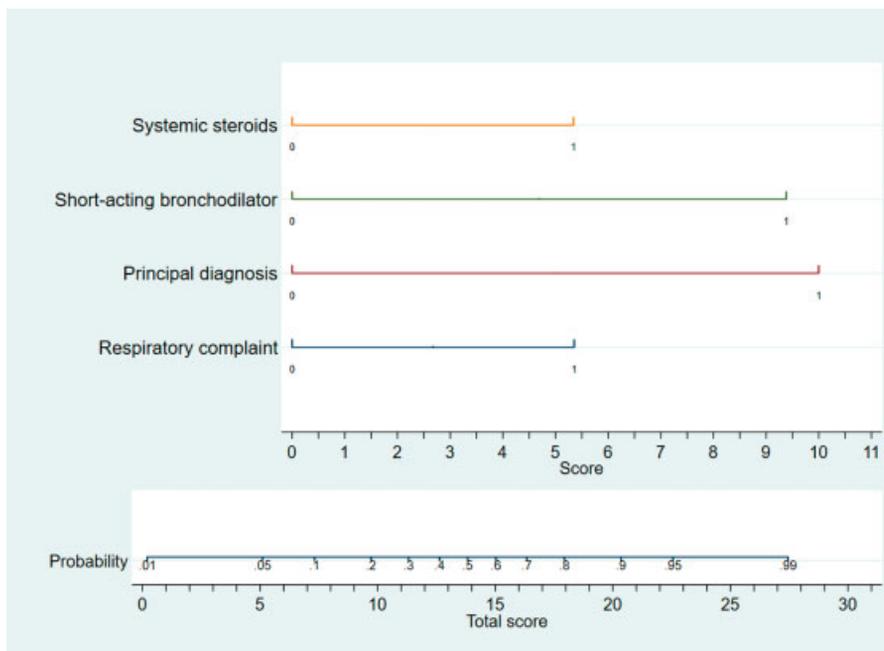
Multiple models conventionally used in studies were evaluated for claims data also, and the final base model chosen was principal and secondary diagnoses, bronchodilator, and oral or intravenous steroid use (→Fig. 4).

**Sensitivity Analysis**

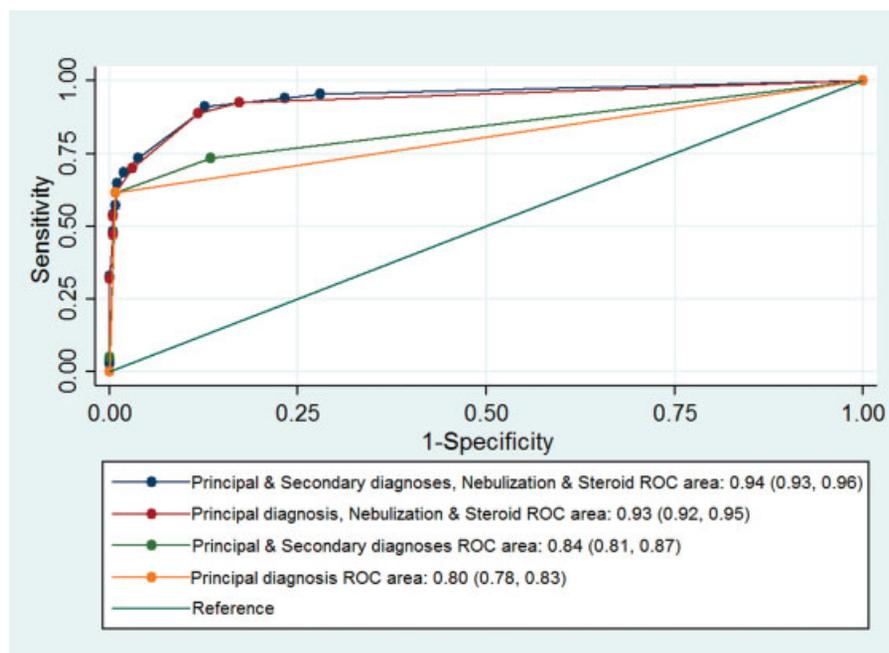
Sensitivity analysis did not change the algorithm performance. Using nebulizer administration instead of any short-acting bronchodilator use for EMR data, reduced the performance of the EMR model nonsignificantly with AUC of 0.92 versus 0.93 ( $p = 0.14$ ). Using same-day claim files along with ED visit claims did not improve the performance of the claims algorithm—AUC of 0.94 for both algorithms,  $p = 0.91$  for the difference.

**Comparison of EMR and Claims Data Algorithms**

We compared the final models from EMR and claims data and there was no statistically significant difference,  $p = 0.54$  (→Fig. 6).



**Fig. 3** Nomogram for electronic medical record data.



**Fig. 4** Receiver operating characteristic (ROC) curves for claims data.

AUC for the model with primary diagnosis only was better for claims data (0.80, 95% confidence interval [0.78, 0.83]) compared with EMR data (0.79 [0.76, 0.81]),  $p = 0.03$ . AUC for the model with primary and secondary diagnoses only was also better for claims data (0.84 [0.81, 0.87]) compared with EMR data (0.77 [0.74, 0.80]),  $p \leq 0.001$ . There was no statistically significant difference in AUC for the model with bronchodilator and steroid only based on EMR (0.88 [0.85, 0.90]) and claims (0.87 [0.84, 0.89]),  $p = 0.45$ . Similarly, no difference was found for model with bronchodilator only. However, AUC for the model with steroid only was better for EMR (0.74 [0.72, 0.77]) than claims (0.71 [0.68, 0.74]),  $p = 0.004$ .

## Discussion

We present the largest study to our knowledge validating algorithms to identify asthma exacerbation-related ED visit in pediatric and adult population with separate algorithms for EMR and claims data. These algorithms use readily

available inputs like chief complaints, diagnosis codes, and medication use. Depending on the proposed utility of the algorithm, different probability cut-offs can be set. While population-based epidemiology studies and quality improvement projects require high sensitivity, performance metrics for health system and genetic epidemiology studies require algorithms with high PPV and specificity. We have provided nomograms which can be utilized with different probability cut-offs for these different purposes. As examples, genetic association study of asthma exacerbation or a performance metric for asthma care will need a high PPV of at least 90%<sup>52</sup> and the possible probability cut-offs for the algorithms would be 0.7 for EMR and 0.6 for claims data to identify asthma exacerbation; however, for a quality improvement project involving escalation of care for asthmatics with an asthma exacerbation leading to ED visit, high sensitivity would be the priority so that most of the eligible patients would receive the needed care; in this situation, probability cut-off of 0.5 might be more appropriate for EMR and claims data. The algorithm could also be integrated into a clinical decision support with a probability cut-off of 0.5, for managing asthma patients with respiratory complaints in an outpatient setting to initiate aggressive management, as a patient with a previous history of asthma exacerbation is at high risk of another exacerbation requiring ED visit.

The only previous algorithms to identify asthma exacerbation leading to ED visit were both performed in pediatric population<sup>31,33</sup> and ours involves both adult and pediatric age ranges. Regarding the performance of our algorithm, as it pertains to generalizability, sensitivity and specificity are fixed test characteristics. However, the predictive values are dependent on the prevalence of asthma exacerbation, which is variable by the region of the United States, ranging from 43.2% in the Northeast, 44.7% Midwest to 48.8% in the South

**Table 4** Sensitivity, specificity, positive predictive value, and negative predictive value for the final models in claims for different asthma exacerbation probability cut-offs

Classification function	Probability cut-offs for claims data final model		
	0.5	0.55	0.6
Sensitivity	91.28%	88.83%	73.57%
Specificity	87.36%	88.19%	96.15%
Positive predictive value	87.93%	88.35%	95.07%
Negative predictive value	90.86%	88.67%	78.30%

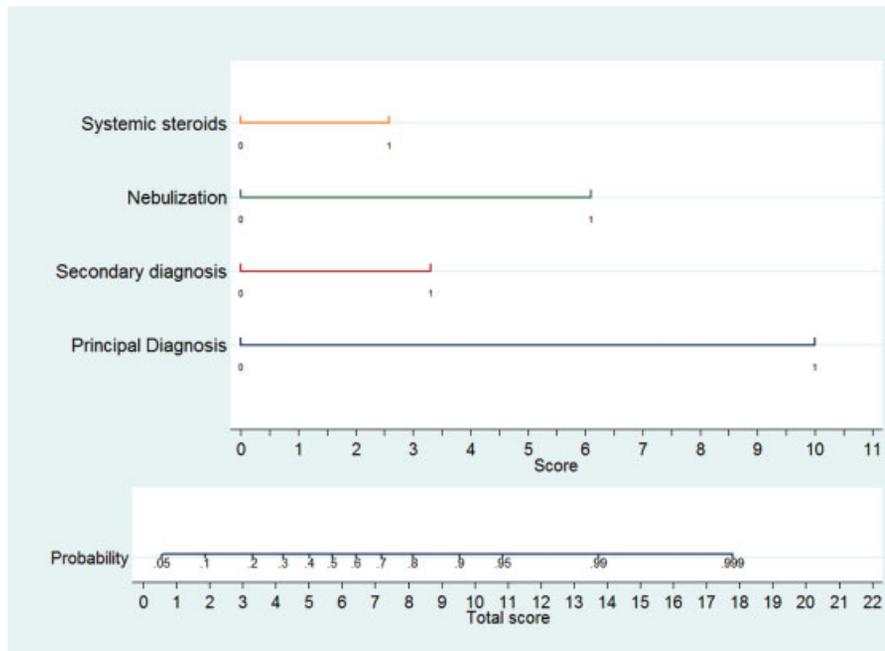


Fig. 5 Nomogram for claims data.

and West.<sup>53</sup> Our study is the first in the Northeast, and has better predictive value for the same sensitivity (94% vs. 79%) even though our asthma attack prevalence is the lowest of the regions of the United States.<sup>31</sup> Our health system also encounters predominantly rural<sup>54</sup> and white population,<sup>9</sup> which has been shown in previous studies to have a lower prevalence of asthma exacerbation and, therefore, could lower the PPV of our algorithm. Despite the availability of effective preventive therapy and guidelines, asthma management and control in the United States is unsatisfactory and costs associated with

asthma are increasing.<sup>55,56</sup> It has been shown that decision support tools, feedback and audit, and clinical pharmacy support were crucial to encourage adherence to asthma guidelines.<sup>57</sup> Our algorithm with acceptable PPV helps to identify the high-risk patients, making these tools possible, and has the potential to improve asthma management.

EMR and claims data each have their advantages and disadvantages, while the former is able to provide better clinical data like vital signs in asthma exacerbation, the latter is able to provide a complete picture of a patient's care. The

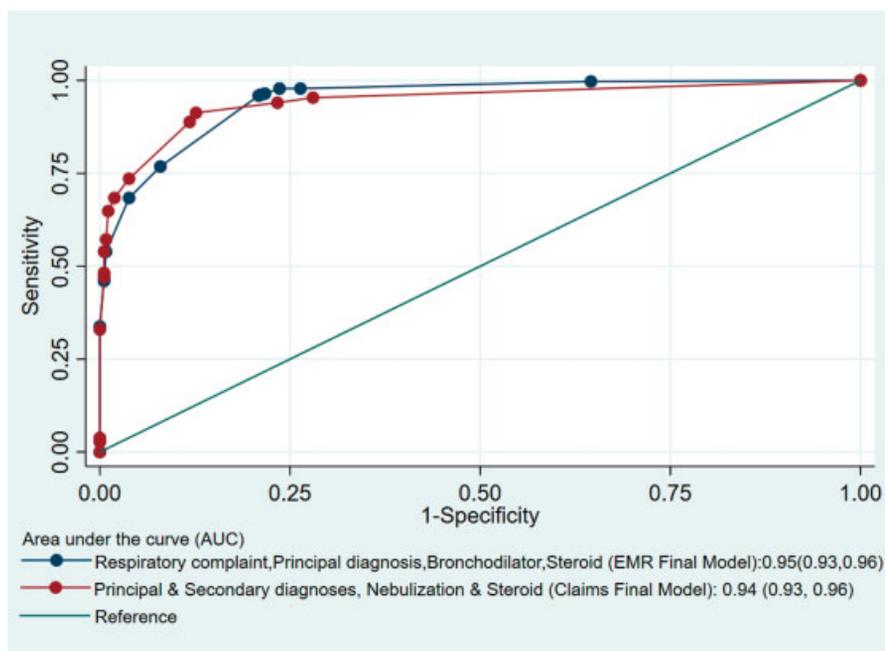


Fig. 6 Comparison of receiver operating characteristic (ROC) curves of the final model for electronic medical record (EMR) and claims data.

concordance between EMR- and claims-based variables for asthma exacerbation identification was variable, with > 90% concordance for primary diagnosis and corticosteroids. When we analyzed how predictive a characteristic that was present in one of the data sets (EMR or claims) but missed by the other was, claim-positive, EMR-negative subset had higher rates of asthma exacerbation for principal diagnosis, secondary diagnosis, and bronchodilator, while EMR-positive, claims-negative subset, had for steroids. Secondary asthma diagnosis was not predictive in the EMR data but predictive in the claims data. Previous studies have compared EMR to claims data and have found variable concordance based on medication or disease condition.<sup>58,59</sup> Models that use diagnostic codes only to predict asthma exacerbation performed better in claims data than EMR, while steroid-only model performed better in EMR data, albeit the AUC of such models were inferior to our final model. This shows the strengths and weaknesses of data elements from EMR and claims data. The final model from EMR and claims were not statistically different. The AUC for the final model for EMR improved from 0.93 (0.91, 0.94) to 0.95 (0.93, 0.96) when the sample used to derive the models changed to a restricted sample with overlapping data. Even though the AUC changed, the change is within acceptable limits of random variability as assessed by the 95% confidence intervals which showed overlapping intervals.

Strengths of our study include separate algorithms for EMR and claims data, validation against a large gold standard population, inclusion of both pediatric and adult population, using appropriate statistical methods to refer to the source population to calculate all performance metrics, and creation of nomograms for use by others. Our study is not without limitations in that the diagnostic codes we used for the algorithm were ICD-9 codes, which at the current time makes our algorithm relevant only for retrospective work, if used as studied. Given the current standard of using ICD-10 codes, the asthma code ICD-10 code J45.xx should be used instead of ICD-9 code 493.xx when using the algorithm, to make it relevant for prospective work. Also, our gold standard did not use peak flow data to confirm asthma exacerbation but we used extensive clinical data to classify our gold standard and validated this.<sup>34</sup> We also oversampled ED visits with classifiers of interest, but we were able to calculate estimates of classification function in the source population, making our predictive values generalizable to an ED asthma population similar to ours, which is predominantly rural and white. We also did not explore data mining and analytics-based techniques for classification such as decision trees, Support Vector Machines, or newer advanced techniques that may provide better performance, and these could be evaluated in future studies. The agreement between EMR and claims data for various data elements could have been influenced by (1) our EMR system, as billing is highly integrated with EMR for Epic Systems Corporation-based software, and (2) by our organization structure which is an integrated health system providing both health care and insurance, influencing generalizability as it pertains to comparison of EMR and claims data elements.

## Conclusion

We were able to create an algorithm based on respiratory complaints listed in chief complaints, principal diagnosis, any short-acting bronchodilator, and steroid (oral or intravenous) use in the ED for EMR data, with an excellent PPV of 94%; and another based on principal and secondary diagnoses, nebulization, and steroid (oral or intravenous) use in the ED for claims data, with an excellent PPV of 95%. We also provide different probability cut-offs to provide a sensitive or specific algorithm based on the needs of a research study or quality improvement program.

## Clinical Relevance Statement

Asthma exacerbation leading to ED visit is prevalent and preventable. Identification of these asthma exacerbations using EMR and claims data has the potential to improve morbidity in asthma by integrated care and better research. This study provides separate algorithms to identify asthma-related ED visit using EMR and claims data with high positive predictive value using readily available data elements like chief complaints, diagnostic codes, and medications.

## Multiple Choice Question

What was the purpose of nomogram in this study?

- To provide a mathematical formula for calculating predicted probability of asthma exacerbation in the ED.
- To provide a visual approach for calculating predicted probability of asthma exacerbation in the ED.
- To provide a mathematical formula for calculating area under the curve for the final model to predict asthma exacerbation in the ED.
- To provide a visual approach for calculating area under the curve for the final model to predict asthma exacerbation in the ED.

**Correct Answer:** The correct answer is option b, to provide a visual approach for calculating predicted probability of asthma exacerbation in the ED. Nomograms were constructed in this study to provide a visual approach for calculating predicted probabilities of asthma exacerbation-related ED visit, although traditionally they have used in medicine to estimate prognosis. The advantage of nomogram is the ability to calculate probabilities based on individual patient's characteristics of different predictor variables. A nomogram can also be integrated into clinical decision support tools and utilized for clinical care. To interpret the nomogram, first the score for each variable in the model should be calculated by lining up the value for the variable with the score scale in the upper part of the nomogram, and then all the scores should be summed. The resulting total score should be matched to the probability of asthma exacerbation in the lower part of the nomogram. Binary probability cut-off to indicate asthma exacerbation can be chosen for each project based on the desired test statistics of sensitivity, specificity, and

predictive values. Each patient's probability of asthma exacerbation obtained from the nomogram can finally be converted to "yes or no" for asthma exacerbation based on the predetermined probability cut-off.

#### Protection of Human and Animal Subjects

Human and/or animal subjects were not included in the project. The study was reviewed by the Geisinger Institutional Review Board.

#### Conflict of Interest

Agnes Sundaresan reported previous grant funding from Teva Pharmaceutical Industries, outside the submitted work. Other authors declare that they have no conflict of interest relevant to this research.

#### Acknowledgments

The authors would like to thank Allison L. Freeman, MD, and Jonathan D. Darer, MD, MPH, for their valuable input into the study conception and conduct.

#### References

- Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention; 2016. Available at: [www.ginasthma.org](http://www.ginasthma.org). Accessed June 19, 2018
- Rui P, Kang K, Albert M. National Hospital Ambulatory Medical Care Survey: 2013 Emergency Department Summary Tables. Available at: [https://www.cdc.gov/nchs/data/ahcd/nhamcs\\_emergency/2013\\_ed\\_web\\_tables.pdf](https://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2013_ed_web_tables.pdf). Accessed June 19, 2018
- Sullivan PW, Ghushchyan VH, Slejko JF, Belozeroff V, Globe DR, Lin SL. The burden of adult asthma in the United States: evidence from the Medical Expenditure Panel Survey. *J Allergy Clin Immunol* 2011;127(02):363–369
- O'Byrne PM, Pedersen S, Lamm CJ, Tan WC, Busse WW; START Investigators Group. Severe exacerbations and decline in lung function in asthma. *Am J Respir Crit Care Med* 2009;179(01):19–24
- Sullivan PW, Smith KL, Ghushchyan VH, Globe DR, Lin SL, Globe G. Asthma in USA: its impact on health-related quality of life. *J Asthma* 2013;50(08):891–899
- Akinbami LJ, Moorman JE, Liu X. Asthma prevalence, health care use, and mortality: United States, 2005–2009. *Natl Health Stat Rep* 2011;32(32):1–14
- Gold LS, Smith N, Allen-Ramey FC, Nathan RA, Sullivan SD. Associations of patient outcomes with level of asthma control. *Ann Allergy Asthma Immunol* 2012;109(04):260–265
- McDonald VM, Gibson PG. Exacerbations of severe asthma. *Clin Exp Allergy* 2012;42(05):670–677
- Sims EJ, Price D, Haughney J, Ryan D, Thomas M. Current control and future risk in asthma management. *Allergy Asthma Immunol Res* 2011;3(04):217–225
- Chippis BE, Zeiger RS, Borish L, et al; TENOR Study Group. Key findings and clinical implications from The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) study. *J Allergy Clin Immunol* 2012;130(02):332–342
- Suruki RY, Daugherty JB, Boudiaf N, Albers FC. The frequency of asthma exacerbations and healthcare utilization in patients with asthma from the UK and USA. *BMC Pulm Med* 2017;17(01):74
- Engelkes M, Janssens HM, de Jongste JC, Sturkenboom MC, Verhamme KM. Medication adherence and the risk of severe asthma exacerbations: a systematic review. *Eur Respir J* 2015;45(02):396–407
- FitzGerald JM, Gibson PG. Asthma exacerbations. 4: prevention. *Thorax* 2006;61(11):992–999
- Malone DC, Lawson KA, Smith DH. Asthma: an analysis of high-cost patients. *Pharm Pract Manag Q* 2000;20(01):12–20
- Yong PL, Werner RM. Process quality measures and asthma exacerbations in the Medicaid population. *J Allergy Clin Immunol* 2009;124(05):961–966
- James T, Fine M. Monitoring asthma control using claims data and patient-reported outcomes measures. *P&T* 2008;33(08):454–466
- Quality of Care Outcomes Research in CVD Stroke Working Groups. Measuring and improving quality of care: a report from the American Heart Association/American College of Cardiology First Scientific Forum on Assessment of Healthcare Quality in Cardiovascular Disease and Stroke. *Circulation* 2000;101(12):1483–1493
- Darrow LA, Hess J, Rogers CA, Tolbert PE, Klein M, Sarnat SE. Ambient pollen concentrations and emergency department visits for asthma and wheeze. *J Allergy Clin Immunol* 2012;130(03):630–638.e4
- Zeiger RS, Schatz M, Li Q, Zhang F, Purdum AS, Chen W. Step-up care improves impairment in uncontrolled asthma: an administrative data study. *Am J Manag Care* 2010;16(12):897–906
- Tsai CL, Lee WY, Hanania NA, Camargo CA Jr. Age-related differences in clinical outcomes for acute asthma in the United States, 2006–2008. *J Allergy Clin Immunol* 2012;129(05):1252–1258.e1
- Rosychuk RJ, Voaklander DC, Klassen TP, Senthilselvan A, Marrie TJ, Rowe BH. Asthma presentations by children to emergency departments in a Canadian province: a population-based study. *Pediatr Pulmonol* 2010;45(10):985–992
- Rowe BH, Voaklander DC, Wang D, et al. Asthma presentations by adults to emergency departments in Alberta, Canada: a large population-based study. *Chest* 2009;135(01):57–65
- Héguy L, Garneau M, Goldberg MS, Raphoz M, Guay F, Valois MF. Associations between grass and weed pollen and emergency department visits for asthma among children in Montreal. *Environ Res* 2008;106(02):203–211
- Himes BE, Kohane IS, Ramoni MF, Weiss ST. Characterization of patients who suffer asthma exacerbations using data extracted from electronic medical records. *AMIA Annual Symposium proceedings/AMIA Symposium*; 2008:308–312
- Hirshon JM, Shardell M, Alles S, et al. Elevated ambient air zinc increases pediatric asthma morbidity. *Environ Health Perspect* 2008;116(06):826–831
- Mohr LB, Luo S, Mathias E, Tobing R, Homan S, Sterling D. Influence of season and temperature on the relationship of elemental carbon air pollution to pediatric asthma emergency room visits. *J Asthma* 2008;45(10):936–943
- Moorman JE, Rudd RA, Johnson CA, et al; Centers for Disease Control and Prevention (CDC). National surveillance for asthma—United States, 1980–2004. *MMWR Surveill Summ* 2007;56(08):1–54
- Silverman RA, Ito K, Stevenson L, Hastings HM. The relationship of fall school opening and emergency department asthma visits in a large metropolitan area. *Arch Pediatr Adolesc Med* 2005;159(09):818–823
- Florin TA, Carron H, Huang G, Shah SS, Ruddy R, Ambroggio L. Pneumonia in children presenting to the emergency department with an asthma exacerbation. *JAMA Pediatr* 2016;170(08):803–805
- Sharifi M, Krishanswami S, McPheeters ML. A systematic review of validated methods to capture acute bronchospasm using administrative or claims data. *Vaccine* 2013;31(Suppl 10):K12–K20
- Sanders DL, Gregg W, Aronsky D. Identifying asthma exacerbations in a pediatric emergency department: a feasibility study. *Int J Med Inform* 2007;76(07):557–564
- Bronstein JM, Santer L, Johnson V. The use of Medicaid claims as a supplementary source of information on quality of asthma care. *J Healthc Qual* 2000;22(06):13–18
- Dexheimer JW, Abramo TJ, Arnold DH, et al. Implementation and evaluation of an integrated computerized asthma management system in a pediatric emergency department: a randomized clinical trial. *Int J Med Inform* 2014;83(11):805–813

- 34 Oelsner EC, Loehr LR, Henderson AG, et al. Classifying chronic lower respiratory disease events in epidemiologic cohort studies. *Ann Am Thorac Soc* 2016;13(07):1057–1066
- 35 Minard J, Dostaler SM, Olajos-Clow JG, et al. Development and implementation of an electronic asthma record for primary care: integrating guidelines into practice. *J Asthma* 2014;51(01):58–68
- 36 Beard JD, Beck C, Graham R, et al. Winter temperature inversions and emergency department visits for asthma in Salt Lake County, Utah, 2003–2008. *Environ Health Perspect* 2012;120(10):1385–1390
- 37 May L, Carim M, Yadav K. Adult asthma exacerbations and environmental triggers: a retrospective review of ED visits using an electronic medical record. *Am J Emerg Med* 2011;29(09):1074–1082
- 38 Berger WE, Legorreta AP, Blaiss MS, et al. The utility of the Health Plan Employer Data and Information Set (HEDIS) asthma measure to predict asthma-related outcomes. *Ann Allergy Asthma Immunol* 2004;93(06):538–545
- 39 Schatz M, Zeiger RS, Yang SJ, et al. Persistent asthma defined using HEDIS versus survey criteria. *Am J Manag Care* 2010;16(11):e281–e288
- 40 Hirsch AG, Stewart WF, Sundaresan AS, et al. Nasal and sinus symptoms and chronic rhinosinusitis in a population-based sample. *Allergy* 2017;72(02):274–281
- 41 Maeng DD, Snyder SR, Davis TW, Tomcavage JF. Impact of a complex care management model on cost and utilization among adolescents and young adults with special care and health needs. *Popul Health Manag* 2017;20(06):435–441
- 42 Pitblado J. Survey Data Analysis in Stata: Canadian Stata Users Group Meeting; 2009. Available at: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjs6\\_2-NzKPOAhXDwiYKHbMqCD4QFggcMAA&url=http%3A%2F%2Fwww.stata.com%2Fmeeting%2Fcanada09%2Fca09\\_pitblado\\_handout.pdf&usq=AFQjCNFHePahdy3vjwjndc5-asNndJdb3Q&bvm=bv.128617741,d.eWE&cad=rja](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjs6_2-NzKPOAhXDwiYKHbMqCD4QFggcMAA&url=http%3A%2F%2Fwww.stata.com%2Fmeeting%2Fcanada09%2Fca09_pitblado_handout.pdf&usq=AFQjCNFHePahdy3vjwjndc5-asNndJdb3Q&bvm=bv.128617741,d.eWE&cad=rja). Accessed August 2, 2016
- 43 Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189. Doi: 10.1136/bmj.i189
- 44 Chantala K, Blanchette D, Suchindran CM. Software to Compute Sampling Weights for Multilevel Analysis. Chapel Hill, NC: Carolina Population Center, UNC at Chapel Hill; 2011
- 45 Little RJ, Lewitzky S, Heeringa S, Lepkowski J, Kessler RC. Assessment of weighting methodology for the National Comorbidity Survey. *Am J Epidemiol* 1997;146(05):439–449
- 46 Dargatz DA, Hill GW. Analysis of survey data. *Prev Vet Med* 1996;28(04):225–237
- 47 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(03):837–845
- 48 Meng Y, Lou H, Wang C, Zhang L. Predictive significance of computed tomography in eosinophilic chronic rhinosinusitis with nasal polyps. *Int Forum Allergy Rhinol* 2016;6(08):812–819
- 49 Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003;29(07):1043–1051
- 50 Zlotnik A, Abaira V. A general-purpose nomogram generator for predictive logistic regression models. *Stata J* 2015;15(02):537–546
- 51 Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* 2008;26(08):1364–1370
- 52 Newton KM, Peissig PL, Kho AN, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. *J Am Med Inform Assoc* 2013;20(e1):e147–e154
- 53 2015 National Health Interview Survey (NHIS). Data National Center for Health Statistics, CDC; updated February 10, 2017. Available at: <https://www.cdc.gov/asthma/nhis/2015/table6-1.htm>. Accessed June 19, 2018
- 54 Smith K, Warholak T, Armstrong E, Leib M, Rehfeld R, Malone D. Evaluation of risk factors and health outcomes among persons with asthma. *J Asthma* 2009;46(03):234–237
- 55 Murphy KR, Meltzer EO, Blaiss MS, Nathan RA, Stoloff SW, Doherty DE. Asthma management and control in the United States: results of the 2009 Asthma Insight and Management survey. *Allergy Asthma Proc* 2012;33(01):54–64
- 56 Bahadori K, Doyle-Waters MM, Marra C, et al. Economic burden of asthma: a systematic review. *BMC Pulm Med* 2009;9:24
- 57 Okelo SO, Butz AM, Sharma R, et al. Interventions to modify health care provider adherence to asthma guidelines: a systematic review. *Pediatrics* 2013;132(03):517–534
- 58 Kharrazi H, Chi W, Chang HY, et al. Comparing population-based risk-stratification model performance using demographic, diagnosis and medication data extracted from outpatient electronic health records versus administrative claims. *Med Care* 2017;55(08):789–796
- 59 Rowan CG, Flory J, Gerhard T, et al. Agreement and validity of electronic health record prescribing data relative to pharmacy claims data: a validation study from a US electronic health record database. *Pharmacoepidemiol Drug Saf* 2017;26(08):963–972