



# New phenotyping of asthmatics using long-term followed measurement data

## Trajectory clustering of lung function data over 15 years



Yujin Park, Pharm.D<sup>1</sup>, Seng Chan You, MD, MS<sup>2</sup>, Youngsoo Lee, MD<sup>3</sup>, Hyun Young Lee, Ph.D<sup>4</sup>, Eun Young Lee, MS<sup>2</sup>, Hae-Sim Park, M.D, Ph.D<sup>3</sup>, Rae Woong Park, MD, Ph.D<sup>1,2,5</sup>

<sup>1</sup>Department of Biomedical Sciences, Ajou University School of Medicine, Suwon, Korea; <sup>2</sup>Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea; <sup>3</sup>Department of Allergy and Clinical Immunology, Ajou University School of Medicine, Suwon, Korea; <sup>4</sup>Department of Statistics, Clinical Trial Center, Ajou University School of Medicine, Suwon, Korea; <sup>5</sup>FEEDER-NET(Federated E-Health Big Data for Evidence Renovation Network), Suwon, South Korea

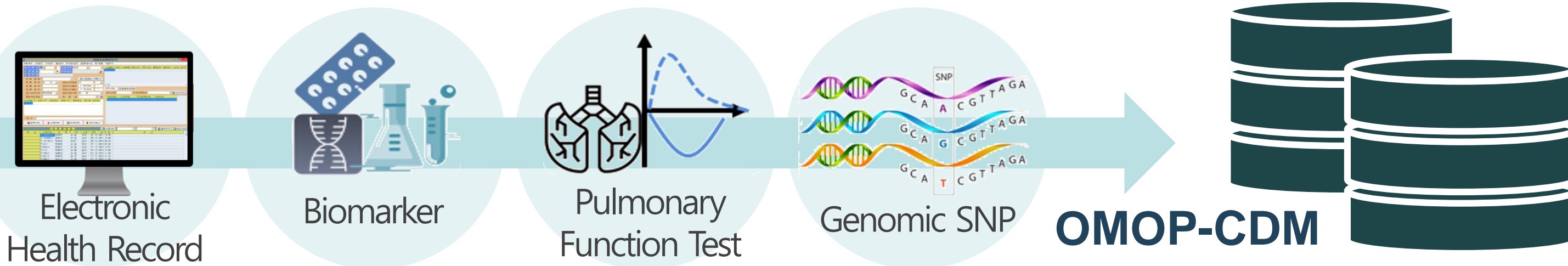
### Why "trajectory clustering" is needed?

- **Trajectory** means the course of measured variable over age or time
- Even though started at **same condition** and treated with **same treatment**, trajectory after the start point can be **different**.
- Recently, the need of new phenotype decided by long-term patterns were emerged
- **Latent class linear mixed model (LCMM)** is an innovative statistical method used to identify subgroups of participants with **heterogeneous trajectories**

### Methods

#### ICARUS Database

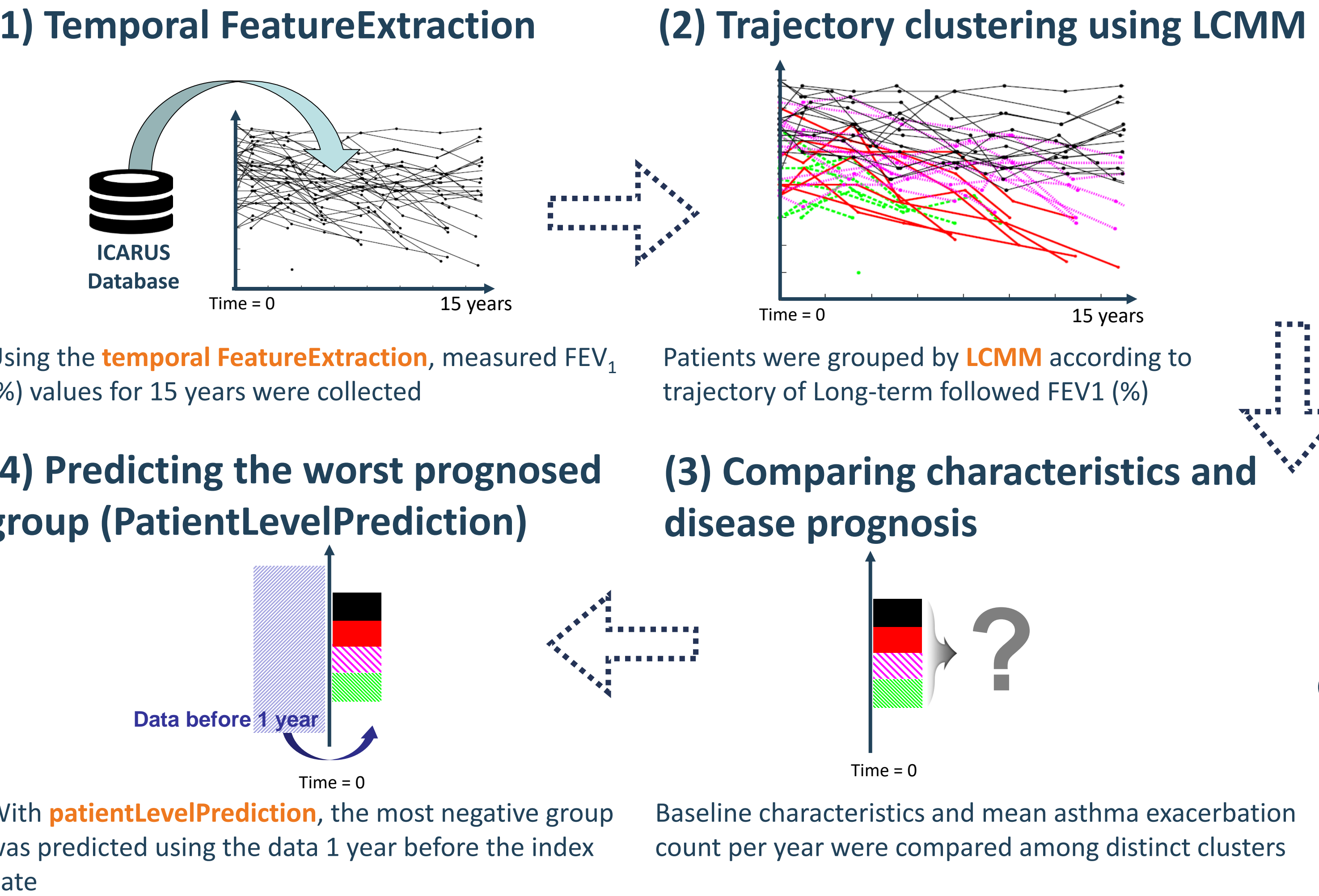
- **Immune/Inflammatory Disease Common Data Model Augmentation for Research Union System (ICARUS) database**
- patients who visited department of allergy and clinical immunology base on the EHR of Ajou University medical center from 1995 to 2017



### Subjects



### Process



### Results

#### Results of trajectory clustering

Severe asthmatics were clustered into 3 classes according to lung function using LCMM

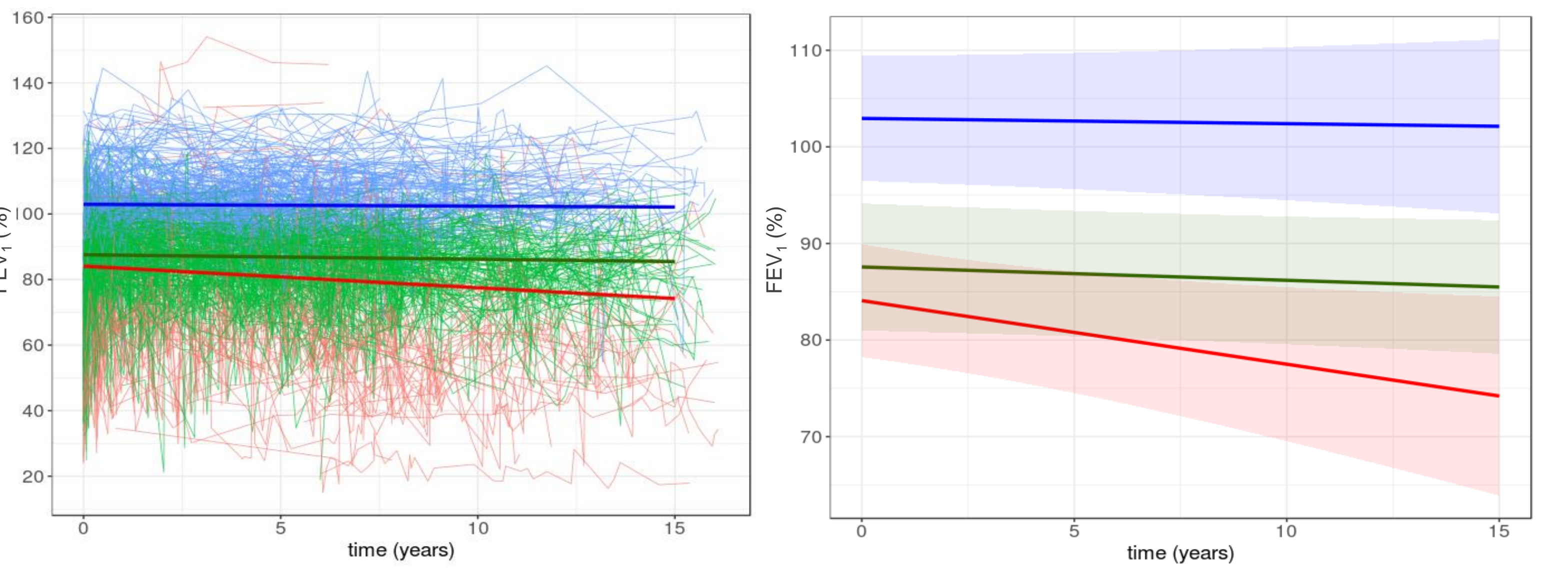


Figure 1-a. Observation individual lung function trajectories and three estimated representative trajectories. Figure 1-b. Three estimated representative trajectories. The shaded areas indicate estimated 95% confidential intervals

	Persistently high (N = 408, 34.75 %)	Persistently low (N = 618, 52.64 %)	Declining (N = 148, 12.61 %)
Intercept (95% CI)	102.947 (96.47, 109.42)	87.57 (80.98, 94.15)	84.08 (78.23, 89.93)
Slope (95% CI)	-0.05 (-0.38, 0.27)	-0.14 (-0.36, 0.09)	-0.66 (-1.27, -0.05)

#### Comparing disease prognosis

The declining group had the most negative prognosis

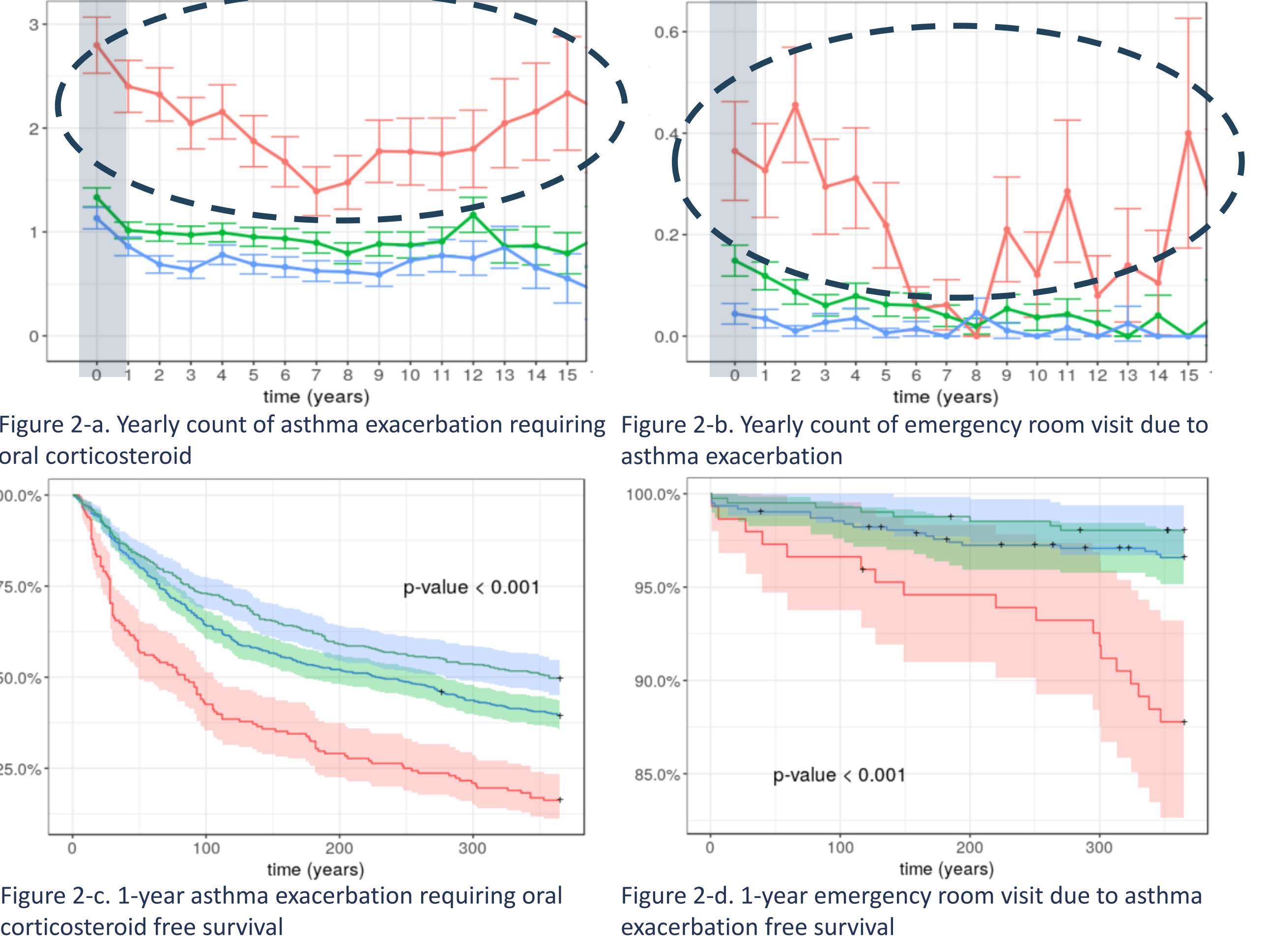


Figure 2-a. Yearly count of asthma exacerbation requiring oral corticosteroid. Figure 2-b. Yearly count of emergency room visit due to asthma exacerbation. Figure 2-c. 1-year asthma exacerbation requiring oral corticosteroid free survival. Figure 2-d. 1-year emergency room visit due to asthma exacerbation free survival

#### Comparing characteristics

Declining group had different baseline characteristics

	Persistently high (N = 408)	Persistently low (N = 618)	Declining (N = 148)	p value
Age, years	38.00 ± 15.03	38.82 ± 13.8	46.42 ± 13.55	<0.001
Follow-up duration, years	7.05 (4.15, 10.74)	6.59 (3.37, 10.72)	8.15 (4.11, 11.57)	0.105
Female, n (%)	252 (61.76)	337 (54.53)	89 (60.14)	0.059
BMI, kg/m <sup>2</sup> (N)	24.12 ± 13.17 (167)	25.71 ± 33.78 (373)	24.24 ± 4.81 (110)	0.705
Total IgE (KU/L)	213.5 (79.3, 512.8) (346)	241 (92, 477) (501)	199 (46.5, 535) (103)	0.688
Blood eosinophil (%)	3.7 (1.9, 6.8) (375)	4.2 (1.8, 7.5) (562)	4.4 (2.7, 8.5) (127)	0.034
Blood neutrophil (%)	55.4 (49.4, 61.9) (374)	54.9 (48.4, 62.7) (560)	55.7 (48.8, 62.1) (126)	0.764
Serum EDN (ng/mL)	46.5 (30.5, 63.3) (61)	50.1 (31.3, 83.1) (102)	65.8 (45.9, 100.2) (24)	0.009
Rhinosinusitis, N (%)	284 (69.6)	383 (62.0)	83 (56.1)	0.005
Urticaria/angioedema, N (%)	36 (8.8)	47 (7.6)	2 (1.4)	0.010
Anaphylaxis, N (%)	17 (4.2)	20 (3.2)	3 (2.0)	0.443
Hypertension, N (%)	13 (3.2)	28 (4.5)	9 (6.1)	0.291
Diabetes Mellitus, N (%)	7 (1.7)	12 (1.9)	3 (2.0)	0.956
Osteoporosis, N (%)	10 (2.5)	9 (1.5)	5 (3.4)	0.257
GERD, N (%)	12 (2.9)	21 (3.4)	6 (4.1)	0.802
Ischemic heart disease, N (%)	2 (0.5)	9 (1.5)	7 (4.7)	0.002

Mean ± Standard; Median (interquartile range); BMI, Body mass index; EDN, Eosinophil-derived neurotoxin; GERD, Gastroesophageal reflux disease.

#### Results of patient level prediction using PLP package

Declining group can be predicted using the data before 365 days

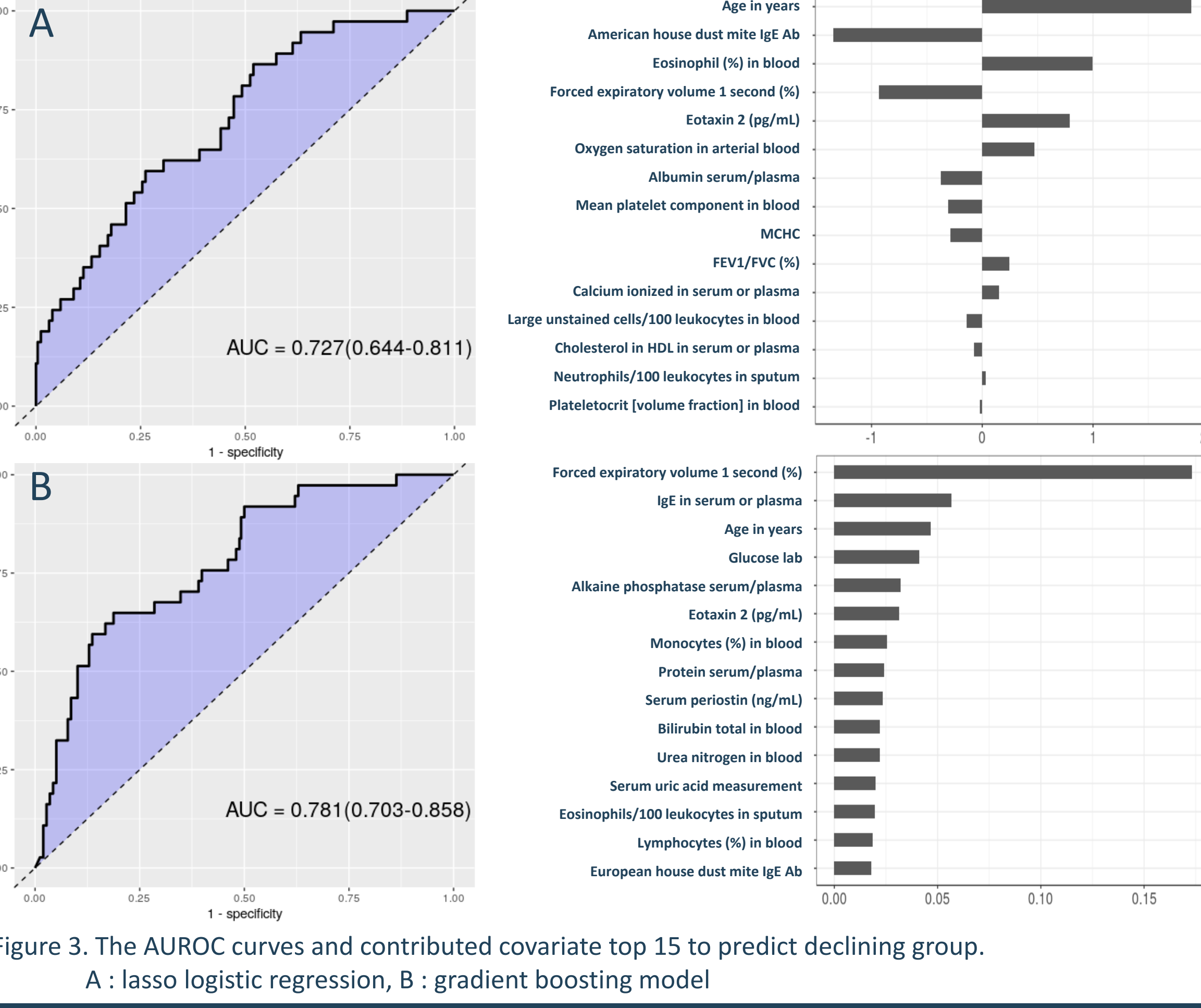


Figure 3. The AUROC curves and contributed covariate top 15 to predict declining group. A: lasso logistic regression, B: gradient boosting model

### Conclusions

1. Severe asthmatics can be classified according to long-term changes of lung function using the LCMM
2. A declining lung function had poor prognosis compared with other trajectories and can be predicted by data before 365 days using the PatientLevelPrediction
3. **Not only cross-sectional phenotyping, it is also important to define phenotype by long-term changes over time**