

# Clinical Significance of Asthma Clusters by Longitudinal Analysis in Korean Asthma Cohort

So Young Park<sup>1,9</sup>, Seunghye Baek<sup>2,9</sup>, Sujeong Kim<sup>1</sup>, Sun-young Yoon<sup>1</sup>, Hyouk-Soo Kwon<sup>1</sup>, Yoon-Seok Chang<sup>3</sup>, You Sook Cho<sup>1</sup>, An-Soo Jang<sup>4</sup>, Jung Won Park<sup>5</sup>, Dong-Ho Nahm<sup>6</sup>, Ho-Joo Yoon<sup>7</sup>, Sang-Heon Cho<sup>3</sup>, Young-Joo Cho<sup>8</sup>, ByoungWhui Choi<sup>9</sup>, Hee-Bom Moon<sup>1</sup>, Tae-Bum Kim<sup>1\*</sup>, for the COREA Study Group

**1** Division of Allergy and Clinical Immunology, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea, **2** Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, Seoul, Korea, **3** Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea, **4** Department of Internal Medicine, Soonchunhyang University Bucheon Hospital, Bucheon, Korea, **5** Department of Internal Medicine, College of Medicine, Yonsei University, Seoul, Korea, **6** Department of Internal Medicine, College of Medicine, Ajou University, Suwon, Korea, **7** Department of Internal Medicine, College of Medicine, Hanyang University, Seoul, Korea, **8** Department of Internal Medicine, College of Medicine, Ewha Womans University, Seoul, Korea, **9** Department of Internal Medicine, College of Medicine, Chung-Ang University, Seoul, Korea

## Abstract

**Background:** We have previously identified four distinct groups of asthma patients in Korean cohorts using cluster analysis: (A) smoking asthma, (B) severe obstructive asthma, (C) early-onset atopic asthma, and (D) late-onset mild asthma.

**Methods and Results:** A longitudinal analysis of each cluster in a Korean adult asthma cohort was performed to investigate the clinical significance of asthma clusters over 12 months. Cluster A showed relatively high asthma control test (ACT) scores but relatively low FEV<sub>1</sub> scores, despite a high percentage of systemic corticosteroid use. Cluster B had the lowest mean FEV<sub>1</sub>, ACT, and the quality of life questionnaire for adult Korean asthmatics (QLQAKA) scores throughout the year, even though the percentage of systemic corticosteroid use was the highest among the four clusters. Cluster C was ranked second in terms of FEV<sub>1</sub>, with the second lowest percentage of systemic corticosteroid use, and showed a marked improvement in subjective symptoms over time. Cluster D consistently showed the highest FEV<sub>1</sub>, the lowest systemic corticosteroid use, and had high ACT and QLQAKA scores.

**Conclusion:** Our asthma clusters had clinical significance with consistency among clusters over 12 months. These distinctive phenotypes may be useful in classifying asthma in real practice.

**Citation:** Park SY, Baek S, Kim S, Yoon S-y, Kwon H-S, et al. (2013) Clinical Significance of Asthma Clusters by Longitudinal Analysis in Korean Asthma Cohort. PLoS ONE 8(12): e83540. doi:10.1371/journal.pone.0083540

**Editor:** Tesfaye Mersha Baye, Cincinnati Children's Hospital Medical center, United States of America

**Received:** May 31, 2013; **Accepted:** November 5, 2013; **Published:** December 31, 2013

**Copyright:** © 2013 Park et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by a grant of the Korea Healthcare Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (grant no. A102065 & H13C0776). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: allergy@medimail.co.kr

These authors contributed equally to this work.

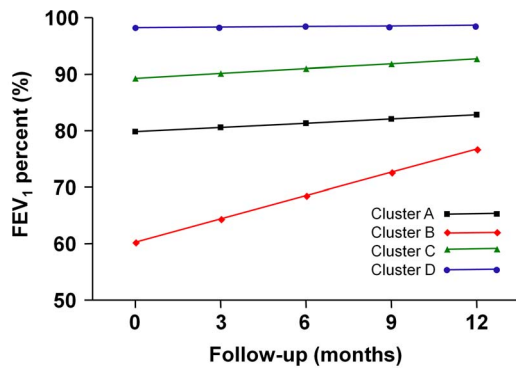
## Introduction

Asthma is a clinical syndrome of intermittent respiratory symptoms characterized by chronic airway inflammation, non-specific bronchial hyperresponsiveness, and reversible airway obstruction [1]. In the past, asthma management depended on the severity of the disease, which is determined according to patient lung function [forced expiratory volume in 1 second (FEV<sub>1</sub>)], daytime and nocturnal symptoms, and the frequency of rescue bronchodilator use [2]. Recent treatment guidelines emphasize the asthma control status, a convenient approach that can easily be applied clinically [1]. However, both the asthma control and severity have demonstrated critical drawbacks by failing to reflect the heterogeneous nature of asthma, which is determined by individually distinct pathophysiological backgrounds [3–5]. The concept of ‘heterogeneity’ has been used on a number of occasions in an attempt to create a novel classification

that integrates heterogeneous clinical characteristics with intrinsic disease severity, facilitating tailored treatment for asthma patients [5–8].

In our previous study [9], we used cluster analysis to identify four distinct groups of asthma in two large independent cohorts of Korean adult asthma patients: (A) smoking asthma, (B) severe obstructive asthma, (C) early-onset atopic asthma, and (D) late-onset mild asthma. Both cohorts were comprised of asthma patients at all levels of severity and had a reasonably large of number of participants [10].

Because the cluster analysis classifications in previous similar studies [5–8] lacked longitudinal evaluation of the population, it was difficult to confirm the clinical significance and to justify the real clinical relevance of the classifications. To overcome this limitation, our previous study revealed the change of FEV<sub>1</sub>



**Figure 1. FEV<sub>1</sub>% predicted values during the 12-month follow-up period in each cluster (adapted from Kim TB et al., Eur Respir J, 2013;41:1308–1314).**

doi:10.1371/journal.pone.0083540.g001

throughout a 12-month follow-up period in the four distinctive groups [10].

The purpose of our present study was to characterize and compare cluster-specific trends in clinical parameters of asthma control over 12 months of follow-up in a Korean asthma cohort.

## Methods

### Study population

This study was performed with 724 subjects from the Cohort for Reality and Evolution of Adult Asthma in Korea (COREA), who were the subjects of our previous cluster analysis [10]. As described in the previous report, patients, recruited by allergists or pulmonologists from 11 tertiary referral centers in Korea, were diagnosed with asthma by the presence of asthma symptoms and airway reversibility. The patients were seen at their hospitals regularly every 3 months and were managed uniformly according to the Global Initiative for Asthma (GINA) guidelines [1]. In every 3-month-visit, patients were supposed to check spirometry, report ACT scores and the investigator reviewed and recorded the medications taken in the previous 3 months. This included whether the patient had taken systemic corticosteroid for asthma exacerbation and whether asthma medication had been stepped up, down or maintained. Also, patients were asked to answer the questionnaire, QLQAKA, at cohort registration and every 1 year afterwards [11].

## Ethics

All study participants were fully informed of the study protocol and provided written, signed statements of informed consent. The protocol was approved by the Institutional Review Board and Ethics Committee of Asan Medical Center.

## Cluster analysis

Hierarchical cluster analysis using Ward's method was performed to estimate the number of likely clusters within the population and was followed by k-means cluster analysis as the principal clustering technique for grouping individuals into clusters [12]. Variables for modeling were selected based on their contribution to the characterization of the asthma phenotype, which were the FEV<sub>1</sub>, body mass index (BMI), age at onset, atopic status, smoking history, and history of hospital utilization due to exacerbation [10].

## Longitudinal outcome measures

We investigated the longitudinal trend of clinical status in terms of FEV<sub>1</sub>% predicted, asthma control test (ACT) score [13], percentage of systemic corticosteroid use (percentage of patients on a brief course of systemic corticosteroid for acute exacerbation at a given timepoint), and quality of life questionnaire for adult Korean asthmatics (QLQAKA) score [14] for each predefined cluster. Data at 3<sup>rd</sup>, 6<sup>th</sup>, 9<sup>th</sup> and 12<sup>th</sup> months were retrieved from the cohort and were retrospectively analyzed.

QLQAKA is the Korean modified version of the Juniper asthma quality of life questionnaire (AQLQ) [15]. In the QLQAKA, answers to each question are scored on a 5-point scale, with a score of 1 representing greatest impairment and a score of 5 representing no impairment. Items are weighted equally and are reported as a mean score for each domain (activity limitations, emotions, symptoms, and exposure to environmental stimuli) along with the overall mean score [14].

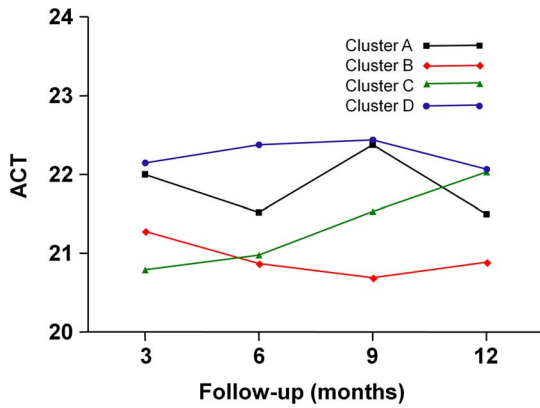
## Statistical analysis

We used the linear mixed effects models [16] for the FEV<sub>1</sub>% predicted, the ACT and QLQAKA scores, and the generalized estimating equations (GEE) [17] for the percentage of systemic corticosteroid use in order to estimate the outcome values and test the differences among clusters [18]. Both statistical methods are used to account for within-individual correlations since the data for each patient were repeatedly measured and assumed to be correlated. The linear mixed effect model included cluster effect, time, and their interaction as fixed effects, and subject and time as random effects. The analyses were applied to the raw data with no

**Table 1. FEV<sub>1</sub>% predicted values during the 12-month follow-up period in each cluster.**

Months	A		B		C		D	
	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)
0	57	79.85 (76.70–83.00)	109	60.31 (58.02–62.60)	159	89.33 (87.46–91.22)	155	98.22 (96.30–100.15)
3	48	80.60 (77.53–83.68)	82	64.43 (62.19–66.66)	125	90.17 (88.33–92.02)	114	98.34 (96.47–100.22)
6	39	81.35 (78.10–84.60)	77	68.55 (66.19–70.91)	102	91.02 (89.06–92.98)	110	98.46 (96.48–100.44)
9	41	82.10 (78.46–85.74)	71	72.67 (70.02–75.31)	106	91.86 (89.66–94.07)	97	98.58 (96.36–100.80)
12	37	82.85 (78.67–87.03)	64	76.78 (73.73–79.84)	94	92.71 (90.17–95.25)	95	98.70 (96.15–101.25)

doi:10.1371/journal.pone.0083540.t001



**Figure 2. ACT scores during the 12-month follow-up period in each cluster.**  
doi:10.1371/journal.pone.0083540.g002

imputation of missing data. To ensure the insensitivity of the results to missing data, analyses were also done after multiple imputations. Multiple imputations with a Markov chain Monte Carlo method were performed for each outcome using PROC MI in SAS, and 10 imputed data sets were then analyzed by using standard procedures for complete data. The results from the data sets were combined using PROC MIANALYZE in SAS. (See Supplementary Appendix) All statistical analyses were performed with SAS software (version 9.2; SAS Institute, Cary, NC, USA).

**Results**

**FEV<sub>1</sub>**

Because the patients whose FEV<sub>1</sub> was only checked at baseline (with no follow-up data from the subsequent 12-month period) were excluded, the total number of subjects included in the analysis was 480. The distribution of the subjects at baseline for clusters A, B, C, and D was 57, 109, 159 and 155, respectively.

Although we previously published the change of the FEV<sub>1</sub> during a 12-month period in four clusters [10], a more in-depth analysis has been performed here. The FEV<sub>1</sub> is consistently lowest in cluster B (the severe obstructive asthma group) and highest in cluster D (the late-onset mild asthma group) (Figure 1 and Table 1). Cluster A, the smoking asthma group and cluster C, the early-onset atopic asthma group ranked between cluster B and D in the order of low FEV<sub>1</sub>. The differences between the four clusters throughout the 12-month follow-up period were statistically significant ( $P < 0.0001$ ) [10]. The improvement in the FEV<sub>1</sub>%

predicted in cluster B over time was statistically significant ( $P < 0.0001$ ), whereas increases in the FEV<sub>1</sub> in other clusters were not significant.

In order to check if improvement of FEV<sub>1</sub> seen in cluster B represents loss of data from patients with a low baseline FEV<sub>1</sub>, we compared the mean values of last recorded FEV<sub>1</sub> between patients who revisited the clinic and those who did not at each follow up point, however, we found no significant differences of the last recorded FEV<sub>1</sub> at 3, 6, 9 months and interestingly, at 12 month, the mean value of last recorded FEV<sub>1</sub> in the patients who did not revisit clinic was significantly higher than those who did.

**ACT score**

The ACT score was measured from the 3<sup>rd</sup> month (Figure 2 and Table 2). Patients who did not fill in the ACT score were excluded from the study, leaving 459 patients. The distribution of patient numbers for clusters A, B, C, and D was 54, 102, 156 and 147, respectively.

All mean ACT scores were greater than 20, even in the most severe group (cluster B). There was a tendency of the patients in cluster D to report higher ACT scores than patients in cluster A, B and C, with significant differences from cluster B and C ( $P = 0.001$  and  $0.007$ , respectively). Patients in cluster B reported the lowest ACT scores throughout the whole follow-up period, with the exception of the 3<sup>rd</sup> month. Cluster A showed relatively high ACT scores compared with its low FEV<sub>1</sub>. A significant increase in ACT score was found in cluster C ( $P = 0.001$ ).

**Percentage of systemic corticosteroid use**

The number of patients used for longitudinal analysis of systemic corticosteroid use was 57, 109, 162, and 157 for clusters A, B, C and D, respectively. Patients for whom we had no information on systemic corticosteroid use were excluded from analysis. Patients in cluster B tended to more frequently use systemic corticosteroids than patients in other clusters (Figure 3 and Table 3). In particular, the use of systemic corticosteroids in cluster B was significantly higher than in clusters C and D over the entire period and at each visit ( $P = 0.0005$  and  $P < 0.0001$ , respectively). The percentage of systemic corticosteroid use in each cluster at each visit displayed a fluctuating pattern.

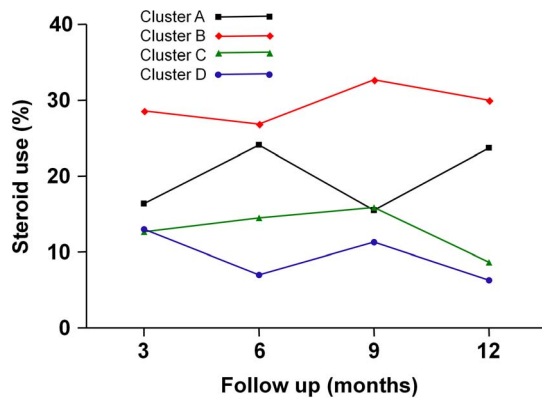
**QLQAKA score**

We measured QLQAKA scores twice, at baseline and at the 12<sup>th</sup> month. Lower QLQAKA scores were consistently seen in cluster B than in other clusters (Figure 4 and Table 4). The differences between cluster B and the other clusters were statistically significant at baseline and the 12<sup>th</sup> month, though no

**Table 2. ACT scores during the 12-month follow-up period in each cluster.**

Months	A		B		C		D	
	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)
3	44	22.00 (21.05–22.96)	58	21.28 (20.49–22.07)	99	20.79 (20.18–21.41)	92	22.15 (21.51–22.78)
6	37	21.52 (20.44–22.61)	63	20.87 (20.05–21.69)	87	20.98 (20.30–21.67)	92	22.38 (21.70–23.06)
9	34	22.38 (21.35–23.41)	62	20.69 (19.93–21.45)	98	21.53 (20.92–22.13)	92	22.44 (21.82–23.07)
12	41	21.50 (20.53–22.46)	68	20.89 (20.15–21.63)	99	22.03 (21.42–22.64)	95	22.07 (21.45–22.69)

doi:10.1371/journal.pone.0083540.t002



**Figure 3. Percentage of use of systemic corticosteroids during the 12-month follow-up period in each cluster.**  
doi:10.1371/journal.pone.0083540.g003

significant differences between cluster A and D were seen at the 12<sup>th</sup> month. The scores from baseline to the 12<sup>th</sup> month significantly increased in clusters B, C, and D ( $P=0.001$ ,  $P<0.0001$ , and  $P<0.003$ , respectively).

#### Analyses after multiple imputations

In order to ensure the insensitivity of the results to missing data, we repeated the analyses after multiple imputations (10 times). Although there were slight differences in the absolute numbers in the results, the trends were almost identical to the original outcomes for the FEV<sub>1</sub>, ACT, and QLQAKA scores. For the percentage of corticosteroid use, the overall trend was similar, though not identical (Figures S1, S2, S3, S4 and Tables S1, S2, S3, S4).

#### Discussion

As a follow-up to our previous report, we analyzed longitudinal outcomes in our Korean cohort to verify the validity and clinical significance of our clustering. We proposed a new differentiated classification model with easily accessible clinical variables on the basis of our previous asthma clustering [10]. However, we needed to confirm the usefulness and clinical application of our clinical phenotype classifications by observing their longitudinal changes, which has not yet been shown in other studies from Europe and the USA [5,6]. To our knowledge, this is the first assessment of the longitudinal outcomes of asthma phenotypes by cluster analysis.

We analyzed the four variables that we considered to be the most clinically important. The FEV<sub>1</sub> is an objective measure that

reflects the severity and control status of asthma as well as the degree of airway obstruction. The use of systemic corticosteroids is also an important factor defining asthma control status because it is prescribed when patients are considered to be uncontrolled or only partly controlled [1]. In addition, we evaluated patterns in the ACT and QLQAKA scores as subjective means to assess asthma control status.

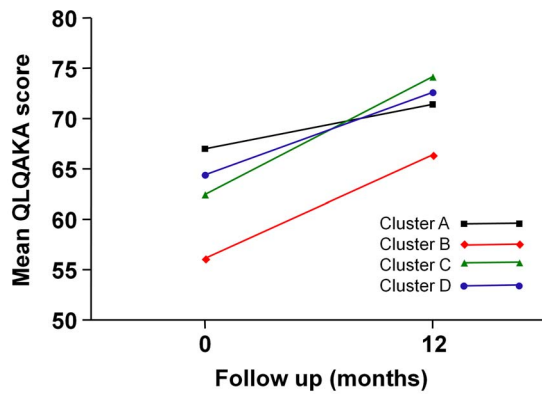
Because we observed changes in the values of each variable over time, we were able to outline the clinical course of each individual cluster. Cluster A (the smoking asthma group) had the second lowest mean FEV<sub>1</sub>% predicted after cluster B (the severe obstructive asthma group). Despite substantial airway obstruction, the ACT scores of cluster A were, after those of cluster D (the late-onset mild asthma group), the second highest. From this speculation, we could hypothesize that the smoking patients might be underestimating the seriousness of their asthma symptoms, believing the symptoms to be related with smoking, or they might have blunted perception of airway obstruction. In fact, it is supported by Kleis et al who have shown smoking asthmatics have reduced dyspnea perception during a nonspecific provocative challenge with methacholine [19]. Moreover, patients in cluster A show high degree of blood eosinophil count compared with other clusters (Table S5) which is a supporting evidence for dissociation between symptom and severe exacerbations in eosinophilic asthma with high degree of airway inflammation as asserted by some researchers [20]. All patients in cluster A were diagnosed as asthma, not COPD, by allergists and pulmonologists which was assured with methacholine bronchial provocation test. Another point that was observed from longitudinal analysis was that, even with regular inhaled corticosteroid treatment, a statistically significant improvement in the FEV<sub>1</sub> was not evident. This may be due to differences in the underlying pathophysiology of smoking asthma, attributable to the deleterious effect of cigarette smoking on asthma [21]. Unfortunately, this has limited clinical value since there is no data available regarding whether subjects have continued, reduced or quit smoking during the follow-up period.

The patients in cluster B showed the lowest mean FEV<sub>1</sub>% predicted, the highest percentage of systemic corticosteroid use, and the lowest ACT and QLQAKA scores throughout the 12-month follow-up period. Considering that the mean FEV<sub>1</sub> consistently increased from 60.31% at baseline to 76.78% at the 12<sup>th</sup> month with a marked improvement in QLQAKA scores, the patients in this cluster appear to respond to asthma treatment very well. In fact, they showed the largest percent increase in the FEV<sub>1</sub> after using short-acting bronchodilators in our previous study [10]. This implicates that we can expect increase in the FEV<sub>1</sub> in the patients of this cluster if they are treated with the optimal asthma medications. However, the final mean FEV<sub>1</sub>% predicted of the patients in this cluster did not reach 80%, implying that they

**Table 3. Percentage of use of systemic corticosteroids during the 12-month follow-up period in each cluster.**

Months	A		B		C		D	
	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)
3	52	0.164 (0.086–0.292)	84	0.286 (0.204–0.386)	130	0.127 (0.081–0.194)	124	0.130 (0.081–0.201)
6	42	0.241 (0.140–0.382)	79	0.269 (0.187–0.372)	108	0.145 (0.091–0.222)	115	0.070 (0.036–0.132)
9	43	0.155 (0.075–0.295)	74	0.327 (0.237–0.431)	112	0.159 (0.101–0.238)	105	0.113 (0.068–0.183)
12	41	0.237 (0.136–0.383)	68	0.300 (0.210–0.408)	101	0.087 (0.045–0.161)	102	0.063 (0.029–0.129)

doi:10.1371/journal.pone.0083540.t003



**Figure 4. QLQAKA scores during the 12-month follow-up period in each cluster.**

doi:10.1371/journal.pone.0083540.g004

may be included in the severe refractory asthma group. Therefore, optimal management of these patients requires close monitoring from physicians. Furthermore, researchers should use this group to study the pathogenesis of severe refractory asthma.

Cluster C (the early-onset atopic asthma group) was, after cluster D, ranked second in terms of mean FEV<sub>1</sub>% predicted and had the second lowest percentage of systemic corticosteroid use. The mean FEV<sub>1</sub> slightly, but not significantly, increased during the 12-month follow-up period. Interestingly, the ACT scores steadily increased from the 3<sup>rd</sup> month to the last follow-up, meaning that the patients with early-onset atopic asthma may have a benign disease course with marked improvement of subjective symptoms. In contrast, Moore et al. [6] demonstrated an atopic subgroup of patients presenting with a severe course of disease. However, there were no severe asthma patients with a low baseline FEV<sub>1</sub> and a poor prognosis in this cluster in our study.

The patients in cluster D consistently showed the highest FEV<sub>1</sub>% predicted, the lowest percentage of systemic corticosteroid use, and high ACT and QLQAKA scores. The longitudinal outcome analysis demonstrated its benign course throughout the 12-month period, which is comparable to the ‘benign asthma’ group in the United Kingdom study [5]. One unique finding was that this mild asthma group of patients was comprised of those with a later onset of disease and had a female predominance. A common feature of female predominant, older subgroups from other studies was that the patients were obese and non-eosinophilic [5,6]. However, our patients were not obese compared to patients from other clusters and they did not show significantly different total serum IgE levels or blood eosinophil counts [10]. This is indicative of the distinctness of the Korean asthma population, consisting of a subgroup of older, female predominant, asthma patients showing a benign course of disease.

Another interesting finding from this study was that all ACT scores were higher than 20. Usually, in practice we assess patients to be fairly well controlled if the scores are equal to or higher than 20 [22]. Most patients recruited in the cohort appeared to be well managed by attending physicians, because most showed reasonably good lung function during the follow-up period. However, because even those patients of cluster B with relatively poor lung function reported ACT scores greater than 20, physicians should recognize the fact that patients may underestimate their status when scoring ACT.

We didn’t provide information on medication use in this article. The patients were managed by the GINA guidelines, and accordingly the medications were stepped up, down or maintained based on their asthma control status. Therefore, the level of controller medication may differ depending on the asthma severity even for patients in the same cluster, and it may also change in one patient at different follow up points. We thought that change in medication use over time does not necessarily charge bias in the overall picture of the clinical course, since all patients were managed under one big principle. Indeed, when we analyzed the patients’ self-reported compliance to their prescriptions by a visual-analogue scale at every 3-month-visit, we found that all mean scores for every cluster was above 80%, which is fairly good. Therefore, we believe that difference in medications did not influence our results since most patients were compliant to their prescriptions that adhered to the GINA guidelines.

A major limitation of the present study is that we could not complete the 12-month database with all subjects who participated in the initial cluster analysis. There was loss to follow-up and missing data for all variables. In order to overcome this problem, we used the linear mixed effect model and GEE, which offer a general framework from which to develop longitudinal analyses under the missing at random (MAR) assumption [23], so these methods are more robust to potential bias from missing data than LOCF (Last Observation Carried Forward). These methods provided us with unbiased estimates of cluster and time effects assuming that the data loss occurred randomly. Simultaneously, they defined the longitudinal trend of each cluster and make comparisons between them [18]. To confirm the sensitivity of the missing data, we performed multiple(10) imputationsto fill in the missing variables and combined the 10 results to infer, obtaining similar results.

In conclusion, we confirmed the clinical significance of asthma clusters by longitudinally analyzing the data of our cohort. Our asthma clusters may not be immediately applied to clinical practice as further studies are needed to precisely characterize each cluster. An algorithm that can help allocate each new asthma patient into an appropriate cluster should be developed and it is hoped that cluster-specific pathogenesis could be revealed in the near future to aid the development of novel therapeutic strategies.

**Table 4. QLQAKA scores during the 12-month follow-up period in each cluster.**

Months	A		B		C		D	
	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)	n	Pred. Mean (95% CI)
0	14	67.02 (60.79–73.26)	33	56.14 (52.07–60.20)	42	62.45 (58.85–66.04)	32	64.41 (60.30–68.52)
12	17	71.44 (65.77–77.12)	28	66.42 (62.02–70.81)	45	74.20 (70.73–77.68)	39	72.64 (68.90–76.38)

doi:10.1371/journal.pone.0083540.t004

## Supporting Information

**Figure S1** FEV<sub>1</sub>% predicted values during the 12-month follow-up period in each cluster after multiple imputations. (DOCX)

**Figure S2** ACT scores during the 12-month follow-up period in each cluster after multiple imputations. (DOCX)

**Figure S3** Percentage of use of systemic corticosteroids during the 12-month follow-up period in each cluster after multiple imputations. (DOCX)

**Figure S4** QLQAKA scores during the 12-month follow-up period in each cluster after multiple imputations. (DOCX)

**Table S1** FEV<sub>1</sub>% predicted values during the 12-month follow-up period in each cluster after multiple imputations. (DOCX)

**Table S2** ACT scores during the 12-month follow-up period in each cluster after multiple imputations.

(DOCX)

**Table S3** Percentage use of systemic corticosteroids during the 12-month follow-up period in each cluster after multiple imputations.

(DOCX)

**Table S4** QLQAKA scores during the 12-month follow-up period in each cluster after multiple imputations.

(DOCX)

**Table S5** Characteristics of the four clusters of COREA patients.

(DOCX)

## Author Contributions

Conceived and designed the experiments: TBK. Performed the experiments: SYP SB TBK. Analyzed the data: SYP SB SK SY Y HSK Y. Chang Y. Cho ASJ JWP DHN HJY SHC YJC BWC HBM TBK. Contributed reagents/materials/analysis tools: SYP SB SK SY Y HSK YS. Chang YS. Cho ASJ JWP DHN HJY SHC YJC BWC HBM TBK. Wrote the paper: SYP TBK.

## References

- Bateman ED, Hurd SS, Barnes PJ, Bousquet J, Drazen JM, et al. (2008) Global strategy for asthma management and prevention: GINA executive summary. *Eur Respir J* 31: 143–178.
- Lenfant C, Clark TJH, Cagnani CB, Bousquet J, et al. (2002) Global Strategy for Asthma Management and Prevention. Global Initiative for Asthma (GINA).
- Miller MK, Johnson C, Miller DP, Deniz Y, Bleecker ER, et al. (2005) Severity assessment in asthma: An evolving concept. *J Allergy Clin Immunol* 116: 990–995.
- Wenzel SE (2006) Asthma: defining of the persistent adult phenotypes. *Lancet* 368: 804–813.
- Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, et al. (2008) Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 178: 218–224.
- Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, et al. (2010) Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 181: 315–323.
- Green RH, Brightling CE, Bradding P (2007) The reclassification of asthma based on subphenotypes. *Curr Opin Allergy Clin Immunol* 7: 43–50.
- Anderson GP (2008) Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet* 372: 1107–1119.
- Kim TB, Jang AS, Kwon HS, Park JS, Chang YS, et al. (2013) Identification of asthma clusters in two independent Korean adult asthma cohorts. *Eur Respir J* 41: 1308–1314.
- Kim TB, Jang AS, Kwon HS, Park JS, Chang YS, et al. (2012) Identification of asthma clusters in two independent Korean adult asthma cohorts. *Eur Respir J*.
- Kim TB, Park CS, Bae YJ, Cho YS, Moon HB (2009) Factors associated with severity and exacerbation of asthma: a baseline analysis of the cohort for reality and evolution of adult asthma in Korea (COREA). *Ann Allergy Asthma Immunol* 103: 311–317.
- Ball GH, Hall DJ (1967) A clustering technique for summarizing multivariate data. *Behav Sci* 12: 153–155.
- Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, et al. (2004) Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 113: 59–65.
- Kwon HS, Lee SH, Yang MS, Lee SM, Chang YS, et al. (2008) Correlation between the Korean version of Asthma Control Test and health-related quality of life in adult asthmatics. *J Korean Med Sci* 23: 621–627.
- Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, et al. (1992) Evaluation of impairment of health related quality of life in asthma: development of a questionnaire for use in clinical trials. *Thorax* 47: 76–83.
- Laird NM, Ware JH (1982) Random-effects models for longitudinal data. *Biometrics* 38: 963–974.
- Zeger SL, Liang KY (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42: 121–130.
- Ibrahim JG, Molenberghs G (2009) Missing data methods in longitudinal studies: a review. *Test (Madr)* 18: 1–43.
- Kleis S, Chanez P, Delvaux M, Louis R (2007) Perception of dyspnea in mild smoking asthmatics. *Respir Med* 101: 1426–1430.
- Veen JC, Smits HH, Ravensberg AJ, Hiemstra PS, Sterk PJ, et al. (1998) Impaired perception of dyspnea in patients with severe asthma. Relation to sputum eosinophils. *Am J Respir Crit Care Med* 158: 1134–1141.
- Polosa R, Thomson NC (2012) Smoking and asthma: dangerous liaisons. *Eur Respir J*.
- Lai CK, Ko FW, Bhome A, DE Guia TS, Wong GW, et al. (2011) Relationship between asthma control status, the Asthma Control Test<sup>TM</sup> and urgent health-care utilization in Asia. *Respirology* 16: 688–697.
- Cnaan A, Laird NM, Slasor P (1997) Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med* 16: 2349–2380.