# Introduction

Asthma is a chronic airway disease, usually associated with airway inflammation, characterized by the conjunction of respiratory symptoms such as dyspnea, chest tightness and wheezing together with excessive airway caliber fluctuation 1. The inflammatory process frequently features an eosinophilic inflammation, often combined with raised FeNO levels and IgE production directed against aeroallergens, defining the T2 high phenotype 2. However, asthma may be present without airway eosinophilic inflammation 3;4. While much work has been done on eosinophilic asthma, non eosinophilic phenotype has been much less studied although chronic infection, pollutant exposure and airway smooth muscle dysfunction are thought to contribute 5;6.

Clustering has become a popular method to identify phenotypes among a large set of asthmatic patients 7;8 and we have recently reported two clusters among a large cohort of eosinophilic asthmatics 9. However, to the best of our knowledge, clustering method has not yet been applied to a cohort of non eosinophilic asthmatics. It is therefore appropriate to further study non-eosinophilic asthma and to investigate how these patients can be grouped into multiple homogenous clusters.

 Here we performed an unsupervised cluster analysis on selected asthmatics displaying sputum eosinophilia < 3% leveraging a large asthma clinic database from a secondary care centre. We first performed clustering analysis on the whole cohort of non eosinophilic asthmatics and then repeated the analysis on the group of corticosteroid naïve patients and the group of patients receiving high doses of inhaled corticosteroids.

# Methods

***Study design and patient characteristics***

A retrospective cross-sectional study was conducted on 588 non-eosinophilic patients recruited between 2011 and 2020 from the asthma clinic at Liege University Hospital and displaying a sputum eosinophil counts <3%. Patients were selected from our general database that contained 1014 patients with successful sputum induction at their first visit in the asthma clinic. Each patient underwent detailed history taking including the age at which symptoms started, smoking habit, treatment characteristics and the number of exacerbations in the previous 12 months. Exacerbations were defined by a course of OCS for at least 3 days or administration of an intramuscular corticosteroid or admission to emergency department for asthma destabilization treated with injection of corticosteroids. All asthmatics had asthma diagnosis based on symptoms (wheezing, breathlessness, chest tightness, cough) and at least one of the following lung function features: FEV1 increase of ≥ 12% and 200 ml after inhalation of 400 µg salbutamol or a provocative concentration of methacholine causing a 20% fall in FEV1 (PC20M) less than 16 mg/ml 10. Diffusion capacity to carbon monoxide (DLCO) was measured by the single breath method 11 and lung volumes by body plethysmography 12 (Geratherm, Germany). The atopic status was measured by positive skin prick tests or a positive specific IgE reaction (> 0.35 kU/l; Phadia; Groot-Bijgaarden, Belgium) to common aeroallergens in Belgium including dermatophagoides pteronyssinus and dermatophagoides farinae, grass and birch pollens, dog and cat danders, alternaria cladosporium and aspergillus. FENO measurements were performed at 50 ml/s of flow rate (NIOX, Aerocrine, Sweden). The Asthma Quality of Life Questionnaire (AQLQ), Asthma Control Questionnaire (ACQ) and Asthma Control Test (ACT) were administered to assess asthma related quality of life and disease control. 13-15. Asthma was considered as being uncontrolled if ACT was < 20 and/or ACQ > 1,5 14;15. Sputum was induced and processed as previously reported 16. None of the patients included in the study were treated with biologics at the time the visit. The study was approved by the Liege University Hospital ethics committee.

***Statistical analyses***

We performed cluster analysis on three groups of asthmatics: i) all non eosinophilic asthmatics, ii) non eosinophilic patients without ICS and iii) non eosinophilic patients with high doses ICS that is >1000 µg/d equivalent beclomethasone. Qualitative variables were presented as count and percentage while quantitative variables were expressed as median and interquartile range (P25 - P75).

The global clustering framework is detailed in eTable 1. The descriptive tables found in supplementary documents also included the number of missing values and their percentages. The handling of missing values was based on multiple imputation in order to account for the uncertainty of missing values. The multiple imputation procedure generates a set of *m* (the most common value for *m* is 100) plausible values for each missing value.

The method of hierarchical clustering on principal components (HCPC) was applied to each imputed dataset to combine the two steps of variable reduction and cluster analysis. The first step was to perform a factor analysis of mixed data (FAMD) method on *m* imputed datasets to reduce the complexity of multidimensional data. Components that explained at least 90% of the variance in the data were selected. The contribution of each variable to the new components of FAMD was also calculated. In the next step, hierarchical clustering was applied using Ward's criterion on the principal components derived from the first step in order to determine the appropriate number of clusters. Then, K-means was applied to classify the patients into the detected clusters. For each imputed dataset, this process has been applied individually, thereby generating 100 results for clustering analysis. Therefore, in the final step, a consensus clustering method based on the mixture multivariate multinomial model (4M) method was considered to assign each patient to a cluster.

In order to compare the characteristics of patients falling in each cluster a Mann-Whitney nonparametric test and a Chi-square test were applied for quantitative variables and qualitative variables, respectively.

 The analyses were conducted using R software and several well-known packages, including MICE (multiple imputation by chained equations), FactoExtra, and FactoMineR. Based on the author's R function, 4M consensus clustering was implemented. Statistical significance was determined at P values <0.05.

This study was approved by the Liège University Hospital ethics committee. Signed informed consent was obtained from patients as soon as they entered the asthma clinic. They agreed to allow their clinical data and the health outcomes they reported in the routine setting to be used for research purposes.

# Results

 Whole cohort demographic and clinical features (N=588) are given in eTable 2. Patients were mainly females (62%) with a median age of 50 years. The median asthma duration before the visit was 5 years. There were 51% of patients who had never smoked while 25% were ex-smokers and 24% were current smokers. Lung function was generally well preserved with spirometric indices within the normal range in 71% of patients. As compared to inhaled corticosteroid (ICS) naïve patients, those receiving maintenance high dose ICS had a longer asthma duration, a greater prevalence of atopy, a higher treatment burden beyond ICS/LABA, lower baseline % predicted FEV1, greater bronchial hyperresponsiveness and worse asthma control and quality of life (eTable2). With respect to blood analysis, those receiving high dose ICS had greater total serum IgE and a lower serum morning cortisol (eTable2).

***Clusters in the whole cohort***

Two subgroups were identified from the cluster analysis. Cluster 1 comprised the large majority of patients (n=417, 71%) while cluster 2 accounted for 29% of the cohort (n=171) (Table 1). Patients in cluster 1 had a median age of 53 years, a low prevalence of atopic status (24%), a frequent smoking history (53% overall with 25% current smokers and 28% ex-smokers), a late disease onset (Figure 1), a low level of treatment (55% without ICS), preserved lung function with median post bronchodilation FEV1 of 92% predicted but uncontrolled asthma for the majority of them (median ACT and ACQ reaching 16 and 1.7 respectively). Twenty-four percent of them reported at least one exacerbation in the previous year. Patients from this cluster displayed dominant airway and systemic neutrophilic inflammation. Cluster 2 included younger patients (median age 39 years) with an early disease onset and being almost exclusively atopic (99%) and reporting infrequent smoking history (35% overall with 16% current smokers and 19% ex-smokers) (Figure 1). They had a higher treatment burden with 64% of patients receiving maintenance ICS (median ICS dose 800 µg/d equivalent beclomethasone) and approximately one third receiving LTRA and H1 antagonist. They had better asthma control (median ACT and ACQ reaching 18 and 1.3 respectively) than patients in cluster 1. They had preserved airflow with a median post bronchodilation of FEV1 of 95% predicted and paucigranulocytic asthma for the majority of them combining sputum eosinophils < 3% and sputum neutrophils < 76% as previously defined 17. Thirty-one % of them reported at least one exacerbation in the 12 months preceding the visit. Asthma quality of life was significantly lower in cluster 1 than in cluster 2 with median AQLQ of 4.5 and 5.3 respectively (p<0.0001) (Table 1).

While spirometric indices were similar between the two clusters, diffusing capacity (DLCO) and transfer coefficient (KCO) were slightly altered and lower in patients from cluster 1 than in those from cluster 2. Bronchial hyperresponsiveness to methacholine was greater in cluster 2 than in cluster 1 (Figure 2).

Total serum IgE levels and sensitization rate towards common aeroallergens were much higher in cluster 2 than in cluster1. Likewise, FeNO levels were significantly higher in cluster 2 than in cluster 1. In contrast to IgE and FeNO, fibrinogen levels were significantly higher in cluster 1 than in cluster 2 (Figure 3).

***Clusters in corticosteroid naïve patients and in patients treated with high dose ICS***

 Clustering on ICS naive patients (n=280) and those treated with high doses ICS (n=135) also revealed two clusters in each of the two cohorts that mainly differed by age, disease onset and atopic status (eTables 3 and 4). As compared with cluster 1, median age was lower in cluster 2 in both corticosteroid naive and those treated with high dose ICS. Disease onset started earlier in life in cluster 1 than in cluster 2 in the two cohorts. The prevalence of atopy was strikingly different between the two clusters in both cohorts with a higher prevalence in cluster 2 (88% in ICS naïve and 89% in ICS high dose) than in cluster 1 (20% in ICS naïve and 29% in ICS high dose). Overall, asthma control was close to be adequate in both clusters of the steroid naive cohort while the patients were uncontrolled and reporting poor quality of life in both clusters from the cohort treated with high doses ICS. Airflow remained in the normal range in the large majority of patients of both clusters in the two cohorts with pre and post bronchodilation baseline FEV1 > 80% predicted and FEV1/FVC > 70%. In the corticosteroid naive cohort, the median sputum neutrophil count was higher in cluster 1 than in cluster 2 (75% vs 54%, p<0.0001 and 740 103/g vs 350 103/g, p=0.002)while in the cohort treated with high doses of ICS the neutrophilic trait of the cluster 1 was essentially reflected by elevated circulating blood neutrophils (60% vs 50%, p<0.0001 and 4696/µl vs 3200/µl, p<0.0001). Fibrinogen and CRP levels were higher in cluster 1 while IgE and FeNO were strikingly greater in cluster 2 both in corticosteroid naive patients and those receiving high dose ICS. In patients receiving high doses ICS, FeNO levels were even in the “high zone” (median FeNO 35 ppb) in patients from cluster 2 despite the absence of eosinophilic trait.

# Discussion

Non eosinophilic asthmatics defined by a sputum eosinophil count < 3% represent a large proportion of our asthmatic cohort with 588 patients over a total of 1014 patients with successful sputum induction at their first visit (58%). Our data provide evidence for two distinct clusters among non-eosinophilic asthmatics sharing similar spirometric values. The cluster 1 is the most numerous and includes mainly female patients the majority of whom had a late disease onset together with a frequent smoking history, yet not satisfying functional criteria for COPD (post bronchodilation FEV1/FVC < 70%), not receiving ICS for the majority of them and displaying neutrophilic airway and systemic inflammation. They had poor asthma control and quality of life. The cluster 2 features a dominantly male group with early disease onset consisting almost exclusively of atopic patients with a classical sensitization profile for our geographical area. They had better asthma control and quality of life and were more often receiving ICS/LABA, LTRA and H1 antagonist as compared to cluster 1.

The reality of asthma combined to cigarette smoking has been firmly established in epidemiological studies 18, yet there always remains a trend in the medical community to consider a smoker with chronic respiratory symptoms as a patient with chronic bronchitis and/or COPD rather than as an asthmatic. Our data reemphasize the importance of this group of patients in daily practice. Our patients were carefully diagnosed based on reversibility to salbutamol and/or on bronchial hyperresponsiveness to methacholine and their spirometric values were considered as being in the normal range with FEV1 % predicted above 80% and post FEV1/FVC ratio well above 70% for the large majority of them (86% of them). Therefore, our finding supports the concept of a smoking asthma phenotype distinct from COPD 19. As a consequence of greater smoking history, patients from this cluster had mildly impaired diffusion capacity, possibly linked to early emphysema 11, and rather low FeNO values as expected 20. The airway inflammatory profile was highly neutrophilic with a median value close to that seen in COPD, a finding in keeping with the demonstrated relationship between pack years and sputum neutrophils 21, even if airway dysbiosis may be another cause of increased sputum neutrophilia in cluster 1 22. This cluster had a poorer asthma control and quality of life compared to the cluster 2 but did not report a greater exacerbation rate the year prior to the visit with almost two thirds of the patients denying any course of OCS in the 12 preceding months in both clusters. This finding indicates that poor day to day asthma control may not necessarily result into greater exacerbation rate, especially in smoking patients.

Cluster 2 highlights the fact that T2 traits such as high FeNO levels and serum IgE may actually be present in non eosinophilic asthmatics. Atopic status is usually associated with an eosinophilic trait 23. The reason why the cluster 2 that had a high atopic prevalence remained non eosinophilic could due to several factors. First it may reflect the impact of ICS on airway eosinophilic inflammation. Up to two third of patients in cluster 2 were receiving ICS combined to LABA as maintenance treatment. It is highly likely that some of these patients were actually eosinophilic prior to starting their treatment with ICS, a class of drug known to be able to sharply decrease sputum eosinophils 24. Second, the lack of airway eosinophilia may also reflect a low allergen exposure in daily life as airway exposure to an allergen in a sensitized patient drives a long lasting eosinophilic infiltrate through mast cell activation 25 and allergen avoidance result in a decrease in sputum eosinophils 26. FeNO values in cluster 2 were often in the normal range 27 which also could support the absence of significant allergen exposure though it may also be the consequence of chronic treatment with ICS 28.

Whether the steroid naïve and uncontrolled patients from cluster 1 would benefit from ICS remains uncertain as both smoking 29 and high sputum neutrophil 30 makes the patient rather resistant to ICS. Our study also points out a significant group of steroid naive patients in cluster 2 who are paucigranulocytic and displaying normal FeNO. For this group of patients in whom asthma control was close to being adequate at the time of sputum sampling but who still may exacerbate, the best treatment strategy is still unclear as regular ICS was not found superior to placebo or LAMA 31. As needed use of ICS/LABA may be a better option than continuous ICS and as needed SABA in these patients 32.

By focusing on the patients treated with high doses of ICS and poor disease control we selected patients deemed to have severe asthma, or at least difficult to treat asthma, as we cannot guarantee adherence to treatment 33. Several national and international registries have shown that most of the severe asthmatics display signs of T2 high inflammation 34. A recent study including sputum analysis and investigating rizankizumab, a p19 IL-23 receptor antagonist, in severe asthma has provided similar finding as it turned out that the majority of recruited patients displayed sputum eosinophils above 3%, although no inflammatory inclusion criteria was mandatory to enter the study 35. Nonetheless, the NOVELTY study has challenged this view showing poor relationship between clinical severity of the disease and the magnitude of the blood eosinophil count 36. The patients in that study were, however, qualified as being asthmatics without any firm functional criteria and only a minority displayed significant reversibility at the bronchodilating test 37. Our study, which has included asthmatics based on either reversibility to salbutamol or on bronchial hyperresponsiveness to methacholine, points out the existence of severe non eosinophilic asthmatics. Smoking certainly contributes to poor asthma control and smoking cessation is undoubtedly the best treatment option for those smoking patients 18. Azithromycin may be of interest to reduce exacerbation in those with dysbiosis and increased amount of haemophilus influenza 38. A convincing relationship between FeNO and sputum eosinophils was found in previous studies 39;40. Therefore, it is worth noting that in those severe patients from cluster 2, FeNO levels were in the high zone (35 ppb) while airway eosinophilia was not prominent and blood eosinophil counts in the normal range 41. In these patients we could argue that the “magnet” is operating in the airways but the “bomb” (eosinophils) is lacking in the blood 42, which may perhaps explain the relative preservation of airflow as opposed to what was found in severe eosinophilic asthmatics 9. Interestingly, one third of the patients reported at least 2 exacerbations in the year prior to the visit. Why FeNO levels remained high despite ICS is unclear but suggest the inability of ICS to repress nitric oxide synthase 43, although poor adherence to ICS might not be discarded in some of them. Whether those non eosinophilic patients with high FeNO experiencing recurrent exacerbations may benefit from the addition of dupilumab, an anti-IL4 R antibody blocking IL-4 and IL-13, needs to be further investigated since the benefit of this treatment has been essentially observed in patients with blood eosinophils > 150/µl whereas the median blood eosinophils was 109/µl in our patients 44.

 ***Strength and Limitations***

The strength of our study is the application of new and powerful method of clustering on a large cohort of non-eosinophilic asthmatics precisely characterized in terms of lung function and airway and systemic inflammation. Our study has, however, several limitations. First, the patients considered as non eosinophilic based on a single sputum analysis might be intermittent eosinophilic asthmatic 45. Second, as this is a real life study, we are uncertain about the compliance of the patients to the treatment which obviously limits our interpretation of the disease severity. Third, we lack accurate data on comorbidities such as psychologic disorders, gastro-oesophagal reflux or chronic rhinosinusitis, that may alter asthma control and quality of life 46;47. Fourth, an external validation in a new cohort from a center using sputum analysis in clinical practice would enhance consistency of the results. Fifth, there was no longitudinal follow-up that could give us insight on the evolution lung function decline or exacerbation trend in our clusters, which are two major traits to be considered in asthma 48.

**Conclusion**

In conclusion non eosinophilic asthmatics are commonly encountered in daily practice and can be distributed in two distinct clusters. The first one, which contains the greater number of patients, features a late disease onset, a significant smoking history with airway and systemic neutrophilic inflammation while the other includes a majority of male atopic patients featuring paucigranulocytic asthma with moderately increased FeNO. Distinct clusters might require different pharmacological approach based on the inflammatory profile but this needs to be further investigated in large scale controlled trials after appropriate selection of patients