PLOS | ONE

RESEARCH ARTICLE

Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study

Emilie Burte[1,2]*, Jean Bousquet[1,2,3], Raphaëlle Varraso[1,2], Frédéric Gormand[4], Jocelyne Just[5,6], Régis Matran[7], Isabelle Pin[8,9,10,11], Valérie Siroux[8,9,10], Bénédicte Jacquemin[1,2,12☯], Rachel Nadif[1,2☯]

1 INSERM, U1168, VIMA: Aging and chronic diseases, Epidemiological and Public health approaches, F-94807, Villejuif, France, **2** Univ Versailles St-Quentin-en-Yvelines, UMR-S 1168, F-78180, Montigny le Bretonneux, France, **3** University hospital, Montpellier, France, **4** CHU de Lyon, Pneumology Department, Lyon, France, **5** Allergology Department, Centre de l'Asthme et des Allergies, Hôpital Armand-Trousseau (APHP), APHP, Paris, France, **6** Université Paris 6 Pierre et Marie Curie, Paris, France, **7** Univ Lille Nord de France, F-59000, Lille, France, **8** INSERM, IAB, Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, F-38000 Grenoble, France, **9** Univ. Grenoble Alpes, F-38000 Grenoble, France, **10** CHU de Grenoble, F-38000 Grenoble, France, **11** CHU de Grenoble, Pediatric Department, F-38000, Grenoble, France, **12** CREAL-Centre for Research in Environmental Epidemiology Parc de Recerca Biomèdica de Barcelona, Barcelona, Spain

☯ These authors contributed equally to this work.
* emilie.burte@inserm.fr

CrossMark
click for updates

OPEN ACCESS

Citation: Burte E, Bousquet J, Varraso R, Gormand F, Just J, Matran R, et al. (2015) Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study. PLoS ONE 10(8): e0136191. doi:10.1371/journal.pone.0136191

Editor: Zheng Liu, Tongji Medical College, CHINA

Received: March 18, 2015

Accepted: July 30, 2015

Published: August 26, 2015

Copyright: © 2015 Burte 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.

Data Availability Statement: Due to third party restrictions, EGEA data are not publicly available. Please see the following URL for more information: https://egeanet.vjf.inserm.fr/index.php/en/contacts-en. Interested researchers should contact egea.cohorte@inserm.fr with further questions regarding data access.

Funding: The authors have no support or funding to report.

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

Abstract

Background

The classification of rhinitis in adults is missing in epidemiological studies.

Objective

To identify phenotypes of adult rhinitis using an unsupervised approach (*data-driven*) compared with a classical *hypothesis-driven* approach.

Methods

983 adults of the French Epidemiological Study on the Genetics and Environment of Asthma (EGEA) were studied. Self-reported symptoms related to rhinitis such as nasal symptoms, hay fever, sinusitis, conjunctivitis, and sensitivities to different triggers (dust, animals, hay/flowers, cold air. . .) were used. Allergic sensitization was defined by at least one positive skin prick test to 12 aeroallergens. Mixture model was used to cluster participants, independently in those without (Asthma-, n = 582) and with asthma (Asthma+, n = 401).

Results

Three clusters were identified in both groups: 1) Cluster A (55% in Asthma-, and 22% in Asthma+) mainly characterized by the absence of nasal symptoms, 2) Cluster B (23% in Asthma-, 36% in Asthma+) mainly characterized by nasal symptoms all over the year,

sinusitis and a low prevalence of positive skin prick tests, and 3) Cluster C (22% in Asthma-, 42% in Asthma+) mainly characterized by a peak of nasal symptoms during spring, a high prevalence of positive skin prick tests and a high report of hay fever, allergic rhinitis and conjunctivitis. The highest rate of polysensitization (80%) was found in participants with comorbid asthma and allergic rhinitis.

Conclusion

This cluster analysis highlighted three clusters of rhinitis with similar characteristics than those known by clinicians but differing according to allergic sensitization, and this whatever the asthma status. These clusters could be easily rebuilt using a small number of variables.

Introduction

Rhinitis is a common respiratory disease worldwide and affects between 20 and 50% of the population depending on the country and on the definition [1–3]. Rhinitis is characterized by nasal congestion, rhinorrhea, itching and/or sneezing [1]. Classically, rhinitis can be divided into two major categories: allergic rhinitis (AR) and non-allergic rhinitis (NAR), with the need of allergic sensitization tests to distinguish between them [1]. Rhinitis is a complex disease, frequently associated with asthma, whatever the allergic sensitization [1] and phenotypes of rhinitis need to be explored.

In a systems biology study (the MeDALL approach, http://medall-fp7.eu/ [4]), classical and novel phenotypes of allergic rhinitis in children ascribed to *hypothesis-driven* and *data-driven* phenotypes were defined using epidemiologic questionnaires [5]. Even if symptoms of rhinitis are similar for children and adults, the disease may differ for comorbidities [6], and till now phenotypes of rhinitis are unexplored in adults.

Unsupervised learning methods (*data driven*) are useful as they allow studying a large data set without historical knowledge, and identifying distinct phenotypes not always detectable by classical approach. On the other hand, these methods can reinforce *hypothesis-driven* approaches and can thus confirm their validity. These methods have already been used with success to identify phenotypes of asthma [7], [8], chronic obstructive pulmonary diseases (COPD) [9], and other respiratory diseases [10]. To our knowledge, only one study has performed cluster analysis in 18 years old participants, all having current rhinitis [11].

The French Epidemiological study of Genetics and Environment of Asthma, bronchial hyperresponsiveness and atopy (EGEA)) is a case-control cohort on asthma. Participants of this study had a very good phenotypic characterization of respiratory health, including allergic sensitization and several specific questions related to rhinitis. The EGEA study offers the unique opportunity to study rhinitis separately in participants with (AS+) and without (AS-) asthma. The objective of this study was to identify distinct types of rhinitis using unsupervised learning methods in adults from the EGEA study.

Methods

Study design

EGEA is a French case-control and family study based on an initial group of asthma cases and their first-degree relatives, and a group of controls (EGEA1 [12,13], n = 2047; https://egeanet.vjf.inserm.fr). A first follow-up was conducted between 2003 and 2007.

Setting

Protocol and descriptive characteristics of the EGEA study have been previously published [12]. Briefly, 2047 children (<16 years) and adult participants were enrolled at baseline, including 348 participants with current asthma from chest clinics, their 1244 first-degree relatives, and 415 population-based controls. Approximately 12 years later, this population was contacted (EGEA2 [14]). Among the alive cohort (n = 2002), 92% (n = 1845) completed a short self-administered questionnaire and among them 1601 had a complete examination. All participants responded to questionnaires based on international standardized tools to diagnose asthma and to determine respiratory and allergic symptoms, treatments, and environmental exposures.

Participants

The present cross-sectional analysis includes adults at EGEA2 (n = 1571 adults, ≥16 years) without missing data on rhinitis, allergic sensitization and asthma (n = 983, 41% with asthma Fig 1).

Ethics

Ethical approval was obtained from the relevant institutional review board committees (Cochin Port-Royal Hospital and Necker-Enfants Malades Hospital, Paris). Written informed consent was signed by all participants.

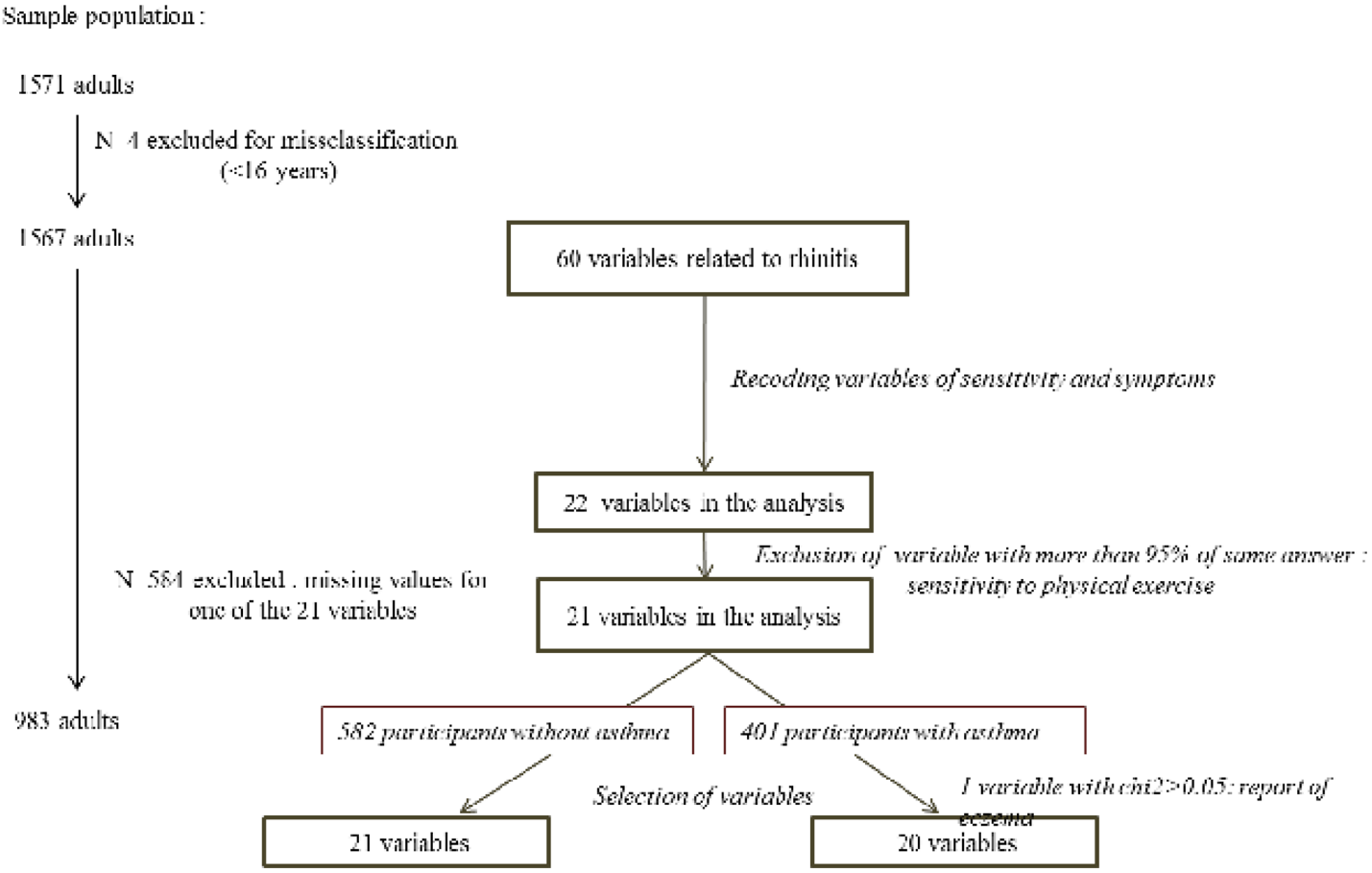


Fig 1. Flow-chart of the variables and of the participants included in the analysis.

doi:10.1371/journal.pone.0136191.g001

Variables

Rhinitis symptoms. Report of nasal symptoms were defined as a positive answer to *"Have you ever had a problem with sneezing, or a runny or a blocked nose when you did not have a cold or the flu?"*. Eyes-associated symptoms were defined as a positive answer to *"Have you ever had itchy or watering eye when you have these nose problems?"*. Current nasal symptoms were defined as a positive answer to the question over the last 12 months. Nasal symptoms were considered as persistent if they occur more than a month per year. They were considered as persistent$_{low}$ if they occur less than 4 days per week and persistent$_{high}$ if they occur more than 4 days per week. Moreover, if the symptoms occurred less than a month per year, persistence of nasal symptoms was considered as intermittent. This classification was built as close as possible to the ARIA guidelines [1], but with some modifications. Answers to the question "*Have these nose problems disturbed you daily activities?*" enabled a score of disturbance from 0 to 3 (0: no, 1: a little bit, 2: moderately, 3: a lot). Participants reported the months in which they had nasal symptoms, and a seasonal pattern was created according to the answer: 0 if no symptom, 1 if symptoms in spring (hay fever), 2 if symptoms in spring/summer, 3 if symptoms in fall/winter, 4 if symptoms all over the year and 5 for the others. Sensitivity to trigger was defined as a positive answer to *"Trigger x usually provoking rhinorrhea"* and *"Trigger x usually provoking sneezing"*. The sensitivity for different triggers was available for animals, weed/flower, dust, cold air, physical exercise, weather, and tobacco smoke exposure (see questionnaires on https://egeanet.vjf.inserm.fr). This sensitivity was coded for the analysis 0: no sensitivity, 1: rhinorrhea or sneezing and 2: rhinorrhea and sneezing. Reports of allergic rhinitis by participants were defined as a positive answer to *"Have you ever had allergic rhinitis?"*, and in the same way for hay fever: *"Have you ever had hay fever?"*. The diagnostic of allergy by a physician was defined as a positive answer to *"Has a doctor ever told you that you are allergic?"*. Positive answers to conjunctivitis, sinusitis and eczema were also considered.

Use of medication for rhinitis. Report of use of medication relative to rhinitis were obtained by a positive answer to either: "*Have you took nasal sprays to treat disorders of the nose in the last 12 months?*" or to "*Have you took pills, capsules, tablets or drugs (other than nasal spray) to treat disorders of the nose in the last 12 months?*".

Asthma. Participants with asthma were defined by a positive answer to either: "*Have you ever had attacks of breathlessness at rest with wheezing?*", or "*Have you ever had asthma attacks?*", or if they were recruited as asthmatic cases at the first survey [12].

Allergic sensitization. Allergic sensitization was defined by a positive skin prick test (SPT +) with a mean wheal diameter $\geq$3mm than the negative control for at least one of 12 aeroallergens (indoor: cat, *Dermatophagoides pteronyssinus*, *Blattela germanica*, outdoor: olive, birch, *Parieteria judaica*, timothy grass, *Cupressus and* ragweed pollen, and molds: *Aspergillus*, *Cladosporium herbarum*, *Alternaria tenuis*). Report of allergic immunotherapy since the first survey (EGEA1) was also available.

Statistical methods

To take into account the specific design of the EGEA study, we conducted the analyses separately in participants without and with asthma.

***Hypothesis-driven*: classical phenotypes.** The analysis based only on the report of nasal symptoms (yes/no) and allergic sensitization (yes (SPT+)/no) enabled to define four profiles separately for participants with and without asthma: phenotype 1: no nasal symptoms and no allergic sensitization, phenotype 2: allergic sensitization only, phenotype 3: nasal symptoms without allergic sensitization and phenotype 4: nasal symptoms and allergic sensitization.

These profiles have already been used to study rhinitis and its relationship with other respiratory diseases [15].

***Data-driven*: novel phenotypes.** *Data and variable selection*. Sixty variables were first considered, known to be commonly associated with rhinitis or allergic sensitization. After recoding and grouping the variables of the sensibility to different triggers, skin prick test (SPT) and symptoms, 22 variables were available. Sensitivity to "physical exercise" having more than 95% of the same answer was excluded. Twenty-one variables were selected for the analysis: report of nasal symptoms, current/ever symptoms, persistence and disturbance of these symptoms, seasonal pattern, sensitivity to seven triggers, report of allergic rhinitis, hay fever, conjunctivitis, sinusitis and eczema, report of diagnostic of allergy by a physician, SPT, report of spray, report of drug except spray, and allergic immunotherapy since the last survey. A variable selection step (chi2 p-value lower than 0.05) led to select 21 variables for As- and 20 for As+ (Fig 1) and finally the analysis included 983 participants (582 As- and 401 As+) with no missing data (Fig 1 and Table B in S1 Supporting Information).

Missing Data. Participants included in the analysis had no missing values, as the data set was built according to that criterion.

Statistical analysis. To describe the phenotypes of disease without the need for historical or a priori assumptions, cluster analysis–or clustering- was used [16]. Cluster analysis is a data mining tool for dividing subjects into several groups so that subjects in the same group are more similar (or related) to each other than to those from others groups. This technique defines the distance of each subject from each other based on the combined values-the multidimensional vector- of their measured characteristics.

Mixture model. The mixture model is a flexible and powerful parametric algorithm of clustering.where each cluster is mathematically represented by a parametric distribution. The entire data set is then modeled by a mixture of these distributions [17]. The number of clusters associated with the lowest Bayesian Information Criterion (BIC) was chosen. As the solution may depend on the initialization, the algorithm was repeated 100 times and the model with the highest likelihood for mixture model was selected. The χ^2 test was used to analyze differences between groups for all qualitative variables. ANOVA was used to compare continuous variable according to the group. To display the subjects in two-dimensional space, multiple correspondence analysis was generated from the dataset; each subject was plotted along the two firsts components.

Tree analysis. To assess which of the 21 or 20 variables were most predictive of the finale cluster, recursive partitioning based on Classification and Regression Tree (CART) was used. The Gini index was used as the splitting index. The dataset was divided into a training set (70% of the original sample) and a validation set (30% of the original sample) to avoid overfitting. Accuracy was used to select the optimal model using the largest value. The validation of the model on the validation set was assessed using the error-rate value of prediction. Results were expressed as percentage of participants assigned to the right cluster (100%-error rate).

Bias. Participants included in the analyses (983) were not significantly different of those not included in the analyses (n = 588, see Table A in S1 Supporting Information) neither for age, sex, body mass index (BMI), nor for nasal symptoms, allergic sensitization, lung function and asthma status.

Due to the familial design of the study, a sensitivity analysis was conducted in a sub-sample of the population with one randomly selected member per family (n = 684 participants, 420 without asthma and 264 with asthma).

All the analyses were performed using the R statistical software. The Rmixmod package (http://cran.r-project.org/web/packages/Rmixmod/index.html) was used to run the algorithm of mixture models, and the rpart package was used to perform the tree analysis.

Table 1. Characteristics of adult participants.

		All (n = 983)	Participants without asthma (n = 582)	Participants with asthma (n = 401)	P value*
Age, mean ± sd		42.6 ± 16.5	45.9 ± 15.9	37.7 ± 16.1	<0.001
Sex, women %		49.5	51.6	46.6	0.13
Tobacco status, %	**Non-smoker**	49.6	48.0	51.9	0.052
	Ex-smoker	26.3	29.1	22.2	
	Smoker	24.1	22.9	25.9	
BMI(kg/m2), %	**<20**	10.7	9.6	12.2	0.09
	[20–25]	49.6	48.5	51.4	
	[25–30]	29.4	32.3	25.2	
	> = 30	10.3	9.6	11.2	
Educational level, %	**Low**	24.5	28.7	18.5	<0.001
	Med	27.7	25.1	31.5	
	high	47.8	46.2	50.0	
Allergic sensitization, %		65.2	46.8	82.0	<0.001
Ever asthma, %		40.8	-	-	
BHR#, n		663	396	n = 267	
%		44.3	27.0	70.0	<0.001
FEV_1% predicted, mean±sd		102 ± 18	106 ± 16	97 ± 18	<0.001
Nasal symptoms, %		58.9	45.5	78.3	<0.001
Reports of AR, %		36.2	21.8	57.1	<0.001
Reports of Hay fever, %		38.8	24.7	59.1	<0.001

BMI = Body Mass Index, FEV1 = Forced Expiratory Volume, AR: allergic rhinitis

#: BHR: Bronchial Hyper Responsiveness (Methacholine test, PD20≤4 mg, Methacholine challenge test was not performed if baseline FEV1 <80% predicted, PD20 = Provocative Dose). BHR was then available for 663 participants (396 without asthma and 267 with asthma).

SPT+: a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens.

* comparing participants without and with asthma (χ_2 test)

doi:10.1371/journal.pone.0136191.t001

Results

The characteristics of the 983 adults according to their asthma status are summarized in Table 1. Participants with asthma had significantly lower Forced Expiratory Volume in one second (FEV_1) level, more often bronchial hyper-responsiveness (BHR), allergic sensitization (SPT+), and reported more often nasal symptoms, AR and hay fever than participants without asthma.

Hypothesis-driven (classical phenotypes) (Table C and D in S1 Supporting Information)

Varied prevalence of the four phenotypes were observed according to the asthma status: Phenotype 1: no symptoms, no allergic sensitization (39% for As- and 4% for As+), phenotype 2: allergic sensitization only (15 for As-and 17% for As+), phenotype 3: nasal symptoms without allergic sensitization (24 for As- and 14% for As+), and phenotype 4: nasal symptoms and allergic sensitization (22 for As- and 65% for As+). Whatever the asthma status, participants of phenotype 4 had the highest rates of hay fever report, allergic conjunctivitis report and sensitivity to hay/flower and animals. Participants of phenotype 3 had the highest rates of sinusitis report and of sensitivity to cold air.

Data-driven (novel phenotypes obtained by cluster analysis)

A three-cluster model was selected as the best model for both As- and As+ participants using the BIC criterion (S1 Fig) and the three clusters were well separated (Fig 2) whatever the asthma status. In participants without asthma: 55% of the participants were in cluster A, 23% in cluster B (23%) and 22% in cluster C. In participants with asthma: 22% of the participants were in the cluster A', 36% in the cluster B' and 42% in cluster C'.

Whatever the asthma status, cluster A and A' were characterized by the absence of nasal symptoms, low reports of AR, hay fever and sensitivity to all triggers (Tables 2 and 3).

In participants without asthma (Table 2). **Cluster B** was characterized by the highest rate of nasal symptoms without eyes-symptoms associated, a high report of sinusitis and eczema, and a low report of AR, hay fever and conjunctivitis as compared to cluster A. The rate of allergic sensitization was lower than for cluster A. Sensitivities to different triggers were lower for hay/flower and animals but higher for cold air, compared to cluster A (Table 2). **Cluster C** was characterized by the highest rate of nasal symptom mostly associated with eyes-symptoms, the highest rate of SPT, the highest rate of sinusitis, eczema and conjunctivitis reports and the highest rates of sensitivity to hay/flower, animals, dust and weather.

The allergic sensitization was mostly monosensitization for clusters A and B while it was mostly polysensitization for cluster C (Table 2). Among participants with allergic sensitization (SPT+), 61% of cluster A, 54% of cluster B, and 30% of cluster C were monosensitized, mostly for *Dermatophagoides pteronyssinus*.

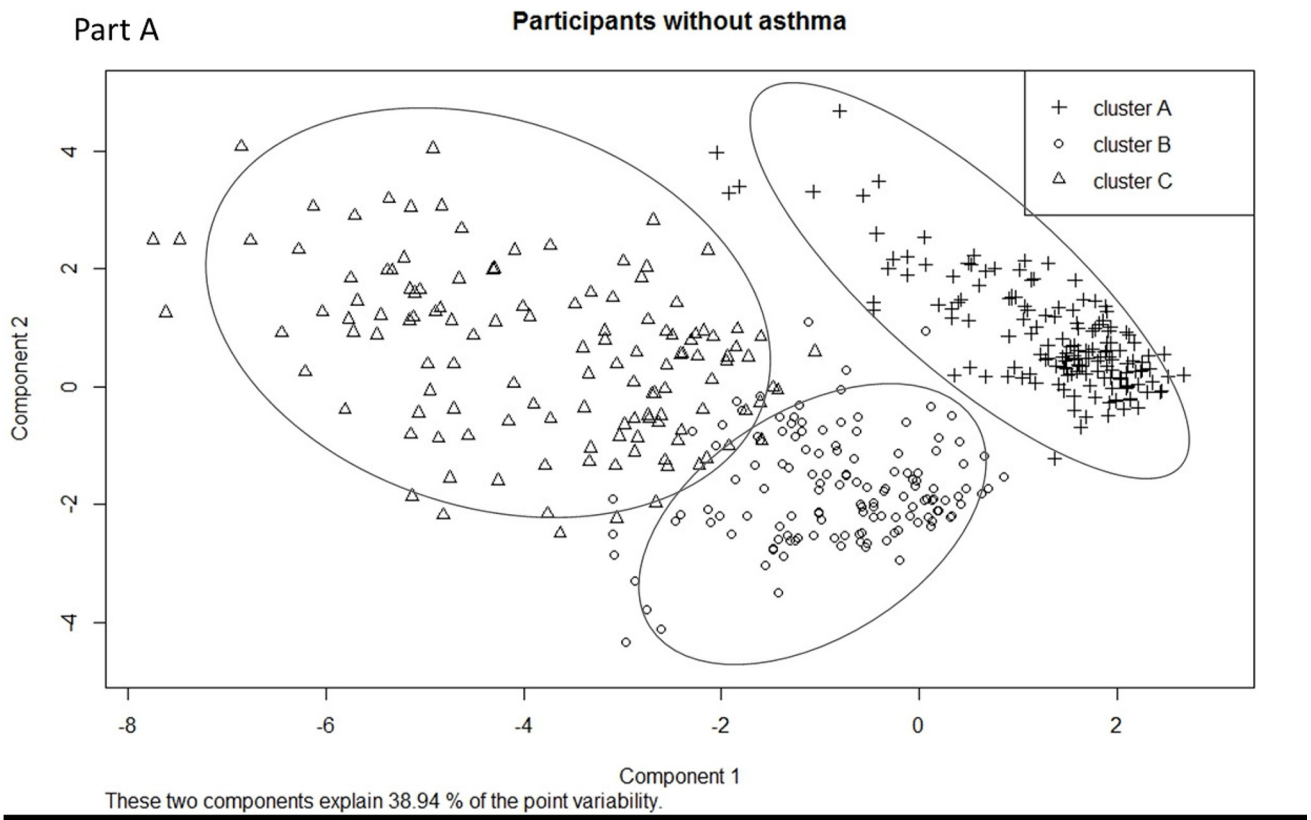
Regarding seasonality of symptoms, cluster B reported symptoms all over the year whereas cluster C reported symptoms mainly during spring (hay/flower season). The score of disturbance was higher for cluster C than for cluster B. No significant difference between clusters was observed in term of persistent or intermittent symptoms.

In participants with asthma (Table 3). **Cluster B'** was characterized by a high rate of nasal symptoms, a low report of AR, hay fever and conjunctivitis and the lowest rate of SPT. Sensitivities to different triggers were low for hay/flower and animals but high for cold air, tobacco and weather. **Cluster C'** was characterized by the highest rate of nasal symptoms with eyes-symptoms, the highest rate of allergic sensitization, the highest rates of report of AR, hay fever, sinusitis and conjunctivitis and the highest rates of sensitivity to hay/flower, animals, dust and weather.

The allergic sensitization rate was high whatever the cluster (Table 3), and mostly characterized by a polysensitization. Similarly to participants without asthma, monosensitization was higher for clusters A' and B'. Among participants with allergic sensitization (SPT+), 34% of cluster A', 29% of cluster B', and 16% of cluster C' were monosensitized mostly for *Dermatophagoides pteronyssinus*. The polysensitization rate was particularly high in cluster C' (80%).

Cluster B' and cluster C' reported symptoms all over the year but cluster C' had a very high peak during spring. The score of disturbance due to nasal symptoms was higher for cluster C' than for cluster B'. Cluster C' declared more $\text{persistent}_{\text{high}}$ than $\text{persistent}_{\text{low}}$ symptoms while cluster B' declared more intermittent or $\text{persistent}_{\text{low}}$ symptoms. The age of onset of asthma was lower for participants of cluster C' than for cluster A' and B'. BHR was higher in participants of cluster B' than in participants of clusters A' or C', but the difference was not significant.

Whatever the asthma status, the report of spray or pills/tablet use to nasal problem was higher for cluster C (respectively C') than for cluster B (resp. B'). Participants of cluster B (resp. B') reported more use of spray than pills/tablet whereas participants of cluster C (resp. C') reported more use of pills/tablet than spray. Age of onset of nasal symptoms was lower in participants with asthma than in those without asthma, and whatever the asthma status,


Part A
Participants without asthma
cluster A
cluster B
cluster C
Component 2
Component 1
-8
-6
-4
-2
0
2
4
These two components explain 38.94 % of the point variability.

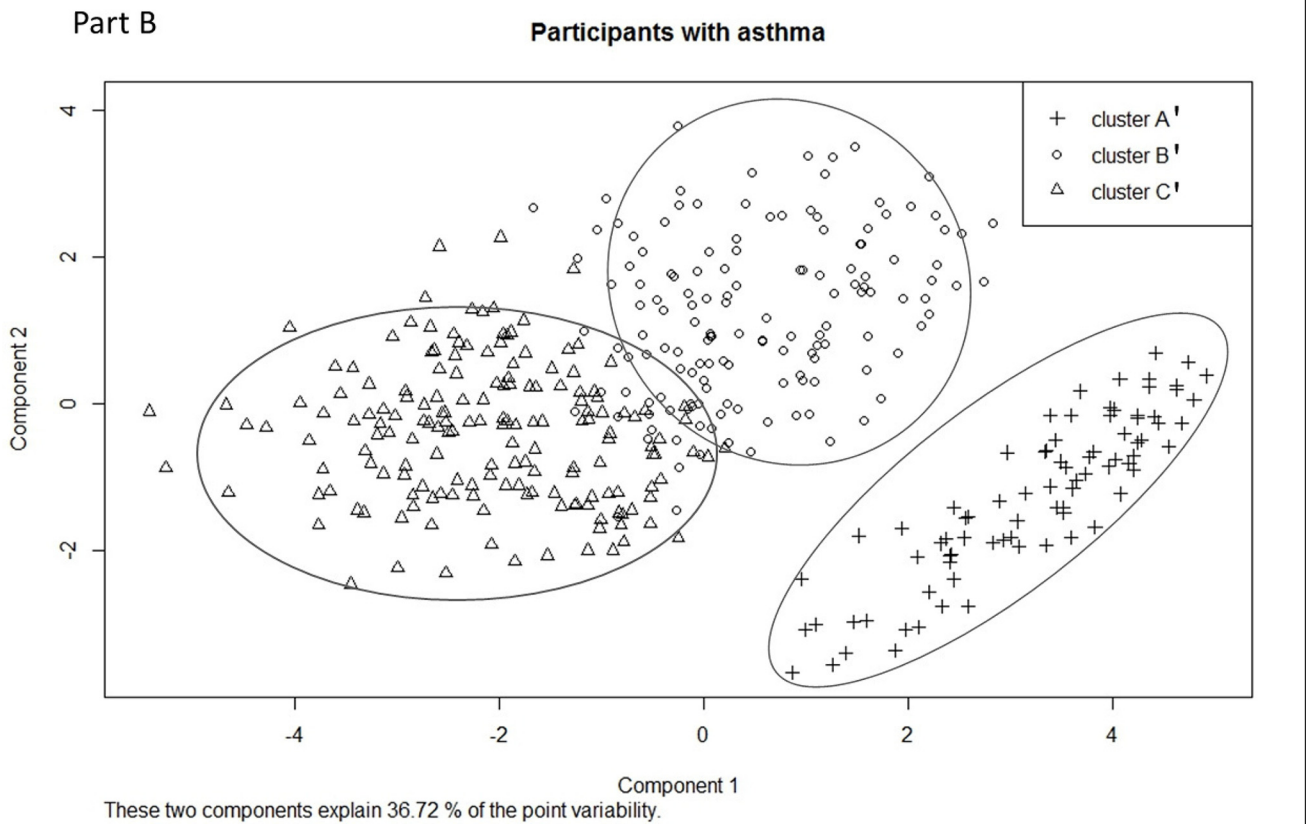

Part B
Participants with asthma
cluster A'
cluster B'
cluster C'
Component 2
Component 1
-4
-2
0
2
4
These two components explain 36.72 % of the point variability.

Fig 2. Visualization of the clusters for participants without asthma (Part A) and participants with asthma (Part B) on the first factorial map.

doi:10.1371/journal.pone.0136191.g002

participants of cluster C (resp. C') had an age of onset of nasal symptoms significantly lower than participants of cluster B (resp. B').

Comparison between *data-driven* clusters and *hypothesis-driven* phenotypes (Tables 4 and 5)

Clusters obtained by *data-driven* approach may be easily assimilated to no rhinitis (NoR: cluster A and A'), non-allergic rhinitis (NAR: cluster B and B') and allergic rhinitis (AR: cluster C and C') based on their characteristics. These clusters are similar to the phenotypes 1, 3 and 4 from the classical *hypothesis-driven* phenotypes a *prima facie* but differ in their internal characteristics and particularly regarding the allergic sensitization. When comparing *data-driven* and *hypothesis-driven* approach, 10% of participants without asthma were not classified in the same category by the two approaches and 26% of participants with asthma. Considering only participants having nasal symptoms, 21% of participants without asthma were not classified in the same category by the two approaches and 30% of participants with asthma.

Decision tree

For participants without asthma, a classification tree on the 21 variables enabled to highlight 4 variables as being the most important to discriminate the cluster and particularly to distinguish cluster B from cluster C: report of nasal symptoms, report of AR, sensitivity to hay/flowers stimuli and type of nasal symptoms-with or without eyes symptoms- (Fig 3, Part A). Using only these 4 variables, 96% of the participants were assigned to the correct cluster.

For participants with asthma, a classification tree on the 20 variables enabled to highlight 4 variables as being the most important to discriminate the cluster and particularly to distinguish cluster B' from cluster C': report of nasal symptoms, sensitivity to "hay/flower" stimuli, diagnosis of allergy by a MD and report of hay fever (Fig 3, Part B). Using only these 4 variables, 87% of the participants were assigned to the correct cluster.

Sensitivity analysis

The cluster analysis on the sub-sample of the population including only one member per family has shown very similar results than the study on the 983 participants (same number of cluster, same characteristics–data not shown-).

Discussion

This study using a clustering approach identified three rhinitis phenotypes in adults with almost no overlap between them. They are similar to the *hypothesis-driven* phenotypes of no rhinitis (NoR: cluster A and A'), non-allergic rhinitis (NAR: cluster B and B') and allergic rhinitis (AR: cluster C and C'). However, hypothesis and data-driven phenotypes differ in terms of allergic sensitization. Near of a quarter of participants without asthma would have been considered as having allergic rhinitis considering the *hypothesis-driven* phenotypes whereas they have a non-allergic rhinitis pattern. Our study was able to highlight the importance of the NAR phenotypes, less understood and which need to be studied [18] and enhanced the importance of the non-allergic component in rhinitis. In participants with asthma, the AR cluster was associated with the highest rate of allergic sensitization and number of allergens, suggesting a comorbid effect of asthma and allergic rhinitis on the polysensitization.

Table 2. Characteristics of participants without asthma according to each cluster.

		Cluster A-no rhinitis- (n = 317)	Cluster B-non-allergic rhinitis- (n = 136)	Cluster C-allergic rhinitis- (n = 129)	p-value*
Age, mean ± sd		46.7 ± 16.2	48.9 ± 15.5	40.9 ± 14.5	<0.001
Sex, women %		46.7	58.1	56.6	0.036
Tobacco, %	**Non-smoker**	47.2	47.0	51.2	0.90
	Ex-smoker	30.0	30.2	25.6	
	Smoker	22.8	22.88	23.3	
BMI (kg/m2), %	**<20**	10.1	8.8	9.3	0.53
	[20–25]	47.0	47.8	52.7	
	[25–30]	33.4	36.0	25.6	
	> = 30	9.5	7.4	12.4	
Educational level, %	**Low**	31.2	34.6	16.3	0.007
	Medium	25.2	23.5	26.4	
	High	43.5	41.9	57.4	
Nasal symptoms, %	**No symptoms**	100.0	0.0	0.0	<0.001
	Symptoms without eyes symptoms	0.0	71.3	17.1	
	Symptoms with eyes symptoms	0.0	28.7	82.9	
Type of nasal symptoms, %	**No symptoms**	100	0.0	0.0	-
	Symptoms: ever but not current	0.0	4.4	0.0	
	Ever and current symptoms	0.0	95.6	100	
Persistence of nasal symptoms, %	**Intermittent**	-	50.0	40.3	0.22
	$Persistent_{low}$	-	18.4	25.6	
	$Persistent_{high}$	-	31.6	34.1	
Disturbance due to nasal	**No**	-	77.2	42.6	<0.001
symptoms, %	**Low**	-	17.7	39.5	
	Medium	-	4.4	14	
	High	-	0.7	3.90	
Allergic sensitization,%	SPT = 0	71.9	80.9	23.3	<0.001
	SPT = 1	17	10.3	23.3	
	SPT = 2	7.6	4.4	19.4	
	SPT>2	3.5	4.4	34.1	
Report of diagnosis of allergy by a physician, %		15.5	18.4	72.1	<0.001
Immunotherapy since first survey (EGEA1)		1.60	0.0	13.2	<0.001
Age of onset of nasal symptoms, mean ± sd	**(n = 224)**	-	33.7 ± 18.2	22.1 ± 14.1	<0.001
Report of allergic rhinitis, %		5.70	6.6	77.5	<0.001
Report of hay fever, %		10.7	10.3	74.4	<0.001
Report of conjunctivitis, %		13.4	22.1	49.6	<0.001
Report of sinusitis, %		34.7	55.1	56.6	<0.001
Report of eczema, %		22.1	30.9	36.4	0.005
Sensitivity to hay/flowers, %	**No sensitivity**	89.3	82.4	29.5	<0.0001
	Rhinorrhea or sneezing	8.2	16.9	30.2	

(*Continued*)

Table 2. (*Continued*)

		Cluster A-no rhinitis- (n = 317)	Cluster B-non-allergic rhinitis- (n = 136)	Cluster C-allergic rhinitis- (n = 129)	p-value*
	Rhinorrhea and sneezing	2.5	0.7	40.3	
Sensitivity to animals, %	**No sensitivity**	98.1	100	78.3	<0.0001
	Rhinorrhea or sneezing	1.60	0.0	13.2	
	Rhinorrhea and sneezing	0.30	0.0	8.50	
Sensitivity to dust, %	**No sensitivity**	74.1	58.1	27.1	<0.0001
	Rhinorrhea or sneezing	24.3	39.7	46.5	
	Rhinorrhea and sneezing	1.60	2.20	26.4	
Sensitivity to tobacco smoke, %	**No sensitivity**	98.1	90.4	87.5	<0.0001
	Rhinorrhea or sneezing	1.60	7.40	11.7	
	Rhinorrhea and sneezing	0.30	2.20	0.80	
Sensitivity to cold air, %	**No sensitivity**	84.2	66.9	67.4	<0.0001
	Rhinorrhea or sneezing	15.1	30.2	27.9	
	Rhinorrhea and sneezing	0.60	2.90	4.70	
Sensitivity to weather, %	**No sensitivity**	97.5	88.2	83.0	<0.0001
	Rhinorrhea or sneezing	1.60	11.8	10.8	
	Rhinorrhea and sneezing	0.90	0.00	6.20	
Use of nasal spray in the last 12 months, %		23.0	39.7	54.3	<0.0001
Use of other drug in the last 12 months, %		17.7	27.9	62.0	<0.0001

BMI = Body Mass Index

*p-value overall

doi:10.1371/journal.pone.0136191.t002

Participants of the study had a very good phenotypic characterization of respiratory health and allergic sensitization that gave us the opportunity to consider several questions related to rhinitis. The design of the study allowed us to compare the characteristics of rhinitis phenotypes according to the asthma status. One of the limitations is that the sample is not big enough to study finest clusters and particularly mixed rhinitis (participants having non allergic and allergic rhinitis). Our analysis did not identify some very specific adult rhinitis phenotypes such as hormonal rhinitis [19] probably because of their low prevalence. Overall, to our knowledge, our study is the first with such detailed questionnaires and allergic sensitization available. As the analyses were performed separately for participants without and with asthma, our results cannot be transposed to population-based studies.

Cluster of rhinitis have consistent characteristics with previous literature and clinician's knowledge. We reported that AR cluster was more related to conjunctivitis and eyes-associated symptoms whereas NAR cluster was more related to sinusitis. NAR cluster was more associated with sensitivity to trigger as cold air whereas AR cluster was more associated with sensitivity to multiple allergens as pet, hay, and flower. Age at onset of nasal symptoms was lower for AR cluster than for NAR cluster. These results are concordant with several papers comparing

Table 3. Characteristics of participants with asthma according to each cluster.

		Cluster A'-no rhinitis- (n = 87)	Cluster B'-non-allergic rhinitis- (n = 144)	Cluster C'-allergic rhinitis-(n = 170)	p-value*
Age, mean ± sd		39.9 ± 16.8	38.5 ± 17.6	35.9 ± 14.3	<0.001
Sex, women %		44.8	44.4	49.4	0.63
Tobacco, %	Non smoker	51.7	45.1	57.7	0.054
	Ex-smoker	28.7	25.0	16.5	
	Smoker	19.5	29.9	25.9	
BMI (kg/m2), %	<20	8.1	11.1	15.3	0.32
	[20–25]	52.9	47.2	54.1	
	[25–30]	25.3	28.5	22.4	
	> = 30	13.8	13.2	8.2	
Educational level, %	Low	25.3	20.8	13.0	0.12
	Medium	31.0	31.9	32.4	
	High	43.7	47.2	55.6	
Nasal symptoms,	No symptoms	100	0.0	0.0	<0.001
%	Symptoms without eyes symptoms	0.0	43.8	7.6	
	Symptoms with eyes symptoms	0.0	56.2	92.4	
Type of nasal symptoms	No symptoms	100	0.0	0.0	<0.001
%	Symptoms: ever but not current	0.0	1.4	0.0	
	Ever and current symptoms	0.0	98.6	100	
Persistence of nasal symptoms, %	Intermittent	-	50.0	20.6	<0.001
	$Persistent_{low}$	-	29.2	31.2	
	$Persistent_{high}$	-	20.8	48.2	
Allergic sensitization,%	SPT = 0	19.5	33.3	4.1	<0.001
	SPT = 1	27.6	19.4	15.3	
	SPT = 2	20.7	14.6	19.4	
	SPT>2	32.2	32.6	61.2	
Report of diagnosis of allergy by a physician, %		58.6	60.4	96.5	<0.001
Immunotherapy since first survey (EGEA1)		8.0	8.3	22.9	<0.001
Age of onset of nasal symptoms, mean ± sd	(n = 290)	-	19.3 ± 86.9	11.5 ± 10.0	<0.001
Report of allergic rhinitis, %		26.4	41.7	85.9	<0.001
Report of hay fever, %		35.6	34.0	92.4	<0.001
Report of conjunctivitis, %		26.4	30.6	69.4	<0.001
Report of sinusitis, %		46.0	49.3	60.0	0.053
Report of eczema, %		42.5	47.2	53.5	0.22
Sensitivity to hay/flowers, %	No sensitivity	77.0	76.4	10.0	<0.0001
	Rhinorrhea or sneezing	12.6	20.1	30.0	
	Rhinorrhea and sneezing	10.3	3.5	60.0	
Sensitivity to animals, %	No sensitivity	88.5	81.3	52.9	<0.0001
	Rhinorrhea or sneezing	6.9	12.5	18.8	

(*Continued*)

Table 3. (*Continued*)

		Cluster A'-no rhinitis- (n = 87)	Cluster B'-non-allergic rhinitis- (n = 144)	Cluster C'-allergic rhinitis-(n = 170)	p-value*
	Rhinorrhea and sneezing	4.6	6.3	28.2	
Sensitivity to dust, %	**No sensitivity**	64.4	50.7	16.5	<0.0001
	Rhinorrhea or sneezing	25.3	37.5	42.9	
	Rhinorrhea and sneezing	10.3	11.8	40.6	
Sensitivity to tobacco smoke, %	**No sensitivity**	95.4	90.2	78.2	0.0002
	Rhinorrhea or sneezing	1.0	9.1	14.1	
	Rhinorrhea and sneezing	0.0	0.7	4.7	
Sensitivity to cold air, %	**No sensitivity**	86.2	71.5	63.5	0.001
	Rhinorrhea or sneezing	13.8	26.4	30	
	Rhinorrhea and sneezing	0.0	2.1	6.5	
Sensitivity to weather, %	**No sensitivity**	94.3	86.8	70.6	<0.0001
	Rhinorrhea or sneezing	5.8	9.7	14.7	
	Rhinorrhea and sneezing	0.0	3.5	14.7	
Disturbance due to nasal symptoms, %	**No**	-	65.3	29.4	<0.001
	Low	-	20.1	34.1	
	Medium	-	11.1	23.5	
	High	-	3.5	12.9	
Use of nasal spray in the last 12 months, %		44.8	47.2	64.1	0.0018
Use of other drug in the last 12 months, %		42.5	42.4	80.6	<0,0001
Age of onset of asthma, mean ± sd		16,8±16,2	15,6±15,7	12,2±13,2	0.6 (A vs B) and 0.02 (A vs C)
BHR, %	**(n = 267)**	64.4	75.0	69.0	0.36
FEV_1% predicted, mean±sd		96 ± 0.18	98 ± 0.16	96 ± 0.22	0.42

BMI = Body Mass Index, FEV1 = Forced Expiratory Volume, #: BHR: Bronchial Hyper Responsiveness (Methacholine test, PD20≤4 mg, Methacholine challenge test was not performed if baseline FEV1 <80% predicted, PD20 = Provocative Dose). BHR was then available for 663 participants (396 without asthma and 267 with asthma).

*p-value overall

doi:10.1371/journal.pone.0136191.t003

allergic rhinitis to non-allergic rhinitis [20], [21], [22]. Overall, it is reassuring that *prima facie*, unsupervised approaches find similar phenotypes than the ones used in the clinical setting.

Interestingly, we observed that AR cluster was associated with more severe symptoms (greater disturbance) than NAR cluster. This result is consistent with the studies by Bachert [23] and Di Lorenzo [21] but discordant with the study by Molgaard [24]. This discordance between studies seems not to be due to the design of the studies, but to the difference in the definitions of the types of rhinitis and particularly in the way that allergic and non-allergic rhinitis were differentiated. Overall, the definitions and particularly the way to define the allergic part of rhinitis seem to be crucial to establish the characteristics of the phenotypes. Furthermore,

Table 4. Comparison of the repartition of the participants without asthma into the different *hypothesis-driven*'s phenotypes and *data-driven*'s cluster.

		Data-driven clusters			
	n (%)	A (No rhinitis)	B (NAR)	C (AR)	
Hypothesis-driven Phenotypes	1 (no symptoms, no SPT)	228 (39)	0	0	228
	2 (no symptoms, SPT+)	89 (15)	0	0	89
	3 (symptoms, no SPT ~ NAR)	0	110 (19)	**30 (5)**	140
	4 (symptoms, SPT+ ~AR)	0	**26 (5)**	99 (17)	125
		317	136	129	**582**

doi:10.1371/journal.pone.0136191.t004

whatever the cluster of rhinitis (NAR or AR), we found that almost all of the participants who reported ever rhinitis also reported current nasal symptoms which suggest that considering rhinitis ever or current rhinitis would give the same result.

Prevalence and repartition of non-allergic and allergic rhinitis are very different according to the study: between 63% and 77% of rhinitis would be of allergic type [22,24], but some other studies argue that over 75% of rhinitis is non-allergic rhinitis or mixed rhinitis [20]. In our study, we found a higher prevalence of rhinitis in participants with asthma. However, within each asthma status, the prevalence of NAR cluster was similar to that of AR cluster.

Whereas rhinitis is classically divided in allergic and non-allergic rhinitis based on the allergic sensitization, our results suggest that allergic sensitization may be insufficient to differentiate correctly AR and NAR and to make the diagnosis of AR. This result is concordant with a paper studying predictor factors to differentiate between allergic and non-allergic rhinitis in children [25], which found out that features of rhinitis as seasonality, moderate/severe symptoms help in the differentiation of rhinitis. Di Lorenzo [21] has showed that several clinical and laboratory parameters may help to reinforce or exclude the diagnosis of AR obtained with SPT, and Quillen said that: "allergy testing is not necessary in all patients but may be useful in ambiguous or complicated cases"[26]. Finally, Berstein [27] said that "taking into account age of symptom onset, family history, quantification of inciting allergic and/or non-allergic triggers, and seasonality followed by aeroallergen skin testing to assess atopic status has been shown to be the most useful approach for clearly differentiate rhinitis subtypes". Overall, these results are concordant with known complexity to define phenotypes of rhinitis.

This study enabled to validate and confirm phenotypes of rhinitis often described in the literature, but for the first time highlighted in a statistical way. Thanks to a classification tree, our results showed the clinical interest of using only a few numbers of questions to classify the participants in the 3 clusters and particularly to distinguish between non-allergic and allergic rhinitis. These questions are often available in respiratory epidemiological study making easier the reconstruction and the use by general physician and pharmacist.

Table 5. Comparison of the repartition of the participants with asthma into the different *hypothesis-driven*'s phenotypes and *data-driven*'s cluster.

		Data-driven clusters			
	n (%)	A' (No rhinitis)	B' (NAR)	C' (AR)	
Hypothesis-driven Phenotypes	1 (no symptoms, no SPT)	17 (4)	0	0	17
	2 (no symptoms, SPT+)	70 (17)	0	0	70
	3 (symptoms, no SPT ~ NAR)	0	48 (12)	**7 (2)**	55
	4 (symptoms, SPT+ ~AR)	0	**96 (24)**	163 (41)	259
		87	144	170	**401**

doi:10.1371/journal.pone.0136191.t005

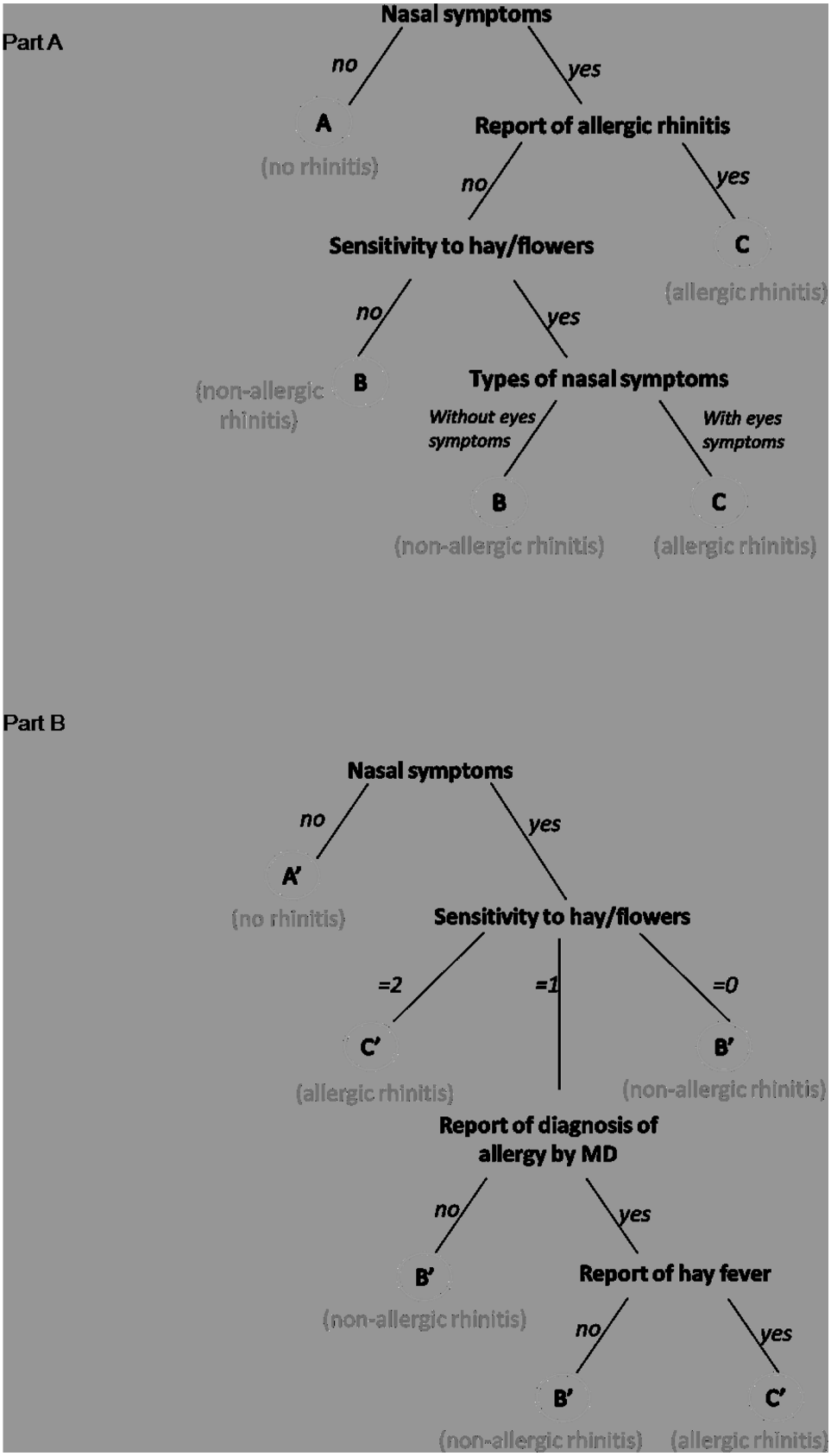

Part A
Nasal symptoms
no
yes
A
(no rhinitis)
Report of allergic rhinitis
no
yes
Sensitivity to hay/flowers
C
(allergic rhinitis)
no
yes
B
(non-allergic rhinitis)
Types of nasal symptoms
Without eyes symptoms
With eyes symptoms
B
(non-allergic rhinitis)
C
(allergic rhinitis)
Part B
Nasal symptoms
no
yes
A'
(no rhinitis)
Sensitivity to hay/flowers
=2
=1
=0
C'
(allergic rhinitis)
B'
(non-allergic rhinitis)
Report of diagnosis of allergy by MD
no
yes
B'
(non-allergic rhinitis)
Report of hay fever
no
yes
B'
(non-allergic rhinitis)
C'
(allergic rhinitis)

Fig 3. Classification tree obtained with the most predictive variables in participants without (Part A) and with asthma (Part B).

doi:10.1371/journal.pone.0136191.g003

In conclusion, taking into account all available specific questions related to rhinitis, a cluster analysis enabled to highlight three clusters of rhinitis with similar characteristics than those known by clinicians but differing according to allergic sensitization, and this whatever the asthma status. The clusters obtained by *data-driven* approach may be considered as "smoothed" phenotypes compared to the ones obtained only using nasal symptoms and allergic sensitization. These clusters could now be used to study the association with biological and environmental factors. Overall, although cluster analysis is thought to be hypothesis generating, studies in asthma, COPD and now rhinitis show that is may also be useful in hypothesis confirmation.

Supporting Information

S1 Fig. BIC criterion according to the number of cluster for participants without (Part A) and with (Part B) asthma.
(EPS)

S1 Supporting Information. Comparison of the characteristics of the participants included and non-included in the analysis (Table A). Missing values for each variables (Table B). Description of the participants without asthma according to the four classical phenotypes (hypothesis driven) (Table C). Description of the participants with asthma according to the four classical phenotypes (hypothesis driven) (Table D).
(DOC)

Acknowledgments

The authors thank all those who participated to the setting of the study and on the various aspects of the examinations involved: interviewers, technicians for lung function testing and skin prick tests, blood sampling, IgE determinations, coders, those involved in quality control, data and sample management and all those who supervised the study in all centers. The authors are grateful to the three CIC-Inserm of Necker, Grenoble and Marseille who supported the study and in which participants were examined. They are also grateful to the three biobanks in Lille (CIC-Inserm), Evry (Centre National du Genotypage) and at Annemasse (Etablissement français du sang) where biological samples are stored.

EGEA cooperative group

Coordination: V Siroux (epidemiology, PI since 2013); F Demenais (genetics); I Pin (clinical aspects); R Nadif (biology); F Kauffmann (PI 1992–2012).

Respiratory epidemiology: Inserm U 700, Paris: M Korobaeff (Egea1), F Neukirch (Egea1); Inserm U 707, Paris: I Annesi-Maesano (Egea1-2); Inserm CESP/U 1018, Villejuif: F Kauffmann, N Le Moual, R Nadif, MP Oryszczyn (Egea1-2), R Varraso; Inserm U 823, Grenoble: V Siroux. Genetics: Inserm U 393, Paris: J Feingold; Inserm U 946, Paris: E Bouzigon, F Demenais, MH Dizier; CNG, Evry: I Gut (now CNAG, Barcelona, Spain), M Lathrop (now Univ McGill, Montreal, Canada).

Clinical centers: Grenoble: I Pin, C Pison; Lyon: D Ecochard (Egea1), F Gormand, Y Pacheco; Marseille: D Charpin (Egea1), D Vervloet (Egea1-2); Montpellier: J Bousquet; Paris Cochin: A Lockhart (Egea1), R Matran (now in Lille); Paris Necker: E Paty (Egea1-2), P Scheinmann (Egea1-2); Paris-Trousseau: A Grimfeld (Egea1-2), J Just.

Data and quality management: Inserm ex-U155 (Egea1): J Hochez; Inserm CESP/U 1018, Villejuif: N Le Moual; Inserm ex-U780: C Ravault (Egea1-2); Inserm ex-U794: N Chateigner (Egea1-2); Grenoble: J Quentin-Ferran (Egea1-2).

Author Contributions

Conceived and designed the experiments: FG BJ JJ RM RN IP VS RV. Performed the experiments: FG JJ RM RN IP. Analyzed the data: EB. Contributed reagents/materials/analysis tools: EB JB BJ RN. Wrote the paper: EB JB BJ RN. Commented on the manuscript, read and approved the final version: EB JB RV FG JJ RM IP VS BJ RN.

References

1. Bousquet J, Khaltaev N, Cruz AA, Denburg J, Fokkens WJ, Togias A, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA2LEN and AllerGen). Allergy: European Journal of Allergy and Clinical Immunology. 2008. p. 8–160. doi: 10.1111/j.1398-9995.2007.01620.x PMID: 18331513
2. Bousquet PJ, Leynaert B, Neukirch F, Sunyer J, Janson CM, Anto J, et al. Geographical distribution of atopic rhinitis in the European Community Respiratory Health Survey I. Allergy Eur J Allergy Clin Immunol. 2008; 63(10):1301–9.
3. Katelaris CH, Lee BW, Potter PC, Maspero JF, Cingi C, Lopatin A, et al. Prevalence and diversity of allergic rhinitis in regions of the world beyond Europe and North America. Clin Exp Allergy. 2012 Feb; 42(2):186–207. doi: 10.1111/j.1365-2222.2011.03891.x PMID: 22092947
4. Bousquet J, Anto J, Auffray C, Akdis M, Cambon-Thomsen a, Keil T, et al. MeDALL (Mechanisms of the Development of ALLergy): an integrated approach from phenotypes to systems medicine. Allergy. 2011 May; 66(5):596–604. doi: 10.1111/j.1398-9995.2010.02534.x PMID: 21261657
5. Pinart M, Benet M, Annesi-Maesano I, von Berg A, Berdel D, Carlsen K-HKCL, et al. Comorbidity of eczema, rhinitis, and asthma in IgE-sensitised and non-IgE-sensitised children in MeDALL: a population-based cohort study. Lancet Respir Med. 2014 Feb; 2(2):131–40. doi: 10.1016/S2213-2600(13)70277-7 PMID: 24503268
6. Izquierdo-Domínguez A, Valero AL, Mullol J. Comparative analysis of allergic rhinitis in children and adults. Curr Allergy Asthma Rep. 2013 Apr; 13(2):142–51. doi: 10.1007/s11882-012-0331-y PMID: 23250586
7. Siroux V, Basagaña X, Boudier a, Pin I, Garcia-Aymerich J, Vesin a, et al. Identifying adult asthma phenotypes using a clustering approach. Eur Respir J. 2011 Aug; 38(2):310–7. doi: 10.1183/09031936.00120810 PMID: 21233270
8. Antó JM. Recent advances in the epidemiologic investigation of risk factors for asthma: a review of the 2011 literature. Curr Allergy Asthma Rep. 2012 Jun; 12(3):192–200. doi: 10.1007/s11882-012-0254-7 PMID: 22422360
9. Burgel P-R, Paillasseur J-L, Caillaud D, Tillie-Leblond I, Chanez P, Escamilla R, et al. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. Eur Respir J. 2010 Sep; 36(3):531–9. doi: 10.1183/09031936.00175109 PMID: 20075045
10. Weatherall M, Travers J, Shirtcliffe PM, Marsh SE, Williams M V, Nowitz MR, et al. Distinct clinical phenotypes of airways disease defined by cluster analysis. Eur Respir J. 2009 Oct; 34(4):812–8. doi: 10.1183/09031936.00174408 PMID: 19357143
11. Kurukulaaratchy RJ, Zhang H, Patil V, Raza A, Karmaus W, Ewart S, et al. Identifying the heterogeneity of young adult rhinitis through cluster analysis in the Isle of Wight birth cohort. Journal of Allergy and Clinical Immunology. Elsevier Ltd; 2015;143–50. doi: 10.1016/j.jaci.2014.06.017 PMID: 25085342
12. Kauffmann F. EGEA—descriptive characteristics. Clin Exp Allergy. 1999; 29:17–21.
13. Kauffmann F, Dizier MH, Pin I, Paty E, Gormand F, Vervloet D, et al. Epidemiological study of the genetics and environment of asthma, bronchial hyperresponsiveness, and atopy: phenotype issues. Am J Respir Crit Care Med. 1997 Oct; 156(4 Pt 2):S123–9. PMID: 9351592
14. Siroux V, Boudier A, Bousquet J, Bresson J-L, Cracowski J-L, Ferran J, et al. Phenotypic determinants of uncontrolled asthma. J Allergy Clin Immunol. 2009 Oct; 124(4):681–7.e3. doi: 10.1016/j.jaci.2009.06.010 PMID: 19665764
15. Shaaban R, Zureik M, Soussan D, Neukirch C, Heinrich J, Sunyer J, et al. Rhinitis and onset of asthma: a longitudinal population-based study. Lancet. 2008 Sep 20; 372(9643):1049–57. doi: 10.1016/S0140-6736(08)61446-4 PMID: 18805333

16. Fraley C, Raftery A. How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. Comput J. 1998 Aug 1; 41(8):578–88.
17. McLachlan GJ, Peel DA. Finite mixture models. 2000.
18. Tran NP, Vickery J, Blaiss MS. Management of Rhinitis: Allergic and Non-Allergic. Allergy, Asthma Immunol Res. 2011; 3(3):148.
19. Settipane RA. Other causes of rhinitis: mixed rhinitis, rhinitis medicamentosa, hormonal rhinitis, rhinitis of the elderly, and gustatory rhinitis. Immunol Allergy Clin North Am. 2011 Aug; 31(3):457–67. doi: 10.1016/j.iac.2011.05.011 PMID: 21737037
20. Bernstein J a. Nonallergic rhinitis: therapeutic options. Curr Opin Allergy Clin Immunol. 2013 Aug; 13 (4):410–6. doi: 10.1097/ACI.0b013e3283630cd8 PMID: 23756873
21. Di Lorenzo G, Pacor ML, Amodio E, Leto-Barone MS, La Piana S, D'Alcamo A, et al. Differences and similarities between allergic and nonallergic rhinitis in a large sample of adult patients with rhinitis symptoms. Int Arch Allergy Immunol. 2011 Jan; 155(3):263–70. doi: 10.1159/000320050 PMID: 21293145
22. Kalpaklioglu a F, Kavut AB. Allergic and nonallergic rhinitis: can we find the differences/similarities between the two pictures? J Asthma. 2009 Jun; 46(5):481–5. doi: 10.1080/02770900902849897 PMID: 19544169
23. Bachert C, Cauwenberge P Van, Olbrecht J, Schoor J Van. Prevalence, classification and perception of allergic and nonallergic rhinitis in Belgium. Allergy. 2006; 61:693–8. PMID: 16677237
24. Molgaard E, Thomsen SF, Lund T, Pedersen L, Nolte H, Backer V. Differences between allergic and nonallergic rhinitis in a large sample of adolescents and adults. Allergy. 2007 Sep; 62(9):1033–7. PMID: 17578499
25. Topal E, Bakirtas A, Yilmaz O, Karagöl IHE, Arslan U, Arga M, et al. Predictive factors to differentiate between allergic and nonallergic rhinitis in children. Int Forum Allergy Rhinol. 2014 Feb 24; 4(6):447–52. doi: 10.1002/alr.21312 PMID: 24574308
26. Quillen DM, Feller DB. Diagnosing rhinitis: allergic vs. nonallergic. Am Fam Physician. 2006 May 1; 73 (9):1583–90. PMID: 16719251
27. Bernstein JA. Characterizing rhinitis subtypes. Am J Rhinol Allergy. 2013; 27(6):457–60. doi: 10.2500/ajra.2013.27.3983 PMID: 24274219

Characterization of rhinitis according to the asthma status in adults using an unsupervised approach in the EGEA study

Emilie Burte[1,2,*], Jean Bousquet,[1,2,3], Raphaëlle Varraso,[1,2], Frédéric Gormand,[4], Jocelyne Just,[5,6], Régis Matran,[7], Isabelle Pin,[8,9,10,11], Valérie Siroux,[8,9,10], Bénédicte Jacquemin,[1,2,12,#], Rachel Nadif,[1,2,$#]

1. INSERM, U1168, VIMA: Aging and chronic diseases. Epidemiological and Public health approaches, F-94807, Villejuif, France
2. Univ Versailles St-Quentin-en-Yvelines, UMR-S 1168, F-78180, Montigny le Bretonneux, France
3 University hospital, Montpellier, France 4. CHU de Lyon, Pneumology Department, Lyon, France
5. Allergology Department, Centre de l'Asthme et des Allergies. Hôpital Armand-Trousseau (APHP), APHP, Paris, France
6. Université Paris 6 Pierre et Marie Curie, Paris, France
7. Univ Lille Nord de France, F-59000, Lille, France
8. INSERM, IAB, Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, F-38000 Grenoble, France.
9. Univ. Grenoble Alpes, F-38000 Grenoble, France.
10. CHU de Grenoble, F-38000 Grenoble, France.
11. CHU de Grenoble, Pediatric Department, F-38000, Grenoble, France
12. CREAL-Centre for Research in Environmental Epidemiology Parc de Recerca Biomèdica de Barcelona, Barcelona, Spain.

***** Corresponding author
E-mail: emilie.burte@inserm.fr (EB)

[#] These authors contributed equally to this work

Methods

A lung function test with methacholine challenge was performed using a standardized protocol with similar equipment across centers according to the ATS/ERS guidelines (S~~E~~1). Methacholine challenge was performed unless baseline FEV1 <80% predicted.

~~**Statistical analysis**~~

~~**Selection of variables to include in the clustering analysis**~~

~~To select the variables to be included in the clustering algorithm, a first clustering -using a mixture model- was performed and the chi2 relationship between the clusters obtained and the 21 variables was estimated. The variables having a p value lower than 0.05 were kept. Therefore, 21 variables were selected for As- and 20 for As+ (Fig 1). A sensitivity analysis was done doing the analysis for As+ on the same 21 variables, but the final partition was more than 99% equal of the one with the 20 variables.~~

References

S~~E~~1. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates a, et al. Standardisation of spirometry. Eur Respir J. 2005 Aug;26(2):319–38.

Table A. Comparison of the characteristics of the participants included and non-included in the analysis

		Not included (n=588)	**Included (n=983)**	**p-value**
age, mean± sd		43.1±16.6	42.6±16.5	0.52
Sex, women %		52.4	49.5	0.28
Tobacco status, %	Non smoker	50.4	49.6	0.40
	Ex smoker	28.2	26.3	
	Smoker	21.3	24.1	
BMI, %	<20	10.8	10.7	0.09
	[20-25]	56	49.6	
	[25-30]	23	29.4	
	>=30	10.2	10.3	
SPT	SPT=0	40.9	44.8	0.61
	SPT=1	20.3	17.9	
	SPT=2	12.8	12.9	
	SPT>2	25.9	24.4	
Ever asthma, %		40.8	40.8	0.99
FEV1, % predict		102.5±0.19	102.4±0.18	0.89
BHR, %		42.9 (n=203)	44.3 (n=663)	0.71
Report of nasal symptoms, %		63.3	58.9	0.09
Reports of AR, %		33.33	36.22	0.26
Reports of hay fever, %		35	38.76	0.14
Educational level, %	Low	24.6	24.5	0.94
	Med	26.9	27.7	
	high	48.5	47.8	

BMI= Body Mass Index, SPT: Skin Prick Test, BHR: Bronchial HyperResponsiveness (Methacholine test, PD20≤4 mg), FEV1= Forced Expiratory Volume in 1s , AR: allergic rhinitis

Table B. Missing values for each variables

	Variable	**Missing (N=)**
Nasal symptoms	Type (associated with eyes symptoms or not)	6
	Current or ever	12
	Persistence	103
	Disturbance	103
	Month profile	127
Report of other related disease	Report of allergic rhinitis	33
	Report of hay fever	25
	Report of conjunctivitis	30
	Report of sinusitis	9
	Report of eczema	10
Sensitivity to stimuli	Animals	64
	Hay/flowers	59
	Tobacco	2
	Cold air	53
	Effort	74
	Dust	58
	Weather	192
Drug consumption	Spray for nasal problem , last 12 month	13
	Other drugs for nasal problem(not spray), last 12 months	107
Report	Desensitization since first survey (EGEA1)	89
	Diagnosis of allergy (by a physician)	86
Allergic sensitization	SPT+	266
Asthma status		0

Table C. Description of the participants without asthma according to the four classical phenotypes (hypothesis driven)

		Phenotype 1: No symptoms, no SPT (n=228)	**Phenotype 2: SPT only(n=89)**	**Phenotype 3: symptoms, no SPT(n=140)**	**Phenotype 4: Symptoms and SPT (n=125)**
age, mean± sd		48.3±15.9	42.7±16.6	50.6±14.8	38.7±13.8
Sex, women %		49.1	40,5	61,4	52.8
Tobacco status, %	**Non-smoker**	45.8	50.6	47.9	50.4
	Ex-smoker	32.6	23.6	33.6	21.6
	Smoker	21.6	25.8	18.6	28
BMI, %	**<20**	9.2	12.4	5.7	12.8
	[20-25]	46.5	48.3	47.1	53.6
	[25-30]	32.9	34.8	36.4	24.8
	>=30	11.4	4.5	10.7	8.8
Nasal symptoms, %	**Symptoms without eye symptoms**			62.9	24.8
	Symptoms with eye symptoms			37.1	75.2
Type of nasal symptoms, %	**ever but not current**			2.9	1.6
	ever and current			97.1	98.4
Report of allergic rhinitis*, %		5.3	6.7	25.7	58.4
Report of hay fever*, %		6.6	21.4	17.9	68
Report of conjonctivitis*, %		12.3	16.9	27.1	44.8
Report of sinusitis*, %		35.1	33.7	58.6	52.8
Report of eczema*, %		17.5	33.7	35.7	31.2
Diagnostic of allergy*, %		13.2	21.4	30.7	60.0
Sensitivity to hay/flowers, %	**No sensitivity**	92.5	80.9	76.4	34.4

	Rhinorrhea or sneezing	6.1	13.5	18.6	28.8
	Rhinorrhea and sneezing	1.3	5.6	5	36.8
Sensitivity to animals, %	**No sensitivity**	98.3	97.8	99.3	78.4
	Rhinorrhea or sneezing	1.3	2.2	0.7	12.8
	Rhinorrhea and sneezing	0.4	0	0	8.8
Sensitivity to dust, %	**No sensitivity**	78.9	69.7	52.9	32
	Rhinorrhea or sneezing	23.3	27	40.7	45.6
	Rhinorrhea and sneezing	0.9	3.3	6.4	22.4
Sensitivity to tobacco smoke, %	**No sensitivity**	97.8	98.9	86.3	92
	Rhinorrhea or sneezing	1.8	1.12	11.5	7.2
	Rhinorrhea and sneezing	0.4	0	2.2	0.8
Sensitivity to cold air, %	**No sensitivity**	84.7	83.2	63.6	71.2
	Rhinorrhea or sneezing	14.9	15.7	32.1	25.6
	Rhinorrhea and sneezing	0.4	1.1	4.3	3.2
Sensitivity to weather, %	**No sensitivity**	97.8	96.6	86.4	84.8
	Rhinorrhea or sneezing	1.8	1.1	10.7	12
	Rhinorrhea and sneezing	0.4	2.3	2.9	3.2
Use of nasal spray in the last 12 months*, %		23.3	22.5	45.7	48

Use of other drug in the last 12 months*, %		16.2	21.4	36.4	53.6

***= p-value<0.001, BMI: Body Mass Index**

Table D. Description of the participants with asthma according to the four classical phenotypes (hypothesis driven)

		Phenotype 1: No symptoms, no SPT (n=17)	**Phenotype 2: SPT only** (n=70)	**Phenotype 3: symptoms, no SPT** (n=55)	**Phenotype 4: Symptoms and SPT** (n=259)
age, mean± sd		47.2±14.2	38.2±17.0	45.1±17.7	35.4±15.0
Sex, women %		58.8	41.4	52.7	46
Tobacco status, %	**Non-smoker**	47	52.9	41.8	54.1
	Ex-smoker	41.2	25.7	30.9	18.2
	Smoker	11.8	21.4	27.3	27.8
BMI, %	**<20**	0	10	7.3	14.7
	[20-25]	41.2	55.7	45.5	52.1
	[25-30]	41.2	21.4	30.9	23.9
	>=30	17.6	12.9	16.4	9.3
Nasal symptoms, %	**Symptoms without eye symptoms**			40.0	21.0
	Symptoms with eye symptoms			60.0	79.0
Type of nasal symptoms, %	**ever but not current**			0	0.8
	ever and current			100	99.2
Report of allergic rhinitis*, %		0	32.9	47.3	69.5
Report of hay fever*, %		11.8	41.4	38.2	71.4
Report of conjonctivitis*, %		11.8	30	32.7	55.6
Report of sinusitis, %		41.2	47.1	60	54.1
Report of eczema*, %		41.2	42.9	25.5	56
Diagnostic of allergy*, %		35.3	64.3	52.7	85.7
BHR=1, %		57.1	67.7	71.0	71.8
FEV1 % predict		1.06±0.19	0.94±0.22	0.97±0.21	0.97±0.16
Sensitivity to hay/flowers*, %	**No sensitivity**	94.1	72.9	81.8	31.7
	Rhinorrhea or sneezing	5.9	14.3	5.5	29.7
	Rhinorrhea and sneezing	0	12.9	12.7	38.6

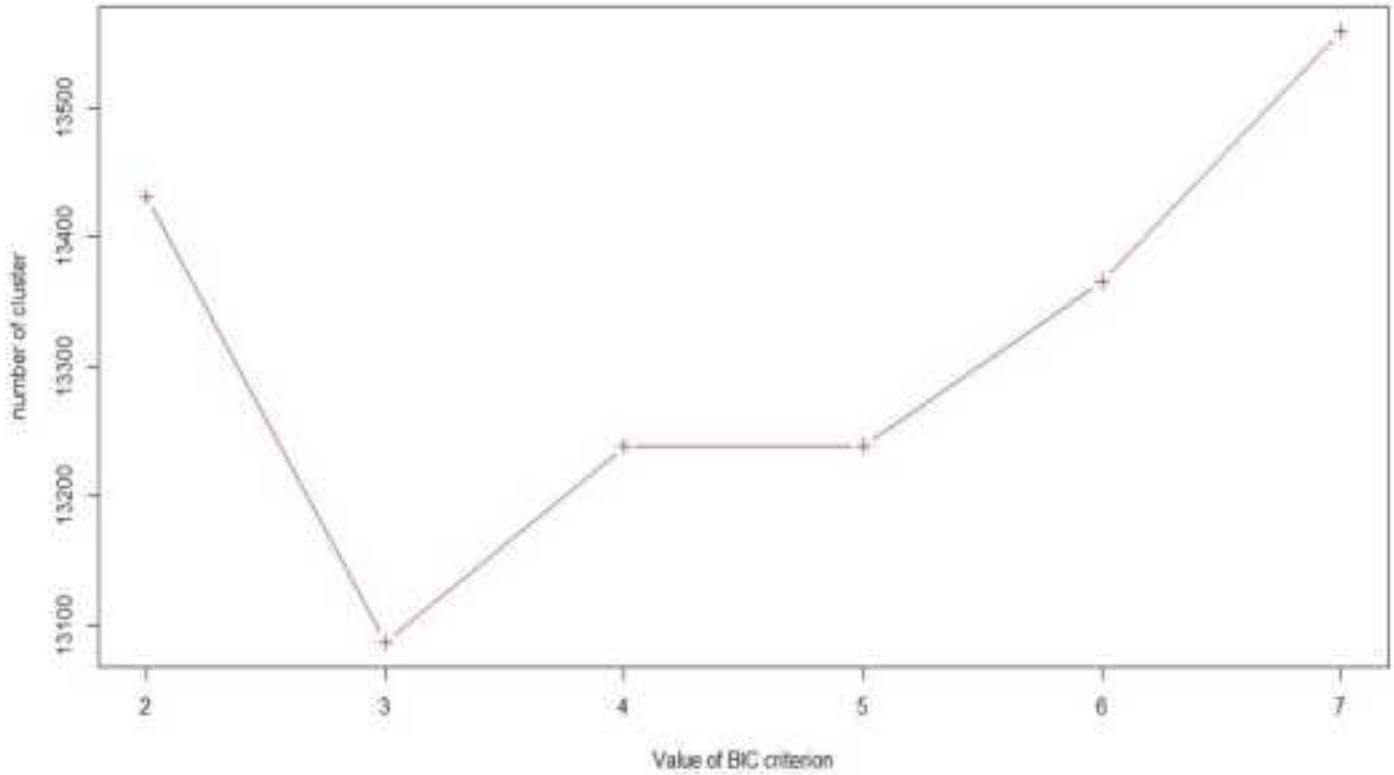
Sensitivity to animals*, %	**No sensitivity**	100	85.7	94.5	59.9
	Rhinorrhea or sneezing	0	8.6	3.6	18.5
	Rhinorrhea and sneezing	0	5.7	1.8	21.6
Sensitivity to dust*, %	**No sensitivity**	82.4	60	49.1	28.6
	Rhinorrhea or sneezing	17.6	27.1	34.6	41.7
	Rhinorrhea and sneezing	0	12.9	16.4	29.7
Sensitivity to tobacco smoke, %	**No sensitivity**	100	94.3	88.9	82.6
	Rhinorrhea or sneezing	0	5.7	7.4	12.7
	Rhinorrhea and sneezing	0	0	3.7	4.6
Sensitivity to cold air, %	**No sensitivity**	88.2	85.7	58.2	69.1
	Rhinorrhea or sneezing	11.8	14.3	38.2	26.3
	Rhinorrhea and sneezing	0	0	3.6	4.6
Sensitivity to weather, %	**No sensitivity**	94.1	94.3	85.5	76.5
	Rhinorrhea or sneezing	5.9	5.7	9.1	13.1
	Rhinorrhea and sneezing	0	0	5.5	10.4
Use of nasal spray in the last 12 months*, %		35.3	44.3	45.5	66.8
Use of other drug in the last 12 months, %		5.9	8.6	5.5	18.5

*= p-value<0.001, BMI= Body Mass Index, BHR: Bronchial HyperResponsiveness (Methacholine test, PD20≤4 mg), FEV1= Forced Expiratory Volume 1s

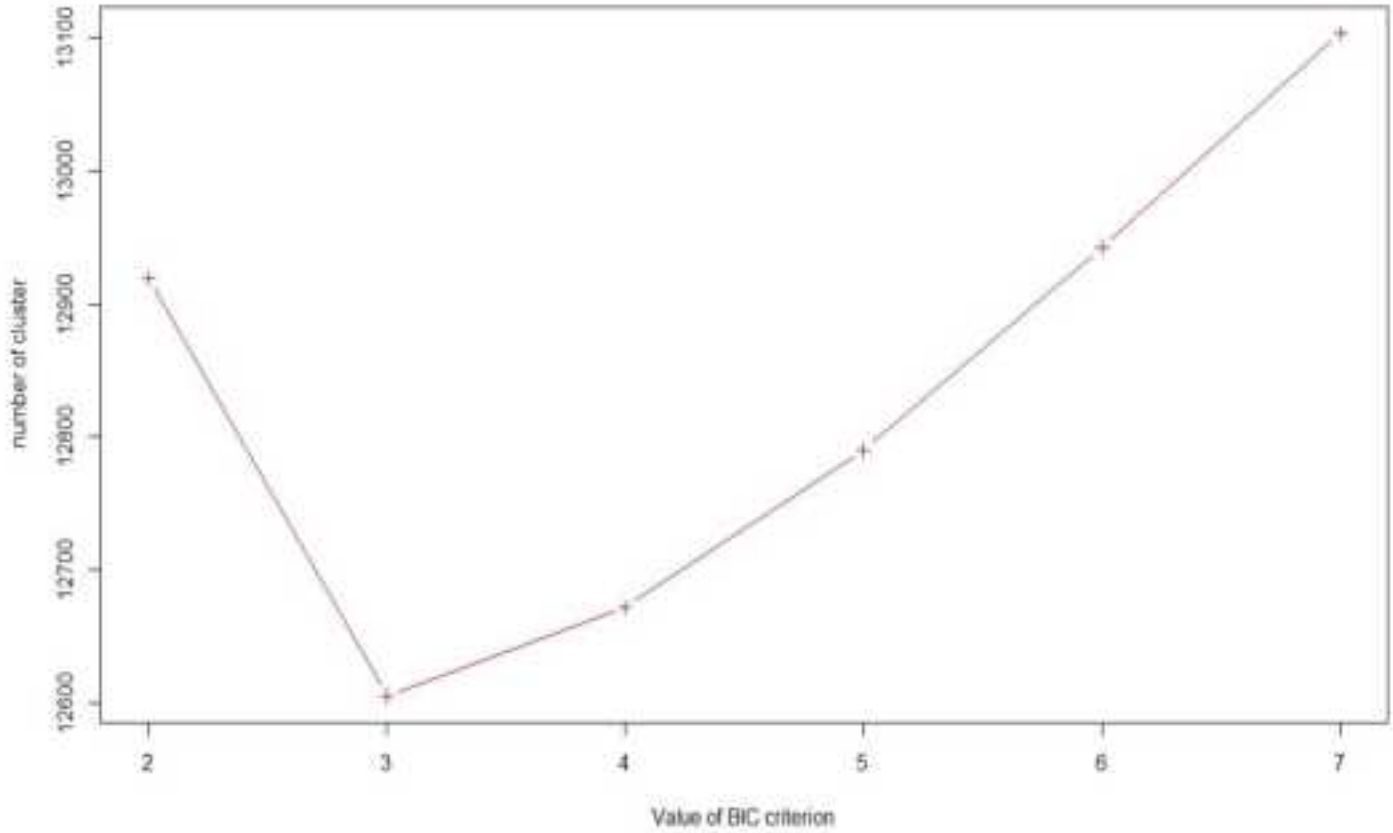
Figure legends

Fig S~~E~~1: BIC criterion according to the number of cluster for participants without (Part A) and with (Part B) asthma

Part A


number of cluster
13100
13200
13300
13400
13500
2
3
4
5
6
7
Value of BIC criterion

Part B


number of cluster
12600
12700
12800
12900
13000
13100
2
3
4
5
6
7
Value of BIC criterion