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

Emilie Burte, Jean Bousquet, Raphaëlle Varraso, Frédéric Gormand, Jocelyne Just, Régis Matran, Isabelle Pin, Valerie Siroux, Bénédicte Jacquemin, Rachel Nadif

To cite this version:

Emilie Burte, Jean Bousquet, Raphaëlle Varraso, Frédéric Gormand, Jocelyne Just, et al.. Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study. PLoS ONE, Public Library of Science, 2015, 10 (8), pp.e0136191. <10.1371/journal.pone.0136191>. <hal-01231006>

HAL Id: hal-01231006
https://hal.sorbonne-universite.fr/hal-01231006
Submitted on 19 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study

Emilie Burte1,2,*, Jean Bousquet1,2,3, Raphaëlle Varraso1,2, Frédéric Gormand4, Jocelyne Just5,6, Régis Matran7, Isabelle Pin8,9,10, Valérie Siroux8,9,10, Bénédicte Jacquemin1,2,12, Rachel Nadif1,2

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

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.

Sample population:

1571 adults

N 4 excluded for misclassification (<16 years)

1567 adults

N 584 excluded, missing values for one of the 21 variables

983 adults

60 variables related to rhinitis

Recoding variables of sensitivity and symptoms

22 variables in the analysis

Exclusion of variable with more than 95% of same answer: sensitivity to physical exercise

21 variables in the analysis

582 participants without asthma

401 participants with asthma

Selection of variables

1 variable with χ² > 0.05: report of

21 variables

20 variables

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 ≥ 3mm than the negative control for at least one of 12 aeroallergens (indoor: cat, Dermatophagoides pteronyssinus, Blattella germanica, outdoor: olive, birch, Pariereria 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 \( \chi^2 \) 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.
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 (FEV1) 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 persistent\textsubscript{high} than persistent\textsubscript{low} symptoms while cluster B’ declared more intermittent or persistent\textsubscript{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

These two components explain 38.94% of the point variability.

Part B

Participants with asthma

These two components explain 36.72% of the point variability.
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.

<table>
<thead>
<tr>
<th></th>
<th>Cluster A-no rhinitis- (n = 317)</th>
<th>Cluster B-non-allergic rhinitis- (n = 136)</th>
<th>Cluster C-allergic rhinitis- (n = 129)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± sd</td>
<td>46.7 ± 16.2</td>
<td>48.9 ± 15.5</td>
<td>40.9 ± 14.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex, women %</td>
<td>46.7</td>
<td>58.1</td>
<td>56.6</td>
<td>0.036</td>
</tr>
<tr>
<td>Tobacco, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smoker</td>
<td>47.2</td>
<td>47.0</td>
<td>51.2</td>
<td>0.90</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>30.0</td>
<td>30.2</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td>22.8</td>
<td>22.88</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m2), %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>10.1</td>
<td>8.8</td>
<td>9.3</td>
<td>0.53</td>
</tr>
<tr>
<td>[20–25]</td>
<td>47.0</td>
<td>47.8</td>
<td>52.7</td>
<td></td>
</tr>
<tr>
<td>[25–30]</td>
<td>33.4</td>
<td>36.0</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>&gt; = 30</td>
<td>9.5</td>
<td>7.4</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>Educational level, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>31.2</td>
<td>34.6</td>
<td>16.3</td>
<td>0.007</td>
</tr>
<tr>
<td>Medium</td>
<td>25.2</td>
<td>23.5</td>
<td>26.4</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>43.5</td>
<td>41.9</td>
<td>57.4</td>
<td></td>
</tr>
<tr>
<td>Nasal symptoms, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No symptoms</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Symptoms without eyes symptoms</td>
<td>0.0</td>
<td>71.3</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>Symptoms with eyes symptoms</td>
<td>0.0</td>
<td>28.7</td>
<td>82.9</td>
<td></td>
</tr>
<tr>
<td>Type of nasal symptoms, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No symptoms</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>Symptoms: ever but not current</td>
<td>0.0</td>
<td>4.4</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Ever and current symptoms</td>
<td>0.0</td>
<td>95.6</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Persistence of nasal symptoms, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermittent</td>
<td>-</td>
<td>50.0</td>
<td>40.3</td>
<td>0.22</td>
</tr>
<tr>
<td>Persistent_low</td>
<td>-</td>
<td>18.4</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>Persistent_high</td>
<td>-</td>
<td>31.6</td>
<td>34.1</td>
<td></td>
</tr>
<tr>
<td>Disturbance due to nasal symptoms, %</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No</td>
<td>-</td>
<td>77.2</td>
<td>42.6</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-</td>
<td>17.7</td>
<td>39.5</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>-</td>
<td>4.4</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-</td>
<td>0.7</td>
<td>3.90</td>
<td></td>
</tr>
<tr>
<td>Allergic sensitization,%</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SPT = 0</td>
<td>71.9</td>
<td>80.9</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>SPT = 1</td>
<td>17.0</td>
<td>10.3</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>SPT = 2</td>
<td>7.6</td>
<td>4.4</td>
<td>19.4</td>
<td></td>
</tr>
<tr>
<td>SPT &gt; 2</td>
<td>3.5</td>
<td>4.4</td>
<td>34.1</td>
<td></td>
</tr>
<tr>
<td>Report of diagnosis of allergy by a physician, %</td>
<td>15.5</td>
<td>18.4</td>
<td>72.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Immunotherapy since first survey (EGEA1)</td>
<td>1.60</td>
<td>0.0</td>
<td>13.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age of onset of nasal symptoms, mean ± sd (n = 224)</td>
<td>-</td>
<td>33.7 ± 18.2</td>
<td>22.1 ± 14.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Report of allergic rhinitis, %</td>
<td>5.70</td>
<td>6.6</td>
<td>77.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Report of hay fever, %</td>
<td>10.7</td>
<td>10.3</td>
<td>74.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Report of conjunctivitis, %</td>
<td>13.4</td>
<td>22.1</td>
<td>49.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Report of sinusitis, %</td>
<td>34.7</td>
<td>55.1</td>
<td>56.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Report of eczema, %</td>
<td>22.1</td>
<td>30.9</td>
<td>36.4</td>
<td>0.005</td>
</tr>
<tr>
<td>Sensitivity to hay/flowers, %</td>
<td>No sensitivity</td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>8.2</td>
<td>16.9</td>
<td>30.2</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
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>
<thead>
<tr>
<th>Table 2. (Continued)</th>
<th>Cluster A-no rhinitis- (n = 317)</th>
<th>Cluster B-non-allergic rhinitis- (n = 136)</th>
<th>Cluster C-allergic rhinitis- (n = 129)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity to animals, %</td>
<td>Rhinorrhea and sneezing</td>
<td>2.5</td>
<td>0.7</td>
<td>40.3</td>
</tr>
<tr>
<td>No sensitivity</td>
<td>98.1</td>
<td>100</td>
<td>78.3</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>1.60</td>
<td>0.0</td>
<td>13.2</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea and sneezing</td>
<td>0.30</td>
<td>0.0</td>
<td>8.50</td>
<td></td>
</tr>
<tr>
<td>Sensitivity to dust, %</td>
<td>No sensitivity</td>
<td>74.1</td>
<td>58.1</td>
<td>27.1</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>24.3</td>
<td>39.7</td>
<td>46.5</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea and sneezing</td>
<td>1.60</td>
<td>2.20</td>
<td>26.4</td>
<td></td>
</tr>
<tr>
<td>Sensitivity to tobacco smoke, %</td>
<td>No sensitivity</td>
<td>98.1</td>
<td>90.4</td>
<td>87.5</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>1.60</td>
<td>7.40</td>
<td>11.7</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea and sneezing</td>
<td>0.30</td>
<td>2.20</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Sensitivity to cold air, %</td>
<td>No sensitivity</td>
<td>84.2</td>
<td>66.9</td>
<td>67.4</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>15.1</td>
<td>30.2</td>
<td>27.9</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea and sneezing</td>
<td>0.60</td>
<td>2.90</td>
<td>4.70</td>
<td></td>
</tr>
<tr>
<td>Sensitivity to weather, %</td>
<td>No sensitivity</td>
<td>97.5</td>
<td>88.2</td>
<td>83.0</td>
</tr>
<tr>
<td>Rhinorrhea or sneezing</td>
<td>1.60</td>
<td>11.8</td>
<td>10.8</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea and sneezing</td>
<td>0.90</td>
<td>0.00</td>
<td>6.20</td>
<td></td>
</tr>
<tr>
<td>Use of nasal spray in the last 12 months, %</td>
<td></td>
<td>23.0</td>
<td>39.7</td>
<td>54.3</td>
</tr>
<tr>
<td>Use of other drug in the last 12 months, %</td>
<td></td>
<td>17.7</td>
<td>27.9</td>
<td>62.0</td>
</tr>
</tbody>
</table>

BMI = Body Mass Index
*p-value overall

doi:10.1371/journal.pone.0136191.002
Table 3. Characteristics of participants with asthma according to each cluster.

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

(Continued)
...allergic rhinitis to non-allergic rhinitis [20], [21], [22]. Overall, it is reassuring *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>
<thead>
<tr>
<th>Cluster</th>
<th>Rhinorrhea and sneezing</th>
<th>No sensitivity</th>
<th>Rhinorrhea or sneezing</th>
<th>Rhinorrhea and sneezing</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>4.6</td>
<td>6.3</td>
<td>28.2</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>64.4</td>
<td>50.7</td>
<td>16.5</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>C</td>
<td>25.3</td>
<td>37.5</td>
<td>42.9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity to tobacco smoke, %</th>
<th>No sensitivity</th>
<th>Rhinorrhea or sneezing</th>
<th>Rhinorrhea and sneezing</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>95.4</td>
<td>90.2</td>
<td>78.2</td>
<td>0.0002</td>
</tr>
<tr>
<td>B</td>
<td>1.0</td>
<td>9.1</td>
<td>14.1</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.0</td>
<td>0.7</td>
<td>4.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity to cold air, %</th>
<th>No sensitivity</th>
<th>Rhinorrhea or sneezing</th>
<th>Rhinorrhea and sneezing</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>86.2</td>
<td>71.5</td>
<td>63.5</td>
<td>0.001</td>
</tr>
<tr>
<td>B</td>
<td>13.8</td>
<td>26.4</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.0</td>
<td>2.1</td>
<td>6.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity to weather, %</th>
<th>No sensitivity</th>
<th>Rhinorrhea or sneezing</th>
<th>Rhinorrhea and sneezing</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>94.3</td>
<td>86.8</td>
<td>70.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>B</td>
<td>5.8</td>
<td>9.7</td>
<td>14.7</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.0</td>
<td>3.5</td>
<td>14.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disturbance due to nasal symptoms, %</th>
<th>No</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*</td>
<td>-</td>
<td>20.1</td>
<td>34.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>-</td>
<td>11.1</td>
<td>23.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-</td>
<td>3.5</td>
<td>12.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 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 |
| FEV1, % predicted, meantsd | 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
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 4. Comparison of the repartition of the participants without asthma into the different hypothesis-driven’s phenotypes and data-driven’s cluster.

<table>
<thead>
<tr>
<th>Hypothesis-driven Phenotypes</th>
<th>Data-driven clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A (No rhinitis)</td>
</tr>
<tr>
<td>1 (no symptoms, no SPT)</td>
<td>228 (39)</td>
</tr>
<tr>
<td>2 (no symptoms, SPT+)</td>
<td>89 (15)</td>
</tr>
<tr>
<td>3 (symptoms, no SPT – NAR)</td>
<td>0</td>
</tr>
<tr>
<td>4 (symptoms, SPT+ – AR)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>317</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0136191.t004

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

<table>
<thead>
<tr>
<th>Hypothesis-driven Phenotypes</th>
<th>Data-driven clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A’ (No rhinitis)</td>
</tr>
<tr>
<td>1 (no symptoms, no SPT)</td>
<td>17 (4)</td>
</tr>
<tr>
<td>2 (no symptoms, SPT+)</td>
<td>70 (17)</td>
</tr>
<tr>
<td>3 (symptoms, no SPT – NAR)</td>
<td>0</td>
</tr>
<tr>
<td>4 (symptoms, SPT+ – AR)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>87</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0136191.t005
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 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).


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 BJ RN.

References