SUPPLEMENTARY MATERIALS

Endotypes identified by cluster analysis in asthmatics and non-asthmatics and their clinical characteristics 10 years later: the case-control EGEA study

Rachel Nadif *et al.*

Nadif R, *et al. BMJ Open Resp Res* 2020; 7:e000632. doi: 10.1136/bmjresp-2020-000632

SUPPLEMENTARY METHODS

Study design

The EGEA is a cohort study based on an initial group of asthma cases recruited in chest clinics from Grenoble, Lyon, Marseille, Montpellier and Paris, along with their first-degree relatives, and a group of controls (1991-1995, EGEA1, https://egeanet.vjf.inserm.fr/). The protocol and descriptive characteristics have been described previously [1, 2]. Briefly, the asthma cases and their first-degree relatives were recruited from respiratory or allergic clinics. The adult cases were recruited in the five cities and children cases were recruited in Paris, Grenoble and Marseille. Control adults were recruited from electoral rolls in Paris, Lyon Montpellier and Grenoble, a check-up centre in Marseille and surgery clinic from the same hospital in Paris and Grenoble. Control children were always recruited from surgery clinics. An overall matching by month of exam, age decade, sex and centre was done.

A 12-year follow-up of the initial cohort was conducted between 2003 and 2007 (EGEA2). Among the alive cohort (n=2002), 92% (n=1845) completed a short self-administered questionnaire, and among them 1602 (n=1571 adults aged ≥16 years) had a complete examination. As a follow-up study of EGEA2, the third survey (EGEA3, 2011-2013, n=1558) was conducted using self-completed questionnaire only**.**

The EGEA collection was certified ISO 9001 from 2006 to 2018 [3]. Ethical approval was obtained from the relevant institutional review board committees (Cochin Port-Royal Hospital and Necker-Enfants Malades Hospital, Paris). All participants signed a written informed consent.

Asthma outcomes at EGEA2 and EGEA3

At EGEA2, the participants with **ever asthma** were those who answered positively to at least one of the two following questions: "*Have you ever had attacks of breathlessness at rest with wheezing*?" or "*Have you ever had asthma attacks?*", or were recruited as asthmatic cases at EGEA1.

Never-asthmatics were those who answered negatively to the two questions above, and were not recruited as asthmatic cases at EGEA1. They were recruited as family members, spouses or control, could report respiratory symptoms but did not fulfil the proposed strict criteria to define asthma. None of them had answered positively to the question "*Have you been (treated/monitored) for emphysema?*" and 2% positively answered to the question "*Have you been (treated/monitored) for chronic bronchitis*?" at EGEA1.

Among participants with ever asthma, **current asthma** was defined by the report of respiratory symptoms (wheeze, nocturnal chest tightness, or attacks of breathlessness following strenuous activity, at rest or at night-time) or asthma attacks or use of inhaled and/or oral medicines because of breathing problems in the past twelve months.

Ranging from 0 to 5 [4, 5], **the asthma symptom score** is based on the number of respiratory symptoms during the past 12 months reported in the questionnaire: (1) breathless while wheezing; (2) woken up with chest tightness; (3) attack of shortness of breath at rest; (4) attack of shortness of breath after exercise; and (5) woken by attack of shortness of breath.

Asthma control at EGEA2 and EGEA3 has been assessed over 3 month period, using responses to survey questions to approximate the Global Initiative for Asthma 2015 definition as closely as possible and as previously used [6, 7]. Participants were defined as having controlled, partly controlled, and uncontrolled asthma if they had none, 1 to 2, or 3 to 4 of the

following criteria, respectively: frequent daytime symptoms (defined by ≥1 asthma attacks or ≥1 episodes of trouble breathing per week in the past 3 months), any night-time symptoms (defined as waking because of asthma or an attack of shortness of breath in the last 3 months), frequent use of reliever medication (defined, on average, as more than twice a week in the past 3 months), and any activity limitation (defined by the following answers: ‘‘totally limited,’’ ‘‘extremely limited,’’ ‘‘very limited,’’ ‘‘moderate limitation,’’ and ‘‘some limitation’’ to the question ‘‘*Overall, among all the activities that you have done during the last two weeks, how limited have you been by your asthma?*’’).

Asthma outcomes at EGEA3

The Asthma Control Test (ACT) is a tool developed by Nathan and collaborators for evaluating asthma control in the last 4 weeks [8]. It consists of five questions on the impact of asthma on everyday functioning, shortness of breath, asthma symptoms, use of rescue medications, and asthma control evaluation with five possible answers each. A maximum score of 25 points indicates complete asthma control. A score between 20 and 25 represents well controlled asthma, while a score of 19 or below represents not well controlled asthma.

Asthma exacerbations were defined by hospital or emergency admissions for asthma or use of oral steroids for breathing difficulties in the past twelve months [6].

Lung function, allergic and inflammatory characteristics (EGEA2)

A lung function test with spirometry and methacholine challenge was performed using standardized protocol with similar equipment across centres according to the American Thoracic Society / European Respiratory Society guidelines [9]. Forced expiratory volume in one second (FEV_1) percent predicted value was based on Quanjer et *al*. reference equations

[10]. For participants with a FEV1 ≥80% of the predicted value, a methacholine bronchial challenge test was performed (maximum dose 4mg) using a Biomedin spirometer (Biomedin Srl, Padua, Italy) in all centres, except in Lyon, where a Pneumotach Jaeger spirometer (Jaeger) was used.

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*, *Blattela germanica*, outdoor: olive, birch, *Parieteria judaica*, timothy grass, *Cupressu*s and ragweed pollen, and moulds: Aspergillus, *Cladosporium herbarum*, *Alternaria tenuis*). Subjects were classified as sensitized if they have one or more SPT+ [11].

Total IgE and blood cell counts (EGEA2)

Total Immunoglobulin E (IgE) determination was assessed by UniCAP system (Pharmacia®) from blood samples in a centralized laboratory, and expressed in international units (IU) per millilitre.

Blood and exhaled breath condensate measurements (EGEA2)

Serums, plasma and exhaled breath condensates were prepared and stored at -80°C according to standardized procedures (http://www.afaq.org/certification=262711141114) until measurements [3].

Measurement of plasma Fluorescent Oxidation Products (FlOPs) level (EGEA2)

Plasma FlOPs level was measured as previously described [12]. Briefly, plasma was extracted into a mixture of ethanol/ether (3/1 v/v) and measured using a spectrofluorimeter (360 nm excitation wavelength, 430 nm emission wavelength). Fluorescence was expressed as a unit of relative fluorescence intensity (RFU).mL^{-1} of plasma.

Measurement of 8-isoprostane level in exhaled breath condensate (EGEA2)

Exhaled breath condensate (EBC) was collected using an RTube according a standardised method as previously described [13]. Briefly, the RTube (TM) was rinsed with deionized water and dried thoroughly. Participants breathed orally at tidal volumes into a mouthpiece attached to a cold condenser (-20°C). They were seated comfortably with a headrest. All headrests and back seats were tilted slightly to avoid any saliva contamination during breathing manoeuvres. Breathing was quiet and regular. After 15 minutes, EBC collection was immediately separated in aliquots and stored at -80·C.

8-isoprostane levels were measured in EBC at the Laboratory of Biochemistry Molecular Biology of CHRU de Lille using a specific enzyme immunoassay kit (8-isoprostanes EIA kit; Cayman Chemical, Ann Arbor, MI, USA) according to the manufacturer's protocol. 50 µL of non-extracted EBC was assayed in duplicate and the 8-isoprostane levels were calculated from a calibration curve obtained from the eight calibration points (0.8–2.0–5.1–12.8–32–80–200–500 pg·mL−1, where 0.8 pg·mL−1 is the lowest point). The lower limit of detection for 8-isoprostanes was 4.0 pg·mL−1 and the intra-assay coefficient of variation was <20%. The measurements were available for 329 CA and 211 NA participants included in the present study.

Measurement of cytokine level in serum (EGEA2)

An inflammation cytokine 9-plex assay was used. This panel included Interleukin(IL)-1Ra, IL-5, IL-6, IL-7, IL-8, IL-10, IL-13, TNF-a and Leptin (Bio-Plex Pro human inflammation cytokine 9-plex assay, Bio-Rad Laboratories, Hercules, California, United States). Determination of serum cytokine levels (at EGEA2) was carried out at the IMIM (Instituto Hospital del Mar de Investigaciones Medicas, Barcelona, Spain) using the Luminex xMAG®

technique in a BioPlex system. Around 50% of the measurements of IL-6 were undetectable, and were not showed. No value above or below the limit of detection (LD) was observed for IL-7 and IL-13 measurements. Less than 3% of the values were above or below the LD for other cytokines. Values below the lower limit of detection (LLD) were assigned the LLD/2 and values above the upper limit of detection (ULD) were assigned the ULD. An aliquot of the same human serum pool was run by duplicate in each plate. In order to minimize inter-plate variability, each serum cytokine value was divided by the pool result obtained from the same plate and expressed as ratio. These ratios were included in the statistical analyses. The measurements were available for 344 CA (86%) and 358 NA (54%) participants included in the present study. For more details see the paper by Akiki *et al.* [14].

Measurement of high sensitivity C-reactive protein (EGEA2)

High sensitivity CRP was measured on a routine clinical chemistry analyser, Synchron, by immunoturbidimetry (Beckman Coulter, Inc.) at the Catholic University of Louvain. Assays were performed using a near-infrared particle immunoassay, with a laser diode at 940 nm. The standard curve was adjusted daily by adding a calibrator. Almost all the measurements of hs-CRP in serum are expected to be below 300 mg·mL−1 with the highest values around 750 or 1000 mg·mL−1. The measurements were available for 612 CA and 372 NA participants included in the present study.

SUPPLEMENTARY RESULTS

The distributions of the clusters according to the centres are provided in Table E6 and E7.

Among asthmatics, the distribution of the CA clusters varies by centre ($p<0.001$), reflecting the study design. Indeed, Marseille, Paris and Grenoble have a higher contribution to the cluster CA3, characterized by childhood-onset asthma, consistent with the fact that these centres have recruited children at EGEA1. Among never-asthmatics, although NA1 was the most prevalent cluster in each centre, a significant difference between centres was observed ($p=0.002$), with more participants assigned to cluster NA2 in Marseille and Montpellier than in the 3 other centres. Noteworthy, all centres have used standardized questionnaires and protocols. The variables that characterized each cluster were not differently recorded according to the centre.

To test the robustness of our results, we used the same mixture approach to identify the endotypes among participants with complete data (318 CA and 545 NA). Briefly, the 318 CA did not differ from the 339 CA on their 28 characteristics (to compare see Table 2 in the main text and supplementary Table E8), and the 545 NA did not differ from the 666 NA on their 23 characteristics (to compare see Table 3 in the main text and supplementary Table E9). Three endotypes were identified among asthmatics and two among never-asthmatics (supplementary Tables E8 and E9) showing similar characteristics as those identified among the participants with imputed data.

REFERENCES

1. Kauffmann F, Dizier MH. EGEA (Epidemiological study on the Genetics and Environment of Asthma, bronchial hyperresponsiveness and atopy)--design issues. EGEA Co-operative Group. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 1995; 25 Suppl 2: 19-22.
2. Kauffmann F, Dizier MH, Annesi-Maesano I, et al. EGEA (Epidemiological study on the Genetics and Environment of Asthma, bronchial hyperresponsiveness and atopy)--descriptive characteristics. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 1999; 29 Suppl 4: 17-21.
3. Nadif R, Bouzigon E, Le Moual, N, et al. EGEA Collection: A biobank devoted to asthma and asthma-related phenotypes. *Open Journal of Bioresources* 2017; 4(5): 1-6.
4. Sunyer J, Pekkanen J, Garcia-Esteban R, et al. Asthma score: predictive ability and risk factors. *Allergy* 2007; 62(2): 142-8.
5. Pekkanen J, Sunyer J, Anto JM, Burney P. Operational definitions of asthma in studies on its aetiology. *The European respiratory journal* 2005; 26(1): 28-35.
6. Siroux V, Boudier A, Bousquet J, et al. Phenotypic determinants of uncontrolled asthma. *Journal of Allergy and Clinical Immunology* 2009; 124(4): 681-7.e3.
7. Siroux V, Boudier A, Dolgopoloff M, et al. Forced mid expiratory flow between 25% and 75% of forced vital capacity is associated with long-term persistence of asthma and poor asthma outcomes. *The Journal of allergy and clinical immunology* 2016; 137(6): 1709-16 e6.
8. Nathan RA, Sorkness CA, Kosinski M, et al. Development of the asthma control test: a survey for assessing asthma control. *The Journal of allergy and clinical immunology* 2004; 113(1): 59-65.
9. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *The European respiratory journal* 2005; 26(2): 319-38.
10. Quanjer PH, Tammeling GJ, Cotes JE, et al. Lung volumes and forced ventilatory flows. *The European respiratory journal* 1993; 6 Suppl 16: 5-40.
11. Burte E, Bousquet J, Varraso R, et al. Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study. *PloS one* 2015; 10(8): e0136191.
12. Andrianjafimasy M, Zerimech F, Akiki Z, et al. Oxidative stress biomarkers and asthma characteristics in adults of the EGEA study. *The European respiratory journal* 2017; 50(6).
13. Rava M, Varraso R, Decoster B, et al. Plasma and exhaled breath condensate nitrite-nitrate level in relation to environmental exposures in adults in the EGEA study. *Nitric Oxide* 2012; 27(3): 169-75.
14. Akiki Z, Rava M, Diaz Gil O, et al. Serum cytokine profiles as predictors of asthma control in adults from the EGEA study. Respiratory Medicine 2017; 125:57-64.

Supplementary Table E1. List of the characteristics included in the cluster analyses

	Among never asthmatics (NA, n=666)	**Among current asthmatics (CA, n=402)**
Age, year *	x	x
Sex, women	x	x
Smoking habits (non-smokers, ex-smokers, smokers)	x	x
Body Mass Index, kg/m^2 *	x	x
Age of asthma onset *	/	x
FEV_1 % predicted *	x	x
FVC % predicted *	x	x
Shortness of breath and wheezing (last 12 months)	x	x
Asthma attacks (last 12 months)	/	x
Attacks of breathlessness at rest (last 12 months)	x	x
Attacks of breathlessness following strenuous activity (last 12 months)	x	x
Nocturnal symptoms (last 12 months)		
Cough	x	x
Chest tightness	x	x
Shortness of breath	x	x
Hospital admissions for asthma, (last 12 months)	/	x
Emergency admissions for asthma, (last 12 months)	/	x
Inhaled corticosteroids (last 12 months)	x	x
Oral corticosteroids (last 12 months)	/	x
Chronic cough	x	x
Chronic phlegm	x	x
Dyspnoea grade 3	x	x
Skin prick test positivity	x	x
Rhinitis (last 12 months)	x	x
Eczema (ever)	x	x
Total IgE, IU/ml *	x	x
White blood cell counts, *		
Eosinophils/mm^3	x	x
Neutrophils/mm^3	x	x
Fluorescent oxidation products, RFU/mL, *	x	x

x: included and /: not included in the cluster analysis. * Included as continuous variables in the cluster analysis. FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity. Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale. Imputed data were smoking in never-

asthmatics, rhinitis in asthmatics, and ever eczema, ICS use, allergic sensitization and total IgE level in both groups.

Supplementary Table E2. Comparison of the 28 characteristics between the 1068 participants included and the 378 not included in the analyses

	Included (n=1068)	**Not included (n=378)**	**P-value**
Age, year	43.9 (16.4)	42.5 (16.7)	0.15
Sex, women	570 (53.4%)	174/378 (46.0%)	0.01
Smoking habits		*n=373*	
non-smokers	532 (49.8%)	178 (47.7%)	
ex-smokers	299 (28.0%)	96 (25.7%)	0.22
smokers	237 (22.2%)	99 (26.6%)	
		n=215	
Body Mass Index, kg/m^2	24.5 (4.2)	24.7 (5.1)	0.5
Asthma ever	402 (37.6%)	156/378 (41.3%)	0.2
Age of asthma onset	*n=402*	*n=123*	0.9
	15.1 (14.9)	15.2 (15.0)	
		n=210	
FEV_1 % predicted	102 (18.4)	103 (17.9)	0.8
FVC % predicted	110 (17.4)	109 (16.9)	0.2
Shortness of breath and wheezing (last 12 months)	280 (26.2%)	95/375 (25.3%)	0.7
Asthma attacks (last 12 months)	183/417 (43.9%)	70/159 (44.0%)	0.9
Attacks of breathlessness at rest (last 12 months)	133 (12.4%)	49/368 (13.3%)	0.7
Attacks of breathlessness following strenuous activity (last 12 months)	435 (40.7%)	165/369 (44.7%)	0.18
Nocturnal symptoms (last 12 months)			
Cough	341 (31.9%)	125/374 (33.4%)	0.6
Chest tightness	303 (28.4%)	99/370 (26.8%)	0.5
Shortness of breath	105 (9.8%)	39/369 (10.6%)	0.7
Hospital admissions (last 12 months)	4/417 (0.96%)	0/153 (0.0%)	0.2
Emergency admissions (last 12 months)	9/417 (2.16%)	1/155 (0.65%)	0.2
Inhaled corticosteroids (last 12 months)	218 (20.5%)	76/324 (23.5%)	0.2
Oral corticosteroids (last 12 months)	63/417 (15.1%)	28/153 (18.3%)	0.3
Chronic cough	84 (7.9%)	45/331 (13.6%)	0.002
Chronic phlegm	73 (6.8%)	38/331 (11.5%)	0.006
Dyspnoea grade 3	114 (10.7%)	41/367 (11.2%)	0.8
Skin prick test positivity	549 (54.2%)	115/189 (60.8%)	0.09
Rhinitis (last 12 months)	371 (34.9%)	129/362 (35.6%)	0.8
Eczema (ever)	383 (36.0%)	114 (30.5%)	0.05
		n=225	
Total IgE, IU/ml, GM (Q1 – Q3)	70.5 (23.8 – 206)	88.1 (29.6 – 221)	0.05
White blood cell counts		*n=218*	

Eosinophils/mm^3	200 (162)	230 (163)	0.01
Neutrophils/mm^3	4015 (1405)	4085 (1370)	0.5
Fluorescent oxidation products, RFU/mL, GM (Q1 – Q3)	93.8 (80.4 – 107)	*n=156* 93.6 (83.0 – 102)	0.9

Data are means (SD) or n (%) unless otherwise stated. FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity. GM= geometric mean, (Q1 – Q3) = first and third quartiles. Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale.

Supplementary Table E3. Description of the characteristics not included in the cluster analysis at EGEA2 according to each CA endotype

	CA 1 (n=53)	**CA 2 (n=219)**	**CA 3 (n=130)**	**P-value**
Asthma symptom score, last 12 months (0-5)	3.22 (1.25)	2.57 (1.31)	1.47 (1.17)	<0.0001
0	0 (0%)	8 (3.6%)	28 (21.5%)	
1 or 2	18 (34.0%)	105 (47.9%)	81 (62.3%)	<0.0001
More than 2	35 (66.0%)	106 (48.4%)	21 (16.2%)	
Asthma control	*n=49*	*n=207*	*n=124*	
Controlled	10 (20.4%)	98 (47.3%)	95 (76.6%)	
Partly controlled	22 (44.9%)	88 (42.5%)	24 (19.4%)	<0.0001
Uncontrolled	17 (34.7%)	21 (10.1%)	5 (4.0%)	
FEV_1 % < 80% predicted	28 (52.8%)	32 (14.6%)	15 (11.5%)	<0.0001
Methacholine challenge,* PD 20 ≤ 4 mg	6/12 (50.0%)	105/135 (77.8%)	58/91 (63.7%)	0.02

Data are means (SD) or n (%) unless otherwise stated.

FEV_1= forced expiratory volume in 1 s.

*Methacholine challenge test was not performed if baseline FEV_1 <80% predicted.

Supplementary Table E4. Overlaps between the three CA endotypes and the four phenotypes previously identified by latent class analysis at EGEA2 (Siroux *et al.* [18])

	CA1 (n=53)	**CA2 (n=219)**	**CA3 (n=130)**
Phenotype E: "*active treated allergic childhood-onset asthma*"(n=222)	18 (34.0%)	109 (49.8%)	38 (29.2%)
Phenotype F: "*active treated adult-onset asthma*" (n=96)	30 (56.6%)	37 (16.9%)	6 (4.6%)
Phenotype G: "*inactive/mild untreated allergic childhood asthma*"(n=159)	1 (1.9%)	17 (7.7%)	62 (47.7%)
Phenotype H: "*inactive / mild untreated adult onset asthma*"(n=164)	4 (7.5%)	56 (25.6%)	24 (18.5%)

Supplementary Table E5. Description of the characteristics not included in the cluster analysis at EGEA2 according to each NA endotype

	NA1 (n=489)	**NA2 (n=177)**	**P-value**
Asthma symptom score, last 12 months (0-5)	0.33 (0.59)	0.94 (1.06)	<0.0001
0	355 (72.6%)	81 (45.8%)	
1 or 2	132 (27.0%)	82 (46.3%)	<0.0001
More than 2	2 (0.40%)	14 (7.90%)	
FEV_1 % < 80% predicted	20 (4.09%)	18 (10.2%)	0.004
Methacholine challenge*			
PD 20 ≤ 4 mg	79/339 (23.3%)	38/94 (40.4%)	0.0009

Data are means (SD) or n (%) unless otherwise stated.
FEV_1= forced expiratory volume in 1 s.
*Methacholine challenge test was not performed if baseline FEV_1 <80% predicted.

Supplementary Table E6. Distributions of CA clusters according to the centres.

	CA cluster			
Centre	**CA1**	**CA2**	**CA3**	Total
Lyon	15 23.1	47 72.3	3 4.6	65
Marseille	11 15.5	37 52.1	23 32.4	71
Montpellier	6 20.7	14 48.3	9 31.0	29
Paris	8 6.7	53 44.5	58 48.8	119
Grenoble	13 11.0	68 57.6	37 31.3	118
Total	53	219	130	402

Supplementary Table E7. Distributions of NA clusters according to the centres.

	NA cluster		
Centre	**NA1**	**NA2**	Total
Lyon	103 77.4	30 22.6	133
Marseille	76 59.8	51 40.2	127
Montpellier	33 68.7	15 31.3	48
Paris	122 76.2	38 23.8	160
Grenoble	155 78.3	43 21.7	198
Total	489	177	666

Supplementary Table E8. Description of the 28 characteristics among the current asthmatics with complete data and according to each endotype

	All (n=318)	**CA1 (n=43)**	**CA2 (n=178)**	**CA3 (n=97)**	**P-value**
Age, year	38.7 (16.5)	55.5 (14.5)	40.74 (15.5)	27.5 (10.3)	<0.0001
Sex, women	158 (49.7%)	24 (55.8%)	101 (56.7%)	112 (34.0%)	0.001
Smoking habits					
non-smokers	158 (49.7%)	20 (46.5%)	80 (44.9%)	58 (59.8%)	
ex-smokers	81 (25.5%)	17 (39.5%)	58 (32.6%)	6 (6.2%)	<0.0001
smokers	79 (24.8%)	6 (14.0%)	40 (22.5%)	33 (34.0%)	
Body Mass Index (BMI), kg/m^2	24.4 (4.34)	27.8 (5.33)	24.8 (4.25)	22.13 (2.41)	<0.0001
Age of asthma onset, %	15.0 (15.1)	30.1 (16.9)	15.0 (14.6)	8.36 (9.30)	<0.0001
FEV_1 % predicted	96 (18.0)	73 (20.0)	99 (16.0)	100 (12.0)	<0.0001
FVC % predicted,	108 (16.0)	93 (17.0)	112 (15.0)	107 (13.0)	<0.003
Shortness of breath and wheezing, (last 12 months)	190 (59.7%)	37 (86.1%)	113 (63.5%)	40 (41.2%)	<0.0001
Asthma attacks in the last 12 months	143 (45.0%)	33 (76.7%)	90 (50.6%)	20 (20.6%)	<0.0001
Attacks of breathlessness at rest, (last 12 months)	92 (28.9%)	22 (51.2%)	60 (33.7%)	10 (10.3%)	<0.0001
Attacks of breathlessness following strenuous activity, (last 12 months)	194 (61.0%)	30 (69.8%)	112 (62.9%)	52 (53.6%)	0.14
Nocturnal symptoms (last 12 months)					
Cough	131 (41.2%)	26 (60.5%)	88 (49.4%)	17 (17.5%)	<0.0001
Chest tightness	171 (53.8%)	32 (74.4%)	121 (68.0%)	18 (18.6%)	<0.0001
Shortness of breath	78 (24.5%)	20 (46.5%)	56 (31.5%)	2 (2.1%)	<0.0001
Hospital admissions, (last 12 months)	3 (0.90%)	3 (7.0%)	0 (0%)	0 (0%)	<0.0001
Emergency admissions, (last 12 months)	7 (2.20%)	7 (16.3%)	0 (0%)	0 (0%)	<0.0001
Inhaled corticosteroids (last 12 months)	150 (47.2%)	36 (83.7%)	87 (48.9%)	27 (27.8%)	<0.0001
Oral corticosteroids (last 12 months)	48 (15.1%)	17 (39.5%)	28 (17.5%)	3 (3.1%)	<0.0001

Chronic cough,	39 (12.3%)	19 (44.2%)	19 (10.7%)	1 (1.00%)	<0.0001
Chronic phlegm,	35 (11.0%)	15 (34.9%)	18 (10.1%)	2 (2.10%)	<0.0001
Dyspnoea grade 3	53 (16.7%)	27 (62.8%)	22 (12.4%)	4 (4.1%)	<0.0001
Skin prick test positivity	259 (81.4%)	27 (62.8%)	139 (78.1%)	93 (95.9%)	<0.0001
Rhinitis (last 12 months)	193 (60.7%)	23 (53.5%)	12 (68.0%)	49 (50.5%)	0.01
Eczema (ever)	158 (49.7%)	16 (37.2%)	86 (48.3%)	56 (57.7%)	0.069
Total IgE, IU/ml, GM (Q1 – Q3)	148 (58.9 – 371)	144 (64.6 – 398)	117 (44.7 – 309)	224 (102 – 525)	0.01
White blood cell counts					
Eosinophils/mm^3	255 (193)	298 (336)	249 (165)	248 (149)	0.24
Neutrophils/mm^3	4007 (1400)	5024 (1636)	4057 (1352)	3466 (1080)	<0.0001
Fluorescent oxidation products, RFU/mL, GM (Q1 – Q3)	91.2 (79.4 – 102)	107 (93.3 – 117)	91.2 (79.4 – 105)	85.1 (74.1 – 95.5)	<0.0001

Data are means (SD) or n (%) unless otherwise stated.

FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity.

GM= geometric mean, (Q1 – Q3) = first and third quartiles.

Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale.

Supplementary Table E9. Description of the 23 characteristics among the never-asthmatics with complete data and according to each endotype

	All (n=545)	**NA1 (n=439)**	**NA2 (n=106)**	**P-value**
Age, year	46.3 (15.8)	46.11 (15.5)	47.29 (17.1)	0.52
Sex, women	302 (55.4%)	232 (52.8%)	70 (66.0%)	0.02
Smoking habits				
non-smokers	270 (49.5%)	223 (50.8%)	47 (44.3%)	
ex-smokers	168 (30.8%)	139 (31.7%)	29 (27.4%)	0.04
smokers	107 (19.6%)	77 (17.5%)	30 (28.3%)	
Body Mass Index (BMI), kg/m^2	24.6 (4.05)	24.34 (3.82)	25.66 (4.85)	0.009
FEV_1 % predicted	107 (16.0)	108 (16.0)	104 (18.0)	0.05
FVC % predicted	113 (18.0)	113 (18.0)	110 (19.0)	0.16
Shortness of breath and wheezing, (last 12 months)	29 (5.30%)	0 (0%)	29 (27.4%)	<0.0001
Attacks of breathlessness at rest, (last 12 months)	14 (2.60%)	0 (0%)	14 (13.2%)	<0.0001
Attacks of breathlessness following strenuous activity, (last 12 months)	145 (26.6%)	91 (20.7%)	54 (50.9%)	<0.0001
Nocturnal symptoms (last 12 months)				
Cough	142 (26.1%)	91 (20.7%)	51 (48.1%)	<0.0001
Chest tightness	65 (11.9%)	42 (9.60%)	23 (21.7%)	0.001
Shortness of breath	6 (1.1%)	0 (0%)	6 (5.7%)	<0.0001
Inhaled corticosteroids (last 12 months)	17 (3.10%)	0 (0%)	17 (16.0%)	<0.0001
Chronic cough	26 (4.80%)	0 (0%)	26 (24.5%)	<0.0001
Chronic phlegm	23 (4.20%)	0 (0%)	23 (21.7%)	<0.0001
Dyspnoea grade 3	32 (5.9%)	1 (0.20%)	31 (29.2%)	<0.0001
Skin prick test positivity	211 (38.7%)	163 (37.1%)	48 (45.3%)	0.15
Rhinitis, (last 12 months)	112 (20.6%)	77 (17.5%)	35 (33.0%)	0.0007

Eczema (ever)	149 (27.3%)	121 (27.6%)	28 (26.4%)	0.9
Total IgE, IU/ml, GM (Q1 – Q3)	45.7 (17.4 – 115)	43.6 (16.6 – 117)	60.2 (21.9 – 151)	0.03
White blood cell counts				
Eosinophils/mm^3	161 (118)	154 (108)	189 (149)	0.03
Neutrophils/mm^3	3937 (1287)	3890 (1237)	4133 (1468)	0.11
Fluorescent oxidation products, RFU/mL, GM (Q1 – Q3)	93.3 (81.3 – 107)	93.3 (81.3 – 107)	93.3 (81.3 – 105)	1.0

Data are means (SD) or n (%) unless otherwise stated.

FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity.

GM= geometric mean, (Q1 – Q3) = first and third quartiles.

Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale.

Supplementary Table E10. Comparison of the 28 characteristics between the 341 participants with current asthma followed-up at EGEA3 and the 61 not followed-up

	Followed-up (n=341)	**Not followed-up (n=61)**	**P-value**
Age, year	39.2 (16.3)	39.9 (17.4)	0.7
Sex, women	176 (51.6%)	23 (37.3%)	0.04
Smoking habits			
non-smokers	178 (52.2%)	23 (37.3%)	
ex-smokers	84 (24.6%)	20 (32.8%)	0.11
smokers	79 (23.2%)	18 (29.5%)	
Body Mass Index, kg/m^2	24.2 (4.4)	25.0 (4.7)	0.2
Age of asthma onset	14.8 (14.7)	17.0 (15.7)	0.3
FEV_1 % predicted	95.1 (17.9)	90.9 (22.8)	0.12
FVC % predicted	107 (16.4)	103 (16.4)	0.06
Shortness of breath and wheezing, (last 12 months)	213 (62.5%)	29 (47.5%)	0.03
Asthma attacks (last 12 months)	155 (45.4%)	27 (44.3%)	0.8
Attacks of breathlessness at rest, (last 12 months)	93 (93%)	22 (36.1%)	0.16
Attacks of breathlessness following strenuous activity, (last 12 months)	212 (62.2%)	43 (70.5%)	0.2
Nocturnal symptoms (last 12 months)			
Cough	140 (41.1%)	24 (39.3%)	0.8
Chest tightness	84 (24.6%)	11 (18.0%)	0.3
Shortness of breath	185 (54.2%)	35 (57.4%)	0.6
Hospital admissions (last 12 months)	2 (0.59%)	2 (3.28%)	0.05
Emergency admissions (last 12 months)	1 (1.64%)	8 (2.35%)	0.7
Inhaled corticosteroids (last 12 months)	174 (51%)	26 (42.6%)	0.2
Oral corticosteroids (last 12 months)	54 (15.8%)	8 (13.1%)	0.6
Chronic cough	42 (12.3%)	9 (14.7%)	0.6
Chronic phlegm	36 (10.6%)	9 (14.7%)	0.3
Dyspnoea grade 3	65 (19.1%)	9 (14.7%)	0.4
Skin prick test positivity	283 (83%)	50 (82%)	0.8
Rhinitis (last 12 months)	212 (62.2%)	31 (50.8%)	0.09
Eczema (ever)	172 (50.4%)	33 (54.1)	0.6
Total IgE, IU/ml, GM (Q1 – Q3)	150 (65.4 – 390)	158 (73.3 – 332)	0.8
White blood cell counts			
Eosinophils/mm^3	263 (201)	258 (212)	0.9
Neutrophils/mm^3	4027 (1411)	4426 (1979)	0.06
Fluorescent oxidation products, RFU/mL, GM (Q1 – Q3)	92.3 (79.5 – 103)	91.7 (76.5 – 102)	0.8

Data are means (SD) or n (%) unless otherwise stated. FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity. GM= geometric mean, (Q1 – Q3) = first and third quartiles. Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale.

Supplementary Table E11. Comparison of the 23 characteristics between the 576 participants without asthma followed-up at EGEA3 and the 90 not followed-up

	Followed-up (n=576)	**Not followed-up (n=90)**	**P-value**
Age, year	47.5 (15.3)	41.8 (17.9)	0.001
Sex, women	331 (57.5%)	40 (44.4%)	0.02
Smoking habits			
non-smokers	290 (50.5%)	40 (44.4%)	
ex-smokers	168 (29.2%)	27 (30.0%)	0.4
smokers	117 (20.3%)	23 (25.6%)	
Body Mass Index, kg/m^2	24.6 (3.9)	24.7 (4.4)	0.7
FEV_1 % predicted	108 (16.2)	102 (17.1)	0.004
FVC % predicted	113 (17.4)	106 (17.3)	<0.0001
Shortness of breath and wheezing, (last 12 months)	34 (5.9%)	4 (4.4%)	0.9
Attacks of breathlessness at rest, (last 12 months)	15 (2.6%)	3 (3.33%)	0.7
Attacks of breathlessness following strenuous activity, (last 12 months)	150 (26%)	30 (33.3%)	0.15
Nocturnal symptoms (last 12 months)			
Cough	143 (24.8%)	34 (37.8%)	0.01
Chest tightness	8 (1.4%)	2 (2.22%)	0.6
Shortness of breath	70 (12.1%)	13 (14.4%)	0.5
Inhaled corticosteroids (last 12 months)	19 (3.3%)	1 (1.11%)	0.5
Chronic cough	28 (4.9%)	5 (5.6%)	0.8
Chronic phlegm	423 (4%)	5 (5.6%)	0.6
Dyspnoea grade 3	34 (5.9%)	6 (6.7%)	0.8
Skin prick test positivity	203 (35.2%)	34 (37.8%)	0.6
Rhinitis (last 12 months)	120 (20.8%)	14 (15.6%)	0.2
Eczema (ever)	159 (27.6%)	21 (23.3%)	0.4
Total IgE, IU/ml, GM (Q1 – Q3)	43.2 (15.9 – 111)	54.3 (19.7 – 134)	0.15
White blood cell counts			
Eosinophils/mm^3	163 (118)	156 (115)	0.6
Neutrophils/mm^3	3982 (1300)	3903 (1534)	0.6
Fluorescent oxidation products, RFU/mL, GM (Q1 – Q3)	94.9 (81.8 – 108)	93.4 (80.5 – 109)	0.5

Data are means (SD) or n (%) unless otherwise stated. FEV_1= forced expiratory volume in 1 s. FVC= forced vital capacity. IgE= Immunoglobulin E. RFU= relative fluorescence intensity. GM= geometric mean, (Q1 – Q3) = first and third quartiles. Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens. Dyspnoea grade 3 was defined according to the Medical Research Council scale.

Supplementary Table E12. Description of the characteristics of the 341 CA at EGEA3 respective to their belonging endotypes identified at EGEA2

	CA 1 (n=41)	**CA 2 (n=188)**	**CA 3 (n=112)**	**P value**
Age, years	59.1 (6.5)	47.8 (15.8)	38.4 (13.5)	<0.0001
Sex, women	27 (65.9%)	104 (55.3%)	45 (40.2%)	0.006
Inhaled corticosteroids (last 12 months)	27/31 (87.1%)	90/146 (61.6%)	34/95 (35.8%)	<0.0001
Oral corticosteroids (last 12 months)	13/28 (46.4%)	44/147 (29.9%)	7/94 (7.5%)	<0.0001
Asthma attacks (last 12 months)	20/36 (51.3%)	87/176 (49.4%)	27/99 (27.3%)	0.0009
Asthma exacerbations (last 12 months)	13/24 (54.2%)	44/145 (30.3%)	8/89 (9.00%)	<0.0001
Asthma symptom score (last 12 months)	*n=34* 2.79 (1.92)	*n=170* 2.04 (1.51)	*n=109* 1.32 (1.27)	<0.0001
Nocturnal symptoms (last 12 months)				
Cough	22/30 (73.3%)	62/148 (41.9%)	32/97 (33.0%)	0.0005
Chest tightness	19/29 (65.5%)	67/151 (44.4%)	28/98 (28.6%)	0.0008
Shortness of breath	13/29 (44.8%)	24/154 (15.6%)	9/98 (9.18%)	<0.0001
Asthma Control Test (last 12 months)	*n=33* 18.2 (5.40)	*n=155* 22.1 (3.18)	*n=94* 23.2 (2.87)	<0.0001
≥20	16 (48.5%)	123 (79.3%)	85 (90.4%)	
<20	17 (51.5%)	32 (20.7%)	9 (9.60%)	<0.0001
Dyspnoea grade 3	19/36 (52.8%)	41/183 (22.4%)	12/109 (11.0%)	<0.0001
Rhinitis (last 12 months)	22/36 (61.1%)	127/171 (74.3%)	71/98 (72.4%)	0.3
Medical visit for asthma (last 12 months)	21/29 (72.4%)	76/136 (55.9%)	24/84 (28.6%)	<0.0001
COPD (medical diagnosis)	4/22 (18.2%)	10/133 (7.5%)	2/92 (2.2%)	0.02

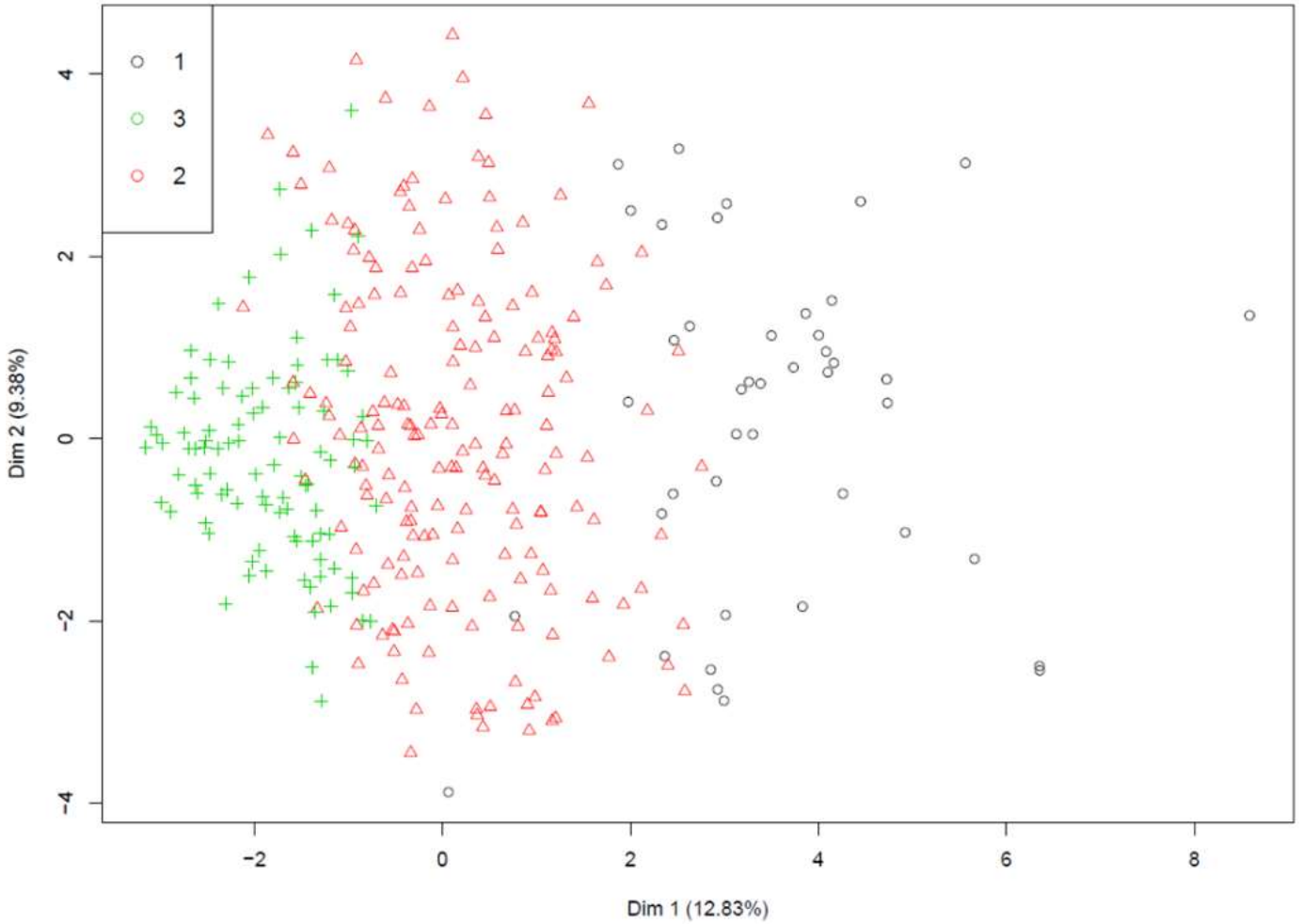
Data are means (SD) or n (%) unless otherwise stated.

Supplementary Table E13. Description of the characteristics of the 576 NA at EGEA3 respective to their belonging endotypes identified at EGEA2

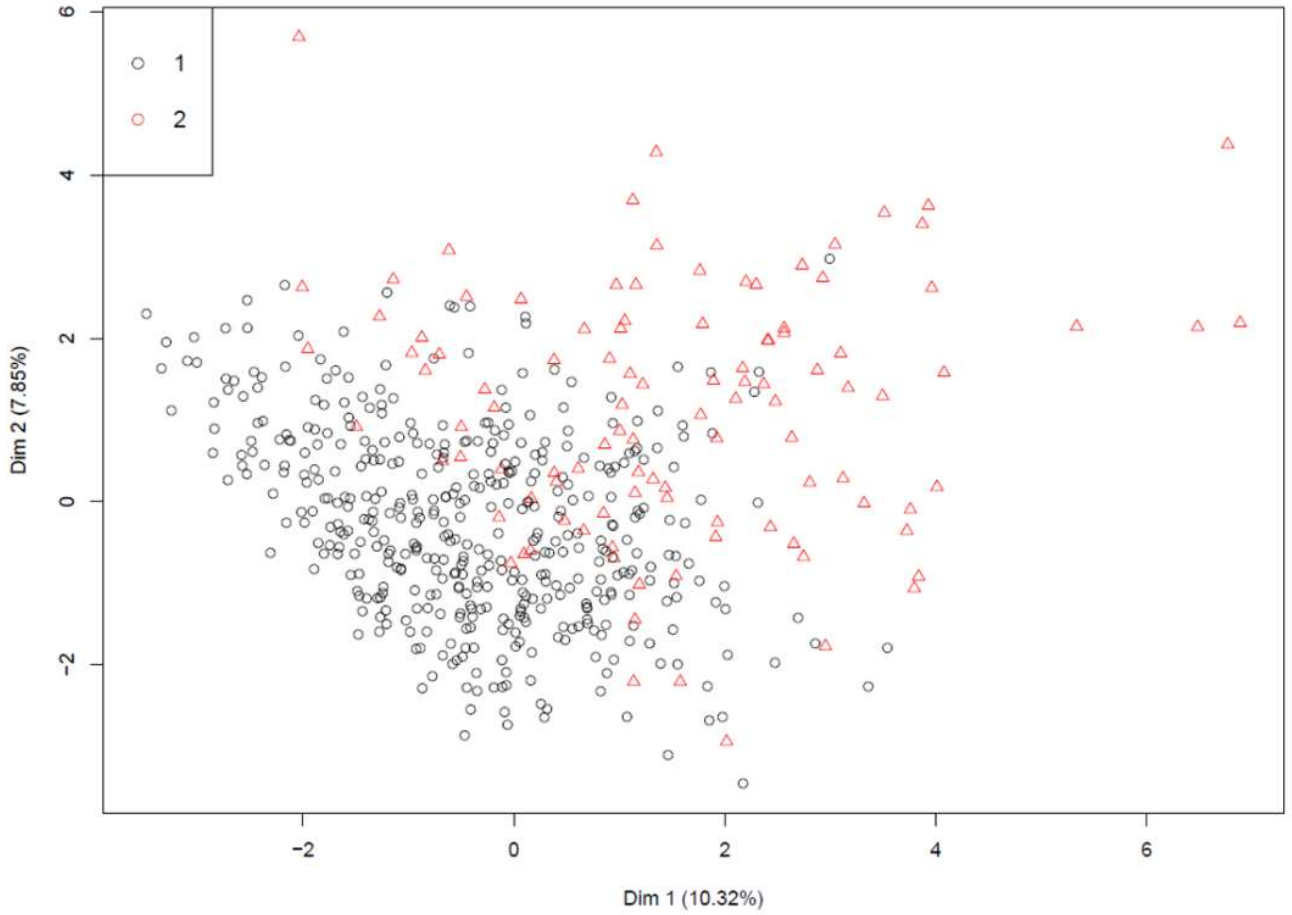
	NA 1 (n=427)	**NA 2 (n=149)**	**P value**
Age, year	53.9 (15.0)	55.6 (16.0)	0.2
Sex, women	236 (55.3%)	95 (63.8%)	0.07
Ever asthma	17/426 (3.99%)	14/148 (9.46%)	0.01
Wheezing (last 12 months)	29/423 (6.86%)	19 (12.7%)	0.03
Attacks of breathlessness at rest, (last 12 months)	16/423 (3.78%)	10/146 (6.85)	0.13
Attacks of breathlessness following strenuous activity , (last 12 months)	86/407 (21.1%)	48/142 (33.8%)	0.002
Nocturnal symptoms (last 12 months)			
Cough	70/364 (19.2%)	33/121 (27.3%)	0.007
Chest tightness	28/382 (7.32%)	14/132 (10.6%)	0.4
Shortness of breath	3/386 (0.78%)	3/133 (3.00%)	0.15
Asthma symptom score, last 12 months (0-5) *	*n=399* 0.38 (0.75)	*n=136* 0.59 (0.96)	0.007
Dyspnoea grade 3	26/419 (6.21%)	25/144 (17.4%)	<0.0001
Rhinitis (last 12 months)	130/396 (32.8%)	69/137 (50.4%)	0.0003

Data are means (SD) or n (%) unless otherwise stated.
Dyspnoea grade 3 was defined according to the Medical Research Council scale.

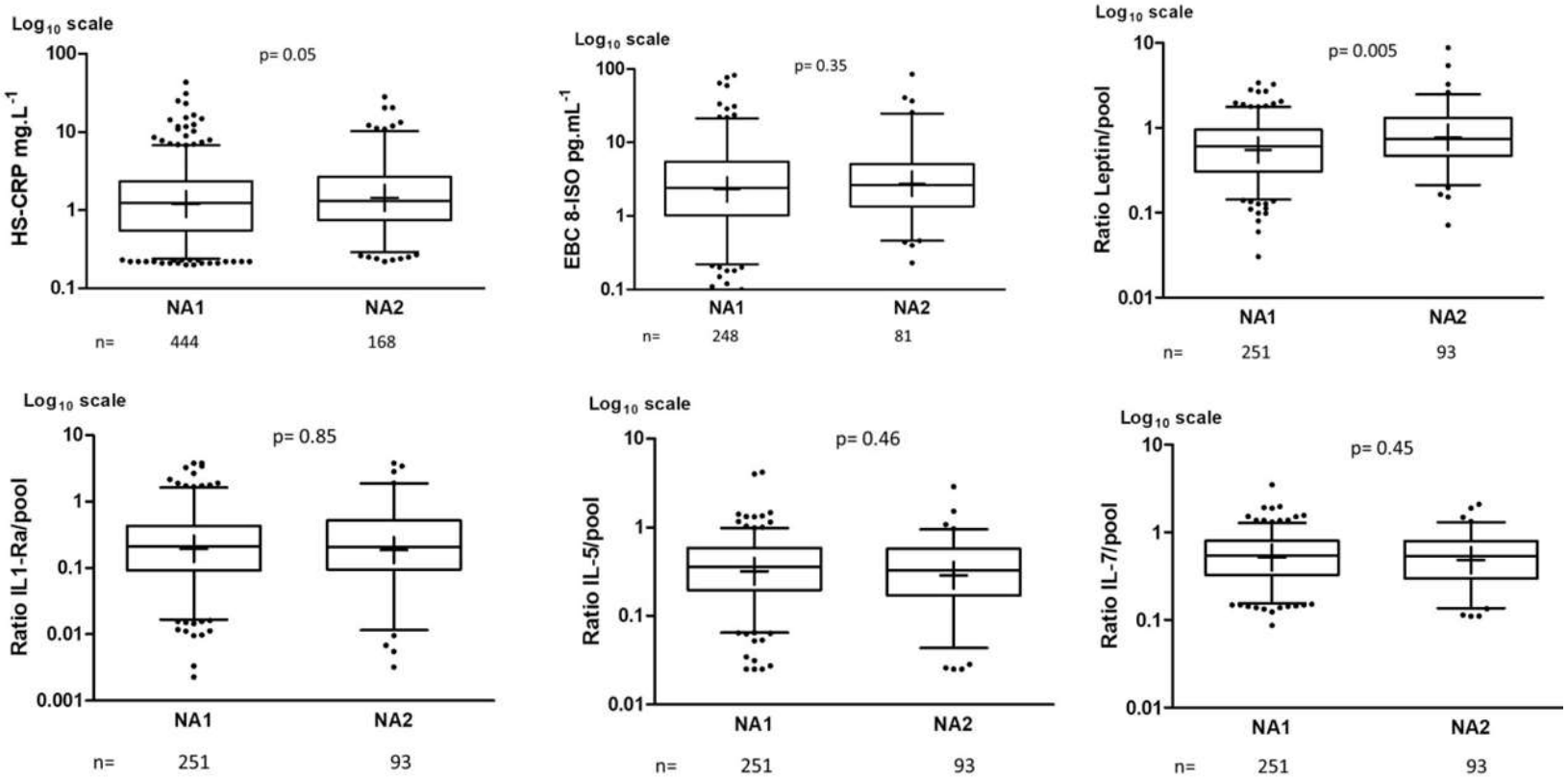
Supplementary Figure E1. Projection of the 3 clusters of asthmatics (n=402) into the factorial plane spawned by the first (DIM1) and second axes (DIM2) of FAMD.

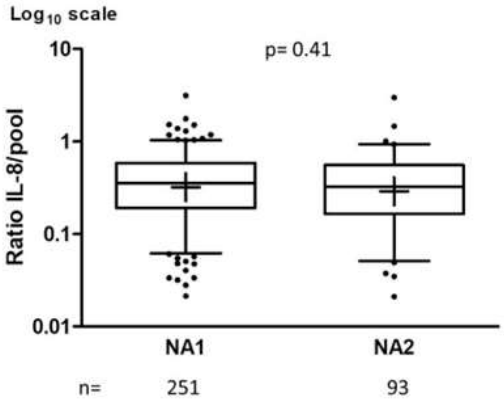


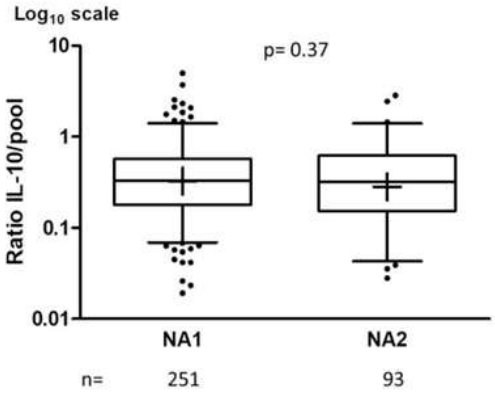
Supplementary Figure E2. Projection of the 2 clusters of non-asthmatics (n=666, step2) into the factorial plane spawned by the first (DIM1) and second axes (DIM2) of FAMD.

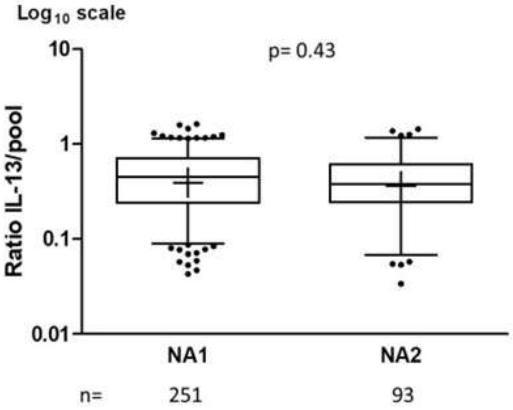


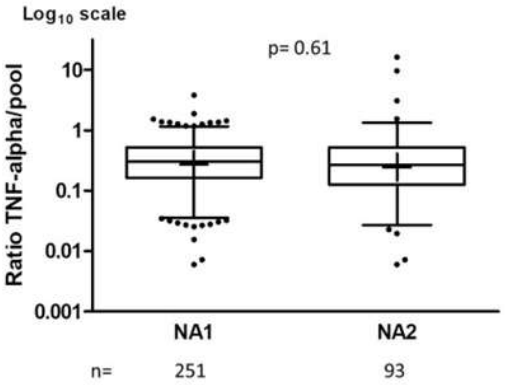
Supplementary Figure E3. Box plots of hs-CRP, 8-isoprostanes, leptin and seven cytokines plotted in NA1 and NA2 endotypes among never-asthmatics. The plots show the median (bar), the first and third quartiles (box), the 1^{rst} and 99th percentiles (whiskers) and the outliers (∗) for each endotype.
P values are adjusted for age and gender.




Log10 scale
p= 0.41
Ratio IL-8/pool
10
1
0.1
0.01
NA1
NA2
n=
251
93


Log10 scale
p= 0.37
Ratio IL-10/pool
10
1
0.1
0.01
NA1
NA2
n=
251
93


Log10 scale
p= 0.43
Ratio IL-13/pool
10
1
0.1
0.01
NA1
NA2
n=
251
93


Log10 scale
p= 0.61
Ratio TNF-alpha/pool
10
1
0.1
0.01
0.001
NA1
NA2
n=
251
93