

ONLINE SUPPLEMENTARY MATERIAL

Title: Association between occupational exposure to irritant agents and a distinct asthma endotype in adults

Authors:

Miora Andrianjafimasy *et al.*

Asthma and clinical characteristics

In the Epidemiological study on the Genetics and Environments of Asthma (EGEA), at both the baseline study (EGEA1, 1991-1995) and the follow-up (EGEA2, 2003-2007), complete examination included a detailed questionnaire (administered by an interviewer) and a medical examination. Lung function tests included measure of Forced Expiratory Volume in one second (FEV1) and methacholine bronchial challenge test (not performed in those with FEV1<80% predicted or post-diluent FEV1<90%). Bronchial responsiveness at 4 mg was evaluated by a decline of 20% in FEV1 (PD20 < 4mg). Skin prick tests (SPT) to 12 aeroallergens (cat, *Dermatophagoides pteronyssinus*, *Blattella germanica*, olive, birch, *Parietaria judaica*, timothy grass, ragweed pollen, *Aspergillus*, *Cladosporium herbarum*, *Alternaria tenuis*, cypress) were performed. SPT to each allergen was positive for a mean wheal diameter > 3 mm than the negative control. At EGEA2, serums, plasma and exhaled breath condensates were prepared and stored at -80°C according to standardized procedures until measurements [1]. Total serum IgE were assessed from blood samples in a centralized laboratory. Several biological markers involved in the inflammation, nitrosative or oxidative stress or immunologic pathways have been measured in blood or in exhaled breath condensate (EBC): the C-reactive protein, fluorescent oxidation products, EBC 8-isoprostanes, three antioxidant enzyme activities, blood and EBC total nitrate/nitrite, exhaled fraction of nitric oxide (FeNO), and nine cytokines [1].

At EGEA2, 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 in chest clinics at EGEA1 according to a standardized procedure as previously described [2]. 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 12 months [3].

Characteristics included in the cluster analysis

For all participants, four demographic characteristics: age (continuous), sex, smoking status (non-, ex- or current smokers) and body mass index (continuous); 20 clinical and functional characteristics: respiratory symptoms in the last 12 months (shortness of breath and wheezing, attacks of breathlessness following strenuous activity, at rest or at night-time, nocturnal chest tightness and cough, chronic cough and phlegm), dyspnea, skin prick tests (SPT) positivity for at least one of 12 aeroallergens, current rhinitis, ever eczema, use of inhaled medicines because of breathing problems in the past 12 months, forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC); and four biological characteristics: blood neutrophil and eosinophil counts, total serum Immunoglobulin E (IgE) level, and plasma fluorescent oxidation products (FIOPs) level, all expressed as continuous were selected in the cluster analysis. In addition, in current asthmatics, five asthma specific characteristics (age of asthma onset (continuous)), asthma attacks in the last 12 months, hospital or emergency admissions in the last 12 months and use of oral medicines because of breathing problems in the last 12 months) were also selected. In summary, 23 characteristics were included in the cluster analysis for the 632 never asthmatics and 28 characteristics for the 367 current asthmatics [4].

Mixture approach for cluster analysis

Cluster analysis was described in detail previously by Nadif et al. [4]. Briefly, to deal with both continuous and categorical data, we relied on Factor analysis of mixed data (FAMD) which combines principal components analysis (PCA) for continuous variables and a multiple

correspondence analysis (MCA) for qualitative variables [5]. FAMD converts all variables into continuous components. We first performed an imputation step. The `{missMDA}` package was used to impute missing data on smoking in never-asthmatics, rhinitis in asthmatics, and ever eczema, ICS use, allergic sensitization and total IgE level in both groups [6]. Then, we used the mixture approach [7] to identify clusters separately among current asthmatics and never-asthmatics. Specifically, we relied on Gaussian Mixture Models (GMMs) for their flexibility, and the availability of variety of covariance structures that can be obtained by means of an Eigen-decomposition. We used the `{mclust}` package which provides a comprehensive strategy for clustering and density estimation, and an integrated approach to finite mixture models with functions that combine model-based hierarchical clustering, expectation-maximization for mixture estimation and several tools for model selection [8]. The selection of the model and the choices of the number of components, of the covariance parameterization and of the numbers of clusters have been addressed by the Bayes Information criterion (BIC) [9,10]. Each participant was assigned to the group where the probability of belonging was the highest. All cluster analyses were performed using R statistical software (version 3.5).

References

- 1 Nadif R, Bouzigon E, Le Moual N, *et al.* EGEA Collection: a biobank devoted to asthma and asthma-related phenotypes. *Open J Bioresour* 2017;**322**:891–921.
- 2 Kauffmann F, Dizier MH, Pin I, *et al.* Epidemiological study of the genetics and environment of asthma, bronchial hyperresponsiveness, and atopy: phenotype issues. *Am J Respir Crit Care Med* 1997;**156**:S123-9.
- 3 Andrianjafimasy M, Zerimech F, Akiki Z, *et al.* Oxidative stress biomarkers and asthma characteristics in adults of the EGEA study. *Eur Respir J* 2017;**50**:1701193.
- 4 Nadif R, Febrissy M, Andrianjafimasy MV, *et al.* Endotypes identified by cluster analysis in asthmatics and non-asthmatics and their clinical characteristics at follow-up: the case-control EGEA study. *BMJ Open Resp Res* 2020;**7**:632.
- 5 Pagès J. Multiple factor analysis by example using R: Chapman and Hall/CRC. 2014.
- 6 Audigier V, Husson F, Josse J. A principal components method to impute missing values for mixed data. *Adv Data Anal Classif* 2016;**10**:5–26.
- 7 McLachlan G, Peel D. Wiley series in probability and statistics. *Finite Mix Model* 2000;**420**–7.
- 8 Scrucca L, Fop M, Raftery AE. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J* 2016;**8**:289–317.
- 9 Schwarz G. Estimating the dimension of a model. *Ann Stat* 1978;**6**:461–4.
- 10 Fraley C, Raftery A. Model-based clustering, discriminant analysis, and density estimation. *J Am Stat Assoc* 2002;**97**:611–31.
- 11 Le Moual N, Zock J-P, Dumas O, *et al.* Update of an occupational asthma-specific job exposure matrix to assess exposure to 30 specific agents. *Occup Env Med* 2018;**75**:507–14.

Figure E1: Flow chart of the study population

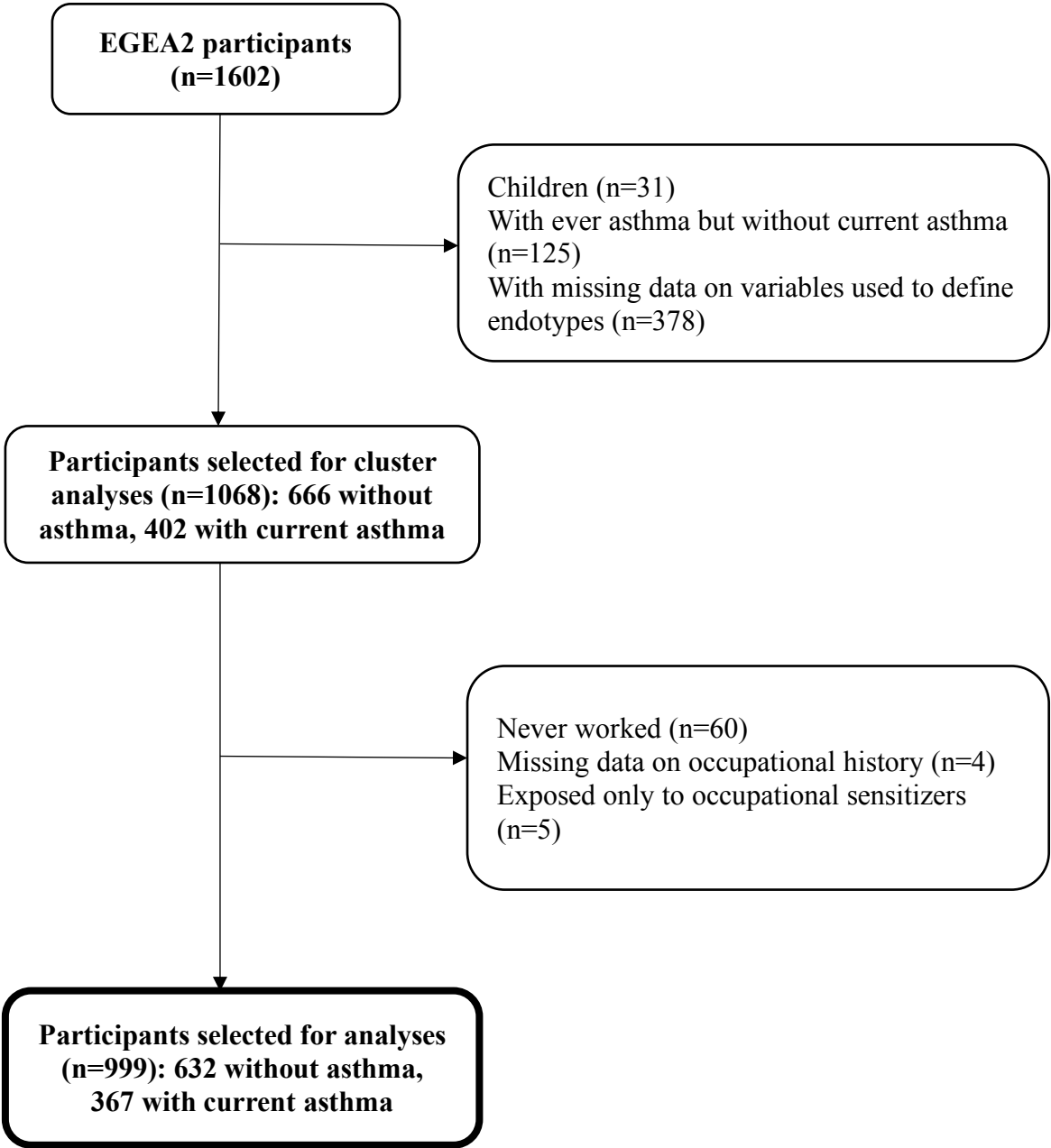


Table E1: List of irritant agents classified in three subgroups – new occupational asthma-specific job exposure matrix (OAsJEM)

Agents (n=19*/30)	Irritant subgroups				
	Any irritants	Highly reactive chemicals	Biocides	Cleaning products / disinfectants§	LMW agents
textiles	1				
moulds	1				
endotoxin	1				
high level chemical disinfectants	1	1	1	1	1
aliphatic amines	1	1			1
isocyanates	1	1			1
acrylates	1	1			1
epoxy resins	1	1			1
persulfates/ henna	1	1			1
wood	1				1
metal	1				1
metal working fluids	1				1
herbicides	1		1		
insecticides	1		1		
fungicides	1		1		
indoor cleaning	1			1	
bleach	1	1	1	1	
organic solvents	1	1			
exhaust fumes	1				

LMW: low molecular weight; some irritant agents have also been classified in the group of LMW agents.

*19 agents among the 30 agents at risk for asthma identified by OAsJEM [11].

§Agents regrouped in the subgroup of cleaning products / disinfectants.

Table E2: Description of the characteristics used to identify the five respiratory endotypes (according to Nadif et al. [4])

	All participants N=999	Never asthmatics (n=632)		Current asthmatics (n=367)		
		NA1 n=463	NA2 n=169	CA1 n=50	CA2 n=203	CA3 n=114
Age, years	44.8 ± 15.8	47.0 ± 14.8	48.5 ± 16.0	53.2 ± 15.5	41.9 ± 15.1	32.4 ± 13.2
Sex, Women	542 (54.3)	255 (55.1)	105 (62.1)	26 (52.0)	113 (55.7)	43 (37.7)
Smoking status						
Non-smokers	487 (48.8)	235 (50.8)	76 (45.0)	24 (48.0)	88 (43.4)	64 (56.1)
Ex-smokers	290 (29.0)	148 (32.0)	42 (24.8)	20 (40.0)	68 (33.5)	12 (10.5)
Current smokers	222 (22.2)	80 (17.3)	51 (30.2)	6 (12.0)	47 (23.1)	38 (33.3)
Body mass index, kg/m²	24.6 ± 4.2	24.4 ± 3.8	25.5 ± 4.4	27.2 ± 5.3	24.8 ± 4.3	23.1 ± 3.8
Age of asthma onset	15.7 ± 15.0	-	-	26.2 ± 17.5	15.9 ± 14.9	10.8 ± 11.2
Median (Q1-Q3)	10.0 (3.0-25.0)	-	-	25.0 (11.0-40.0)	10.0 (4.0-25.0)	7.0 (3.0-15.0)
FEV1 % predicted	102 ± 18.5	108 ± 16.1	104 ± 17.7	76.2 ± 21.7	97.2 ± 17.3	97.0 ± 15.6
FVC % predicted	111 ± 17.6	113 ± 17.7	111 ± 18.0	95.4 ± 20.5	110 ± 16.4	106 ± 12.9
Shortness of breath and wheezing, last 12m	252 (25.2)	2 (0.4)	32 (18.9)	42 (84.0)	126 (62.1)	50 (43.9)
<i>Asthma attacks, last 12m</i>	162 (44.1)	-	-	37 (74.0)	99 (48.8)	26 (22.8)
Attacks of breathlessness at rest, last 12m	122 (12.2)	3 (0.6)	12 (7.1)	26 (52.0)	67 (33.0)	14 (12.3)
Attacks of breathlessness following strenuous activity, last 12m	405 (40.6)	101 (21.9)	71 (42.0)	36 (72.0)	130 (64.0)	67 (58.8)
Nocturnal symptoms, last 12m						
Cough	320 (32.0)	97 (21.0)	71 (42.0)	30 (60.0)	98 (48.3)	24 (21.0)
Chest tightness	281 (28.1)	50 (10.8)	30 (17.7)	37 (74.0)	133 (65.5)	31 (27.2)
Shortness of breath	97 (9.7)	2 (0.4)	7 (4.1)	23 (46.0)	59 (29.1)	6 (5.3)
Hospital admissions, last 12m	4 (1.1)	-	-	3 (6.0)	1 (0.5)	0

<i>Emergency admissions, last 12m</i>	<i>7 (1.9)</i>	<i>-</i>	<i>-</i>	<i>6 (12.0)</i>	<i>1 (0.5)</i>	<i>0</i>
Inhaled corticosteroids, last 12m	201 (20.1)	1 (0.2)	18 (10.6)	41 (82.0)	106 (52.2)	35 (30.7)
<i>Oral corticosteroids, last 12m</i>	<i>57 (15.5)</i>	<i>-</i>	<i>-</i>	<i>18 (36.0)</i>	<i>34 (16.7)</i>	<i>5 (4.4)</i>
Chronic cough	81 (8.1)	1 (0.2)	31 (18.3)	21 (42.0)	24 (11.8)	4 (3.5)
Chronic phlegm	70 (7.0)	0	27 (16.0)	16 (32.0)	22 (10.8)	5 (4.4)
Dyspnea grade 3	110 (11.0)	4 (0.9)	35 (20.7)	28 (56.0)	32 (15.8)	11 (9.6)
Skin prick test positivity	523 (52.4)	167 (37.0)	56 (33.1)	35 (70.0)	159 (78.3)	106 (92.9)
Rhinitis	359 (36.0)	83 (18.0)	48 (28.4)	27 (54.0)	140 (69.0)	61 (53.5)
Eczema ever	362 (36.2)	130 (28.2)	46 (27.2)	19 (38.0)	103 (50.7)	64 (56.1)
Total IgE, IU/mL, GM (Q1-Q3)	67.5 (23.2-195)	40.3 (15.8-101)	50.6 (16.1-135)	144 (65.4-372)	125 (46-279)	203 (95.2-536)
Eosinophil counts, cells/mm³	198 ± 163	155 ± 108	177 ± 138	301 ± 320	255 ± 181	253 ± 187
Neutrophil counts, cells/mm³	4012 ± 1410	3915 ± 1278	4177 ± 1487	4986 ± 1651	4130 ± 1584	3522 ± 1074
FLOPs level, RFU/mL, GM (Q1-Q3)	94.2 (80.7-107)	94.5 (81.8-109)	97.3 (82.6-111)	106 (92.1-118)	91.7 (79.3-105)	88.9 (77.6-99.2)

Data are presented as n (%) or mean ± SD, unless otherwise stated. GM: geometric mean, (Q1 – Q3) = first and third quartiles;

In italics: variables not included to define endotypes in never asthmatics;

FEV1: forced expiratory volume 1s; IgE: immunoglobulin E; FLOPs: fluorescent oxidation products, biomarker of damages related to oxidative stress. Skin Prick Test positivity was defined by a mean wheal diameter ≥3mm than the negative control for at least one of 12 aeroallergens.

Table E3: Description of current occupational exposures to irritant agents in each respiratory endotype

	All participants	Never asthmatics		Current asthmatics		
	N=999	NA1 n=463	NA2 n=169	CA1 n=50	CA2 n=203	CA3 n=114
Any irritant agents						
No exposure	714 (71.5)	333 (72.1)	122 (72.2)	34 (68.0)	150 (73.9)	75 (65.8)
Medium	206 (20.6)	96 (20.8)	36 (21.3)	8 (16.0)	38 (18.7)	28 (24.6)
High	78 (7.8)	33 (7.1)	11 (6.5)	8 (16.0)	15 (7.4)	11 (9.6)
Highly reactive chemicals						
No exposure	714 (78.8)	333 (77.6)	122 (80.8)	34 (79.1)	150 (82.9)	75 (73.5)
Medium	154 (17.0)	76 (17.7)	25 (16.6)	5 (11.6)	25 (13.8)	23 (22.6)
High	38 (4.2)	20 (4.7)	4 (2.6)	4 (9.3)	6 (3.3)	4 (3.9)
Biocides						
No exposure	714 (84.6)	333 (83.9)	122 (85.3)	34 (85.0)	150 (87.2)	75 (81.5)
Medium	99 (11.7)	49 (12.3)	18 (12.6)	3 (7.5)	17 (9.9)	12 (13.0)
High	31 (3.7)	15 (3.8)	3 (2.1)	3 (7.5)	5 (2.9)	5 (5.4)
Cleaning products / disinfectants						
No exposure	714 (81.9)	333 (82.6)	122 (81.9)	34 (82.9)	150 (82.9)	75 (76.5)
Medium	134 (15.4)	57 (14.1)	24 (16.1)	5 (12.2)	28 (15.5)	20 (20.4)
High	24 (2.7)	13 (3.2)	3 (2.0)	2 (4.9)	3 (1.7)	3 (3.1)

Data are presented as n (%).

Table E4: Description of the job titles of the eight participants in CA1 endotype with high occupational exposure to irritants

	Irritants	HRC	Biocide	Cleaning products	LMW	Job title*
Worker #1	High	High	High	High	Medium	Institution-based personal care workers
Worker #2	High	.	High	.	.	Tree and shrub crop growers
Worker #3	High	.	High	.	.	Gardeners, horticultural and nursery growers
Worker #4	High	High	.	Medium	Medium	Painters and related workers
Worker #5	High	High	.	.	High	Machine-tool setters and setter-operators
Worker #6	High	.	.	.	High	Metal melters, casters and rolling-mill operators
Worker #7	High	High	.	.	High	Electrical-equipment assemblers
Worker #8	High	Medium	Medium	High	Medium	Helpers and cleaners in offices, hotels and other establishments

*job title according to the International Standard Classification of Occupations (ISCO-88)

HRC: highly reactive chemicals; LMW: low molecular weight

Table E5: Comparison of the number of participants in each category of occupational exposure to irritants estimated without and with the expert reassessment step

Without expert reassessment step				
With expert reassessment step	No exposure n (%)	Medium n (%)	High n (%)	n total
No exposure n (%)	683 (68.6)	24 (2.4)	4 (0.4)	711
Medium n (%)	0	173 (17.4)	33 (3.3)	206
High n (%)	0	17 (1.7)	61 (6.1)	78
n total	683	214	98	995

Missing value n=4

Interpretation: 2.4% means that 24 participants classified as medium exposure without the expert step were classified as no exposure with the expert reassessment step.

The expert reassessment step was carried out for 43% of all jobs.

Table E6: Associations between occupational exposure to any irritant agents estimated without the expert reassessment step and respiratory endotypes (N=999)

	Occupational exposure to any irritant agents				
	No exposure	Medium	p [§]	High	p [§]
NA2 vs. NA1					
n [#]	432 (117 vs. 315)	140 (34 vs. 106)		57 (17 vs. 40)	
crudeOR (95% CI)	1	0.86 (0.56- 1.34)	0.51	1.14 (0.65- 2.02)	0.66
*adjustedOR (95% CI)	1	0.86 (0.54- 1.36)	0.51	1.09 (0.60- 1.96)	0.79
CA1 vs. NA1					
n [#]	349 (34 vs. 315)	111 (5 vs. 106)		51 (11 vs. 40)	
crudeOR (95% CI)	1	0.42 (0.16- 1.14)	0.08	2.73 (1.26- 5.92)	0.01
*adjustedOR (95% CI)	1	0.41 (0.15- 1.10)	0.08	2.79 (1.19- 6.56)	0.02
CA2 vs. NA1					
n [#]	458 (143 vs. 315)	150 (44 vs. 106)		56 (16 vs. 40)	
crudeOR (95% CI)	1	0.89 (0.60- 1.33)	0.66	0.86 (0.45- 1.66)	0.68
*adjustedOR (95% CI)	1	0.93 (0.62- 1.40)	0.73	0.87 (0.44- 1.71)	0.69
CA3 vs. NA1					
n [#]	389 (74 vs. 315)	132 (26 vs. 106)		54 (14 vs. 40)	
crudeOR (95% CI)	1	1.04 (0.63- 1.72)	0.87	1.47 (0.77- 2.79)	0.23
*adjustedOR (95% CI)	1	1.09 (0.63- 1.89)	0.76	1.30 (0.62- 2.73)	0.49

Significant results in bold.

#n total (n in NA2, CA1, CA2 or CA3 vs. n in NA1); OR: odds ratio for the association between occupational exposure and each endotype (NA2, CA1, CA2 or CA3 vs. NA1) *adjusted for age, sex and smoking status; §for the univariate model (crudeOR), p-values correspond to p-values of Chi-squared test or Fisher exact test

Table E7: Associations between occupational exposure to any irritant agents and respiratory endotype CA1 after further adjustment for body mass index (n=999)

	Occupational exposure to any irritant agents									
	Without expert reassessment step					With expert reassessment step				
	No exposure	Medium	p	High	p	No exposure	Medium	p	High	p
CA1 vs. NA1										
n#	349 (34 vs 315)	111 (5 vs 106)		51 (11 vs 40)		367 (34 vs 333)	104 (8 vs 96)		41 (8 vs 33)	
adjusted OR (95% CI)	1	0.42 (0.16- 1.10)	0.08	2.59 (1.05- 6.38)	0.04	1	0.79 (0.35-1.75)	0.55	2.57 (0.87-7.58)	0.09

OR: odds ratio for the association between occupational exposure and each endotype (NA2, CA1, CA2 or CA3 vs. NA1) adjusted for age, sex, smoking and BMI; #n total (n in NA2, CA1, CA2 or CA3 vs. n in NA1).

Significant results in bold.

Table E8: Associations between occupational exposure to subgroups of irritants and the endotype CA1

Irritant subgroups	CA1 vs. NA1				
	No exposure	Medium exposure		High exposure	
<i>Without expert reassessment step</i>					
Highly reactive chemicals					
n [#]	349 (34 vs. 315)	88 (5 vs. 83)		23 (4 vs. 19)	
crudeOR (95% CI), p [†]	1	0.52 (0.19-1.43)	0.23	2.14 (0.72- 6.42)	0.28
adjustedOR* (95% CI), p	1	0.54 (0.20-1.46)	0.22	1.93 (0.52-7.15)	0.33
Biocides					
n [#]	349 (34 vs. 315)	59 (3 vs. 56)		21 (5 vs. 16)	
crudeOR (95% CI), p [†]	1	0.48 (0.13-1.72)	0.25	3.36 (1.27-8.89)	0.06
adjustedOR* (95% CI), p	1	0.47 (0.13-1.65)	0.24	3.89 (1.32-11.5)	0.01
Cleaning products / disinfectants					
n [#]	349 (34 vs. 315)	59 (2 vs. 57)		22 (4 vs. 18)	
crudeOR (95% CI), p [†]	1	0.30 (0.07-1.37)	0.11	2.27 (0.79-6.55)	0.26
adjustedOR* (95% CI), p	1	0.31 (0.07-1.43)	0.13	2.25 (0.71-7.16)	0.17
<i>With expert reassessment step</i>					
Highly reactive chemicals					
n [#]	367 (34 vs. 333)	81 (5 vs. 76)		24 (4 vs. 20)	
crudeOR (95% CI), p [†]	1	0.61 (0.23-1.64)	0.37	2.02 (0.65-6.24)	0.27
adjustedOR* (95% CI), p	1	0.60 (0.22-1.68)	0.33	2.15 (0.62-7.50)	0.23
Biocides					
n [#]	367 (34 vs. 333)	52 (3 vs. 49)		18 (3 vs. 15)	
crudeOR (95% CI), p [†]	1	0.58 (0.17-1.95)	0.40	2.26 (0.71-7.20)	0.40
adjustedOR* (95% CI), p	1	0.53 (0.15-1.91)	0.33	3.13 (1.00-9.83)	0.05
Cleaning products / disinfectants					
n [#]	367 (34 vs. 333)	62 (5 vs. 57)		15 (2 vs. 13)	
crudeOR (95% CI), p [†]	1	0.89 (0.34-2.33)	0.76	1.71 (0.39-7.41)	0.64
adjustedOR* (95% CI), p	1	0.87 (0.32-2.38)	0.78	2.12 (0.45-9.98)	0.34

OR: odds ratio for the association between occupational exposure and each endotype (NA2, CA1, CA2 or CA3 vs. NA1)

*adjusted for age, sex, smoking status; #n total (n in CA1 vs. n in NA1); †p-value of Chi-squared test or Fisher exact test for small sample size (n<5).

Significant results in bold.