Journal of Breath Research

ACCEPTED MANUSCRIPT • OPEN ACCESS

Increased breath naphthalene in children with asthma and wheeze of the All Age Asthma Cohort (ALLIANCE)

To cite this article before publication: Pedram Shahrokny *et al* 2023 *J. Breath Res.* in press https://doi.org/10.1088/1752-7163/acf23e

Increased breath naphthalene in children with asthma and wheeze of the All Age Asthma Cohort (ALLIANCE)

P Shahrokny[1,*], N Maison[2,3,*], L Riemann[4,5,6], M Ehrmann[2], D DeLuca[4], S Schuchardt[7], D Thiele[8], M Weckmann[8,9], AM Dittrich[4], B Schaub[2], F Brinkmann[8], G Hansen[4], MV Kopp[10], E von Mutius[2,3], KF Rabe[11], T Bahmer[11,12], JM Hohlfeld[1,13], R Grychtol[4,*], O Holz[1,*] and the Alliance study group

*contributed equally

[1] Fraunhofer ITEM, Department of Clinical Airway Research, German Center for Lung Research (BREATH, DZL), Hannover, Germany
[2] Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University, German Center for Lung Research (CPC-M, DZL), Munich, Germany
[3] Institute of Asthma and Allergy Prevention, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[4] Department of Paediatric Pneumology, Allergology and Neonatology, Hannover Medical School, German Center for Lung Research (BREATH, DZL), Hannover, Germany
[5] Clinician Scientist Program TITUS, Else-Kröner-Fresenius-Stiftung, Hannover Medical School, Hannover, Germany
[6] Institute of Immunology, Hannover Medical School, Hannover, Germany
[7] Fraunhofer ITEM, Bio- and Environmental Analytics, Hannover, Germany
[8] Division of Pediatric Pulmonology and Allergology, University Children's Hospital, German Center for Lung Research (ARCN, DZL), Luebeck, Germany
[9] Epigenetics of Chronic Lung Disease, Priority Research Area Chronic Lung Diseases, Leibniz Lung Research Center Borstel, Borstel, Germany
[10] Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland
[11] LungenClinic Grosshansdorf and Department of Medicine, Christian-Albrechts-University Kiel, German Center for Lung Research (ARCN, DZL), Grosshansdorf, Germany
[12] Pulmonary Research Institute at LungenClinic Grosshansdorf, German Center for Lung Research (ARCN, DZL), Grosshansdorf, Germany,
[13] Department of Respiratory Medicine, Hannover Medical School (MHH), Hannover, Germany

Funding: German Center for Lung Research (DZL), Else-Kröner-Fresenius-Stiftung

Wordcount: 3544

Revised Version 18.07.23

Corresponding author:

Prof. Dr. JM Hohlfeld
Department of Clinical Airway Research,
Fraunhofer Institute for Toxicology and Experimental Medicine,
30625 Hannover, Germany
Phone: +49-511-5350-8101
Fax: +49-511-5350-8250
jens.hohlfeld@item.fraunhofer.de

Abstract

Background

Exhaled breath contains numerous volatile organic compounds (VOCs) known to be related to lung disease like asthma. Its collection is non-invasive, simple to perform and therefore an attractive method for the use even in young children. We analysed breath in children of the multicenter All Age Asthma Cohort (ALLIANCE) to evaluate if "breathomics" have the potential to phenotype patients with asthma and wheeze, and to identify extrinsic risk factors for underlying disease mechanisms.

Methods

A breath sample was collected from 142 children (asthma: 51, pre-school wheezers: 55, healthy controls: 36) and analysed using gas chromatography–mass spectrometry (GC/MS). Children were diagnosed according to GINA guidelines and comprehensively examined each year over up to seven years. Forty children repeated the breath collection after 24 or 48 months.

Results

Most breath VOCs differing between groups reflect the exposome of the children. We observed lower levels of lifestyle-related VOCs and higher levels of the environmental pollutants, especially naphthalene, in children with asthma or wheeze. Naphthalene was also higher in symptomatic patients and in wheezers with recent inhaled corticosteroid use. No relationships with lung function or TH2 inflammation were detected.

Conclusion

Increased levels of naphthalene in asthmatics and wheezers and the relationship to disease severity could indicate a role of environmental or indoor air pollution for the development or progress of asthma. Breath VOCs might help to elucidate the role of the exposome for the development of asthma.

The study was registered at ClinicalTrials.gov (NCT02496468).

Word count: 235

Key words: exhaled air, VOC, pediatric asthma, wheeze, ALLIANCE

Introduction

Asthma is one of the most prevalent pediatric chronic diseases worldwide and causes significant burden on patients, family, society and healthcare systems [1–3]. Characteristic clinical features are chronic airway inflammation and variable expiratory airflow obstruction presenting as cough, wheeze, chest tightness and dyspnea usually in response to specific triggers, e.g. viral infections and exposure to allergens [4].

Despite similar clinical manifestation, the pathobiology and course of the disease is very heterogeneous [5]. Many factors contribute to asthma development, among them genetics, epigenetics [6], prenatal influences as maternal smoking, viral respiratory tract infections in early life [7] and unfavorable environmental exposures to cigarette smoke, air pollution, allergens, or mold [8, 9]. There is particular need for non-invasive and simple to assess biomarkers for early disease detection and characterization, especially in children. While exhaled nitric oxide is considered as an established non-invasively accessible biomarker [10, 11], it mainly reflects the level of eosinophilic airway inflammation. For a more comprehensive analysis, volatile organic compounds (VOC) in exhaled breath have gained considerable interest as biomarkers for lung diseases, especially in asthma [12–14].

Breath collection is non-invasive and only requires tidal breathing which makes it particularly attractive for the use in children even at a young age. A number of studies on breath VOCs in adult and pediatric asthma have been published in the last decade and are considered in recent reviews [12–14]. For example, Dallinga et al. showed that exhaled VOCs can distinguish asthmatic from non-asthmatic children [15], and there is evidence that breath VOCs identify distinct inflammation phenotypes [16] or predict exacerbations in asthmatic children [17, 18]. Despite a large body of literature with a number of VOCs potentially discriminating between healthy controls and asthma patients or reflecting disease activity or treatment, there is currently no breath VOC biomarker or biomarker pattern that supports physicians in the diagnosis, treatment, and phenotyping of disease or in preventing exacerbations.

Here, we collected and analyzed breath VOCs from a subgroup of children of the All Age Asthma cohort (ALLIANCE), a multicenter prospective observational cohort recruiting children with pre-school wheeze and children and adults with asthma [19].

We hypothesized that VOCs or VOC patterns are distinct between children with asthma or pre-school wheeze compared to healthy children and contribute to identify extrinsic risk factors, and underlying disease mechanisms. Additionally, we investigated if VOCs were linked to

inflammatory phenotypes, clinical features such as lung function, asthma treatment and exacerbation rate.

Methods

Study design

The ALLIANCE cohort of the German Center for Lung Research (DZL) is a prospective multicenter asthma cohort [19]. For this study, we collected 182 breath samples from 142 children (51 children with asthma, 55 children with wheeze and 36 healthy controls) at two pediatric specialist centers (Hannover, Munich) from October 2016 until spring 2020 (**figure 1, table S1**). In a subgroup of 40 patients, a second breath collection was performed after one or two years (asthma n=20, wheeze n=20). The study was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with local statutory requirements. The study was registered at ClinicalTrials.gov (NCT02496468) and approved by all local ethics committees. All parents of study participants <18 years as well as study participants ≥8 years gave their written informed consent.

Children aged 6 months to 5 years were included if they had at least two episodes of wheeze during the past 12 months ('pre-school wheezer') as indicated by the parents in the respective questionnaire. Children ≥6 years were included based on doctor-diagnosed asthma according to the Global Initiative for Asthma (GINA) guidelines and German guidelines [4, 20]. We also recruited healthy control subjects who had never been diagnosed with asthma or pre-school wheeze. Further inclusion and exclusion criteria are specified in the supplement or have been published [19]. Laboratory tests included differential blood count, and specific immunoglobulin E against 36 allergens measured by Euroline™ (Euroimmun, Germany). All children performed exhaled nitric oxide (FeNO) measurements and spirometry from age 4 years onwards and VOC measurements from 3 years onwards. For the breath and near-subject room air collection we used an in-house developed breath collection device [21]. Further details regarding all procedures and study design are presented in the **online supplement**.

Analysis by gas chromatography-mass spectrometry (GC-MS)

GC-analysis of the samples was performed at Fraunhofer ITEM, Hannover, within eight days of initial breath collection. Turbo Mass Software 5.4 (PerkinElmer, USA) was used for automated identification of individual VOCs based on retention time of specific masses, comparing values with reference compounds and the NIST database (National Institute of

Standards and Technology Mass Spectral Search Program Version 2.2 (NIST, USA)) [21, 22]. A total of 158 VOCs were quantified using peak height of specific target ions which most commonly matched with m/z signals of highest intensity in the respective VOC mass spectrum (total Ion content (TIC)) [22]. Further details regarding measurement and analysis are available in the **online supplement**.

Statistics

We computed the mean value of the 158 target VOCs for all technical replicate samples. Data was then log-transformed. We used the Mann-Whitney-U-Test (MWU) for univariate comparisons between patients (asthma, wheeze) and healthy controls and the Pearson correlation coefficient where appropriate. The false discovery rate was assessed by correcting raw p-values with the *p.adjust* function in R (using the Benjamini and Hochberg (BH) option). Corrected p-values are indicated as pBH. Site effects were accounted for by batch correction using the *ComBat* function of the R package *sva* to our data. For some analyses that were limited to data from one site we used both the original (uncorrected) and the batch corrected "site-effect–free data" to test whether similar results were obtained.

Due to the exploratory nature of our analysis, we focused on the complete VOC dataset (n=158 VOCs). However, to avoid missing important observations by the multiplicity correction we also ran the analysis with subsets of our 158 VOCs. A detailed description of these subsets is provided in the **online supplement**. In addition we dimensionally reduced the dataset using an unsupervised clustering approach called Cytomod [23]. Briefly, VOCs were clustered and assigned to co-varying modules based on the pairwise Pearson-correlations between log-transformed VOC values. For more details, please refer to the **online supplement**. Module scores reflecting the expression of each module in each participant (see **online supplement**) were compared between participant groups with different characteristics using MWU tests. P-values were adjusted for the number of modules tested. A multivariable logistic regression model was used to analyze the association between module 6 and disease status. We focused our analysis on module 6, because most VOCs considered as air pollutants were found in module 6.

Results

Demographics

As shown in **figure 1**, we included 55 children with pre-school wheeze, 51 children with asthma and 36 healthy controls (**table 1**). Total IgE and FeNO were higher ($p<0.01$) in asthmatics than in healthy controls while FEV_1 and FEV_1/FVC values were lower ($p<0.01$). Blood eosinophils were higher in asthmatic children compared to controls ($p=0.01$), and atopy was more prevalent in subjects with wheeze and asthma (**table 1**).

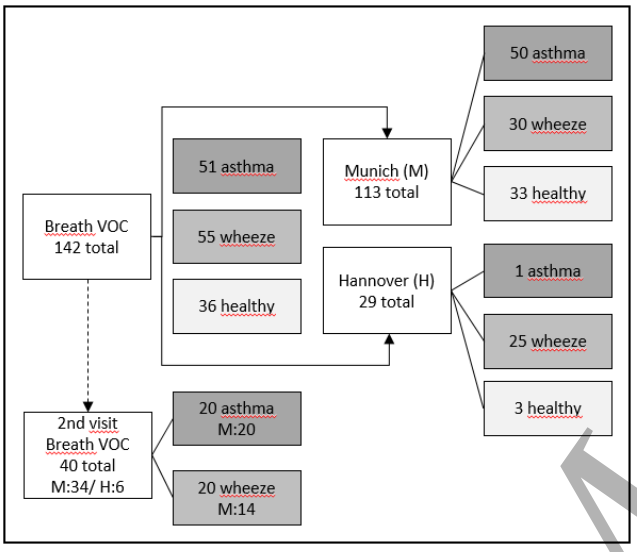

Figure 1. Study population. Breath VOCs were analyzed from 142 subjects. 40 subjects were sampled on two visits, which were at least one year apart. The Hannover site had a focus on wheezers and started the recruitment for VOC sampling later during the study. The one patient with asthma from Hannover was planned to be included as wheezer but the visit could not be scheduled before the 6th birthday, therefore based on the definition criteria, this subject was considered as child with asthma.

Basic data description, quality, and plausibility of VOC data

As expected, acetone and isoprene were major VOCs in breath of children, while disinfectants like 1-propanol, 2-propanol, ethanol, 2-phenoxyethanol, and 1-phenoxypropan-2-ol were the major VOCs detected in the room air samples. In line with the "owncloud concept" [22] we found the expected correlations between near-subject room air and breath for VOCs, that are frequently used in lifestyle products (personal care, cosmetics, home care) or for environmental pollutants. Among these were geranyl acetone, siloxanes, as well as benzene (**figure S1**). Furthermore, patients exposed to second-hand smoke and in one patient to active smoking showed increased levels of cigarette smoke related VOCs (**table S2**) [21, 24]. Isoprene

levels in exhaled breath significantly correlated with age (r=0.36, p<0.0001), as previously shown [25].

Potential confounders: site, age, gender

The median levels of most VOCs were different between study sites. We consider technical reasons, like instrument drift, as unlikely for this observation (details are presented in the **online supplement)** and adjusted the data as outlined in material and methods. To evaluate the impact of age we checked the correlation between VOC values and age using the site effect-free data for all subjects. Besides the already mentioned positive isoprene correlation, we only found a negative correlation for 1,2-propandiol and age (r=-0.27), with borderline significance (pBH=0.068). No significant gender-specific differences for breath VOCs were observed (pBH>0.05).

Reproducibility between visits

For 40 children (asthma n=20, wheeze n=20) we were able to collect a second sample in a follow-up visit one year after taking the initial sample. In three cases, the second sample was taken two years later. Initial and follow-up VOCs values correlated significantly for 13 VOCs (pBH<0.05, **figure S2**, **table S3**). Most of these VOCs appear to be environment- and lifestyle-related. The correlation between the visits was not significant for naphthalene (**figure S3**, p=0.06, pBH=0.27).

Cluster analysis

In the unsupervised cluster analysis, we identified a total of 9 modules (**figure 2**), consisting of 3 to 39 VOCs. The VOCs contained in each module are listed in the supplement (**table S4**). Interestingly, all BTEX (benzene, toluene, ethylbenzene, xylenes) VOCs and other VOCs considered as air pollutants, like naphthalene, were found in module 6. A large number of VOCs in module 4 were substances used as fragrances and flavors typically found in many lifestyle and cleaning products.

VOCs differ between children with asthma or pre-school wheeze and healthy controls

We found 24 significantly different VOCs (pBH<0.05) between healthy controls and asthma patients and 23 significantly different (pBH<0.05) VOCs between healthy controls and children

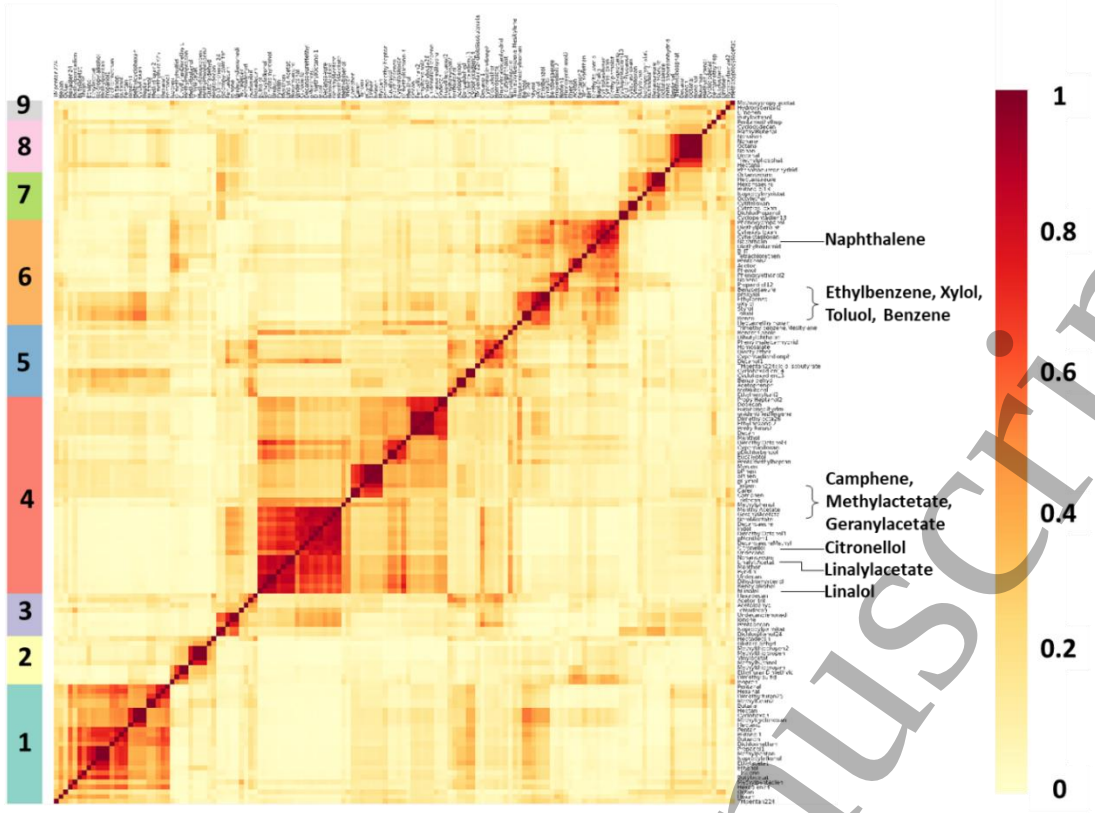

Figure 2: Heatmap of VOC modules: VOCs are first clustered together based on pairwise Pearson correlations. A reliability score (i.e., the fraction of times that a given VOC is assigned to the same cluster) is calculated over 1,000 permutations of participants and is used to assign VOCs to modules. The color bars on the left side depict the module membership for each VOC (a full list is given in supplement **table S4**), and the coloring in the heatmap represents the reliability score of each VOC.

with wheeze (**table 2**). There was a large overlap, with 13 VOCs having significantly different levels both in asthma patients and wheezers compared to healthy controls (**table 2**). As clinical characteristics differed between healthy controls and patients (**table 1**), we performed additional subgroup analyses which are described in detail in the **online supplement.** We did not find any evidence that these differences influenced the results.

Most of the VOCs with significantly higher levels in breath of healthy controls are used as fragrances or flavors and can therefore be found in lifestyle products and foods (**table 2**). Some of these VOCs with suspected fragrance or flavor origin correlated among each other, which could indicate common sources (**table S5**). This result was also mirrored in the cluster analysis data, as most of the VOCs, which were increased in healthy controls, were found in module 4 as indicated in **table 2** and **figure 2**. Module 4 values were also significantly higher in healthy controls compared to asthmatic children (pBH=0.006) and children with wheeze (pBH=0.03), respectively (**figure S4**).

Breath naphthalene levels showed the largest difference between children with asthma or wheeze and healthy controls (**table 2, figure 3**). Furthermore, other VOCs considered as air pollutants, such as 1,2-propanediol and ethylhexanol were also increased in asthma or wheeze. Although not statistically significant (pBH>0.05), we additionally observed higher breath levels for other aromatic air pollutants (**table 2**). These were significantly correlated (all p<0.01) among each other and with naphthalene (**figure S5**), again suggesting a common source. All these environmental pollutants were found in module 6 as indicated in **table 2.** Consistent with this result, overall module 6 values were higher in asthmatic children (pBH=0.07) and children with wheeze (pBH=0.03) compared to healthy controls (**figure S4**).

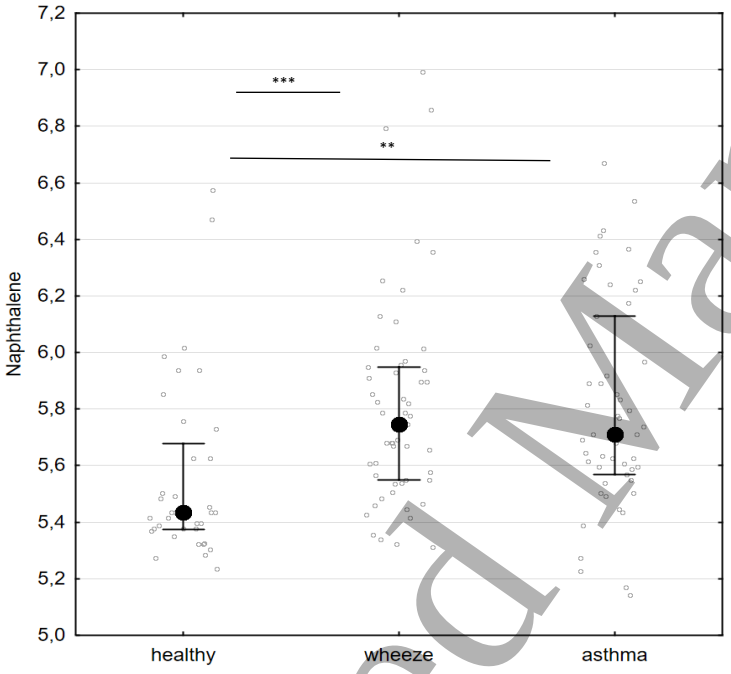

Figure 3. Differences in the levels of naphthalene between groups. Median and interquartile ranges. Individual data is shown in grey open circles. *** p<0.001, ** p<0.01.

<u>VOC profiles associated with disease characteristics</u>

VOCs or module 4 and 6 values did not differ depending on presence of atopy, blood eosinophilia, or increased FeNO (see **online supplement** for definition and further details). Similarly, we did not find any correlations between VOCs and FeNO or absolute blood eosinophil levels. No correlations were found for VOCs or module values and FEV_1 z-score, neither for asthma patients, wheezers or both groups combined. There were also no differences when comparing participants with FEV_1 z-scores <-1 to those with z-scores >-1. The

same results were found for the FEV_1/FVC data. For FEF_{25-75} the results were also similar, except for asthma patients. Here we found higher levels for six VOCs (all in module 4), among them capric-acid, undecane and undecanal and for module 4 values (pBH<0.05) in asthma patients with a z-score <-1.

VOC profiles associated with disease control

No significantly different VOC levels were found between asthmatic children and wheezers with controlled disease according to GINA guideline definition [4] (n=87 and those with "partly-controlled" or "uncontrolled" asthma" (n=19) when testing the different VOC subsets as defined in material and methods (pBH>0.05)). Testing only naphthalene showed significantly higher levels in the partly- and uncontrolled group (p<0.008), which is compatible with the observed trend to increased module 6 values (p=0.03, pBH 0.251). No difference was observed for module 4.

VOCs or module 4 and 6 values did not differ between children with or without exacerbations or wheeze episodes in the past 12 months. In wheezers with current use of inhalative corticosteroids (ICS), we found significantly higher levels for 1-phenoxypropan-2-ol (pBH=0.03) and for naphthalene (p=0.01). Both are found in module 6, which also showed higher values in this group (pBH=0.06). In addition, higher levels for an unidentified terpene (pBH≤0.008), and three alkanes (undecane, dodecane and decane) were observed (pBH=0.09) Two of these VOCs were found in module 4, which showed a trend for higher module values in this group (pBH=0.06).

Association of cluster module 6 with asthma or pre-school wheeze

As VOC module 6 was increased in children with asthma or pre-school wheeze, we were interested in investigating this association while controlling for other known risk factors for asthma development. In a multivariable analysis, increased values for VOC module 6 were significantly associated with asthma and pre-school wheeze, also when adjusting for age, parental asthma, gender, presence of atopy, elevated blood eosinophil counts, and secondary cigarette smoke exposure (**table 3**). Only atopy was also significantly associated to asthma or wheeze status in our multivariable regression model.

Discussion

In this study, we showed that breath VOCs differ between healthy children and patients with asthma or pre-school wheeze. Most of the differential VOCs were related to the exposome of the patients. Lifestyle-related VOCs were generally lower in children with wheeze or asthma compared to healthy children. More importantly, several air pollutants, especially naphthalene, showed higher breath levels in children with wheeze or asthma. The association between pollutant-related VOCs and asthma or wheeze held true in a multivariable model with adjustment for other known disease risk factors. The most prominent VOC linked to asthma or wheeze was naphthalene, however cluster analysis also identified several other pollutants like benzene, toluene, ethylbenzene, and xylene (BTEX), which were linked to asthma or wheeze. Benzene and ethylbenzene originate mainly from traffic emissions and incomplete combustion, while other BTEX are related to indoor exposure caused by proximity to road traffic, cigarette smoking or cooking methods [26, 27]. In agreement with our results, other authors have also shown increased levels of these VOCs in patients with asthma [28, 29].

Naphthalene is a very volatile polycyclic aromatic hydrocarbon (PAH) predominantly present in the atmosphere in its vapor form [27]. It's a natural constituent in coal tar and oil, and a ubiquitous pollutant in the environment [27]. Large amounts are produced worldwide [30], primarily as an intermediate for other chemicals and airborne emissions originate from incomplete combustion, e.g. in traffic exhaust, during burning of biomass, or tobacco smoke [27]. Inhalation is being considered as the most significant route of exposure [31]. We did not find differences in breath naphthalene levels for children who lived in the proximity of larger roads or who were exposed to environmental tobacco smoke, but this analysis was only based on questionnaire data which is only a crude estimate of the real life exposure. Naphthalene concentrations can also be increased indoors [27], especially in buildings with attached garages or integrated fireplaces [27, 32]. A range of indoor building materials has been reported to emit naphthalene, e.g. caulking, carpeting, rubber or vinyl flooring [33]. Cabins of diesel- and gasoline-fuel cars have also been reported as relevant sources [34]. In our study, we did not have information about these potential exposures. Therefore, it would be interesting for future studies to include personal sampling devices. Diet has also been

described as a potential source of naphthalene [27], however, this contribution will be difficult to control for in cohort studies.

In children with pre-school wheeze, increased breath naphthalene levels have been observed before [35], but its potential role was not further discussed. Increased naphthalene levels have also been reported in blood and urine of children with asthma. In a study comparing PAH levels in 195 children, naphthalene serum levels showed one of the strongest associations with asthma [36] , while increased 2-naphthol levels, a naphthalene metabolite, could be found in urine of asthmatic children [37]. Increased PAH metabolite levels in urine of children with asthma were associated with higher disease-specific symptoms and consequently decreased lung function parameters [38]. All substances mentioned above were part of module 6 with considerable correlation among each other, which suggests regular exposure to a potential common source. Little is known so far regarding the potential pathomechanisms linking naphthalene and asthma. Murine studies showed that naphthalene could have local effects on the respiratory epithelium as well as systemic affects due to high plasma concentrations after inhalation [30, 39, 40], however, the exact role in asthma development is still uncertain.

In contrast to the strong pollution signature in children with asthma and wheeze, we found increased levels of flavor and lifestyle related VOCs in healthy subjects. Although speculative, one possible explanation could be a higher awareness of parents of asthmatic or wheezing children limiting exposure to strong flavorings, cosmetics, or cleaning products while parents of healthy children might be less restrictive.

Eleven of the compounds found to be different between healthy children and patients (**table 2)** have been related to asthma in previous studies [13, 14], however the overlap with our results is quite small. Recent reviews show that this is not uncommon in studies with asthma or wheezing patients [12, 13] and very likely due to differences in patient populations and in sampling and analysis methods. Furthermore, as we used a predefined target VOC dataset comprising of 158 VOCs, we might not have detected some of the previously reported VOCs in comparable trials with asthma patients.

We found no relationships of VOC profiles with clinical characteristics such as asthma control or risk for exacerbations [17, 18]. Similarly, VOCs were not associated with inflammatory

phenotypes, i.e. type 2 inflammation, confirming previous negative results from the adult arm of the ALLIANCE cohort [41]. In this respect it is important to note that about 73 % of the asthmatic children were under anti-inflammatory treatment when the breath collection was performed. We cannot exclude that this has reduced our chance of finding VOC patterns associated with inflammation, however we still observed significantly higher FeNO levels in the asthmatic children, indicating ongoing airway inflammation.

Naphthalene and two other VOCs of module 6 were increased in pre-school wheezers treated with ICS, supporting the proposed link between air pollution as a disease driving factor, as ICS treatment in this age group identifies children with a more severe phenotype. Additionally, children with pre-school wheeze and regular ICS treatment showed higher levels of three different long chain alkanes, which have previously been identified in COPD patients with respiratory tract infections caused by Rhinovirus [42]. This could potentially point towards an inflammatory phenotype driven by recurrent viral infections. However, it needs to be mentioned that one of these alkanes (undecane) was also part of module 4 and increased in healthy controls thus illustrating the complexity of the data and the difficulties in distinguishing exogenous and endogenous sources of VOCs.

A strength of our study is the choice of method for breath collection and VOC analysis which has been benchmarked in a previous trial [43]. Furthermore, our results were in concordance with several expected findings. For example, we were able to identify cigarette smoke related VOCs in children exposed to environmental tobacco smoke, and we saw significant correlations for lifestyle related VOCs between repeated measurements over at least one year. Correlations between near-subject room air and breath for VOCs related to hygiene product and detergents are in line with our “owncloud” concept [22].

It’s well known that the VOC composition of breath is influenced by many factors [44], among them demographic, diet, life style and recent activity. Therefore, we interpret the observed correlations of VOCs between the repeated measurements as remarkable, despite rather low correlation coefficients ($r \leq 0.6$) and consider these as an indication for stable living and lifestyle conditions.

However, we also have to report some limitations. Our data showed a marked center effect which we had observed in a previous multicenter study as well [24]. All analyses were performed using batch-corrected data and original data whenever possible (supplementary

material) and the comparable results indicate that batch correction did not introduce an unwanted bias to our data.

In line with our previous studies [21, 22, 24] we limited our GC-MS analysis to preselected 158 VOCs instead of performing a broad analysis of all potential peaks in a chromatogram. While the latter is likely to include more features for comparison, there is a risk for false identification of VOCs, which coelute at the same retention time. On the downside, by including only preselected VOCs, we cannot exclude to have missed VOCs which could have been relevant for the comparison between groups. Unfortunately, healthy control children were not age-matched to children with asthma or wheeze. However, only two VOCs out of 158 VOCs were actually age dependent, which makes a significant influence of age on the result unlikely. Overall, we focused our analysis on single VOC comparisons as well as the two modules 4 and 6, which comprised most of the significantly different VOCs found in the single VOC analysis. Some other VOCs modules were also significantly different between the study groups, however, further detailed analysis of all modules was beyond the scope of this manuscript.

In this study, we were able to identify VOCs associated with asthma and pre-school wheeze indicating an increased exposure of these children to environmental pollutants. Due to the observational nature of our study, we cannot deduce a causal relationship of these environmental exposures to the development of asthma. Still, the strong association confirmed in a multivariable analysis as well as the fact that the result was already seen in pre-school wheezers and in the case of naphthalene particularly in those more severely affected supports previous findings that exposure to pollutants might be an early life factor driving asthma pathology [7]. Although we could not find associations between VOCs and clinically relevant features as inflammatory phenotypes or risk for exacerbations, we consider non-invasive breath analysis an interesting tool to assess the exposome of children and to decipher the role of environmental exposures in the development of airway diseases. The increased levels of breath naphthalene and other PAH in asthmatic and wheezing children would be a strong argument to include breath sampling and potentially also personal sampling devices into future studies to better understand the role of environmental and especially indoor air pollutants.

Acknowledgment

We would like to express our gratitude to Manfred Elend († 17.04.21). Without his extraordinary experience in GC-MS analysis and his valuable advice, we would not have been able to collect the data of this study.

References

[1] Song P, Adeloye D, Salim H, Dos Santos J P, Campbell H, Sheikh A and Rudan I 2022 Global, regional, and national prevalence of asthma in 2019: a systematic analysis and modelling study *Journal of global health* **12** 4052

[2] Aumann I, Prenzler A, Welte T and Gillissen A 2014 Epidemiologie und Kosten von Asthma bronchiale in Deutschland - eine systematische Literaturrecherche *Pneumologie (Stuttgart, Germany)* **68** 557–67

[3] 2020 Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019 *Lancet (London, England)* **396** 1204–22

[4] GINA Main Report 2022 Front Cover

[5] Habib N, Pasha M A and Tang D D 2022 Current Understanding of Asthma Pathogenesis and Biomarkers *Cells* **11**

[6] Alashkar Alhamwe B, Alhamdan F, Ruhl A, Potaczek D P and Renz H 2020 The role of epigenetics in allergy and asthma development *Current opinion in allergy and clinical immunology* **20** 48–55

[7] Bush A 2021 Impact of early life exposures on respiratory disease *Paediatric respiratory reviews* **40** 24–32

[8] Mukharesh L, Phipatanakul W and Gaffin J M 2023 Air pollution and childhood asthma *Current opinion in allergy and clinical immunology* **23** 100–10

[9] Tiotiu A I, Novakova P, Nedeva D, Chong-Neto H J, Novakova S, Steiropoulos P and Kowal K 2020 Impact of Air Pollution on Asthma Outcomes *International journal of environmental research and public health* **17**

[10] Khatri S B *et al* Use of Fractional Exhaled Nitric Oxide to Guide the Treatment of Asthma: An Official American Thoracic Society Clinical Practice Guideline 2021 *American journal of respiratory and critical care medicine* **204** e97-e109

[11] Gaillard E A *et al* 2021 European Respiratory Society clinical practice guidelines for the diagnosis of asthma in children aged 5-16 years *The European respiratory journal* **58**

[12] Ibrahim W, Natarajan S, Wilde M, Cordell R, Monks P S, Greening N, Brightling C E, Evans R and Siddiqui S 2021 A systematic review of the diagnostic accuracy of volatile organic compounds in airway diseases and their relation to markers of type-2 inflammation *ERJ open research* **7**

[13] Peel A M, Wilkinson M, Sinha A, Loke Y K, Fowler S J and Wilson A M 2020 Volatile organic compounds associated with diagnosis and disease characteristics in asthma - A systematic review *Respiratory medicine* **169** 105984

[14] Neerincx A H, Vijverberg S J H, Bos L D J, Brinkman P, van der Schee M P, Vries R de, Sterk P J and Maitland-van der Zee A-H 2017 Breathomics from exhaled volatile organic compounds in pediatric asthma *Pediatric pulmonology* **52** 1616–27

[15] Dallinga J W, Robroeks C M H H T, van Berkel J J B N, Moonen E J C, Godschalk R W L, Jöbsis Q, Dompeling E, Wouters E F M and van Schooten F J 2010 Volatile organic compounds in exhaled breath as a diagnostic tool for asthma in children *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **40** 68–76

[16] Schleich F N *et al* 2019 Exhaled Volatile Organic Compounds Are Able to Discriminate between Neutrophilic and Eosinophilic Asthma *American journal of respiratory and critical care medicine* **200** 444–53

[17] Robroeks C M, van Berkel J J, Jöbsis Q, van Schooten F-J, Dallinga J W, Wouters E F and Dompeling E 2013 Exhaled volatile organic compounds predict exacerbations of childhood asthma in a 1-year prospective study *The European respiratory journal* **42** 98–106

[18] Olopade C O, Zakkar M, Swedler W I and Rubinstein I 1997 Exhaled pentane levels in acute asthma *Chest* **111** 862–5

[19] Fuchs O *et al* 2018 The all age asthma cohort (ALLIANCE) - from early beginnings to chronic disease: a longitudinal cohort study *BMC pulmonary medicine* **18** 140

[20] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften 2020 *Nationale VersorgungsLeitlinie Asthma – Langfassung* www.asthma.versorgungsleitlinien.de

[21] Küppers L, Holz O, Schuchardt S, Gottlieb J, Fuge J, Greer M and Hohlfeld J M 2018 Breath volatile organic compounds of lung transplant recipients with and without chronic lung allograft dysfunction *Journal of breath research* **12** 36023

[22] Holz O, van Vorstenbosch R, Guenther F, Schuchardt S, Trinkmann F, van Schooten F-J, Smolinska A and Hohlfeld J M 2022 Changes of breath volatile organic compounds in healthy volunteers following segmental and inhalation endotoxin challenge *Journal of breath research* **16**

[23] Cohen L *et al* 2019 A Modular Cytokine Analysis Method Reveals Novel Associations With Clinical Phenotypes and Identifies Sets of Co-signaling Cytokines Across Influenza Natural Infection Cohorts and Healthy Controls *Frontiers in Immunology* **10**

[24] Gaida A *et al* 2016 A dual center study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD *Journal of breath research* **10** 26006

[25] Lechner M, Moser B, Niederseer D, Karlseder A, Holzknecht B, Fuchs M, Colvin S, Tilg H and Rieder J 2006 Gender and age specific differences in exhaled isoprene levels *Respiratory physiology & neurobiology* **154** 478–83

[26] Vardoulakis S, Giagloglou E, Steinle S, Davis A, Sleeuwenhoek A, Galea K S, Dixon K and Crawford J 2021 *Indoor Exposure to Selected Air Pollutants in the Home Environment: A Systematic Review*

[27] Jia C and Batterman S 2010 A critical review of naphthalene sources and exposures relevant to indoor and outdoor air *International journal of environmental research and public health* **7** 2903–39

[28] Gahleitner F, Guallar-Hoyas C, Beardsmore C S, Pandya H C and Thomas C P 2013 Metabolomics pilot study to identify volatile organic compound markers of childhood asthma in exhaled breath *Bioanalysis* **5** 2239–47

[29] Arif A A and Shah S M 2007 Association between personal exposure to volatile organic compounds and asthma among US adult population *International archives of occupational and environmental health* **80** 711–9

[30] Yost E E, Galizia A, Kapraun D F, Persad A S, Vulimiri S V, Angrish M, Lee J S and Druwe I L 2021 Health Effects of Naphthalene Exposure: A Systematic Evidence Map and Analysis of Potential Considerations for Dose-Response Evaluation *Environmental health perspectives* **129** 76002

[31] Preuss R, Angerer J and Drexler H 2003 Naphthalene--an environmental and occupational toxicant *International archives of occupational and environmental health* **76** 556–76

[32] Hoffmann K, Krause C, Seifert B and Ullrich D 2000 The German Environmental Survey 1990/92 (GerES II): sources of personal exposure to volatile organic compounds *Journal of exposure analysis and environmental epidemiology* **10** 115–25

[33] Kang D H, Choi D H, Won D, Yang W, Schleibinger H and David J 2012 Household materials as emission sources of naphthalene in Canadian homes and their contribution to indoor air *Atmospheric Environment* **50** 79–87

[34] Jo W-K and Lee J-H 2009 In-Vehicle Levels of Naphthalene and Monocyclic Aromatic Compounds According to Vehicle Type *Environmental Engineering Research* **14** 180–5

[35] van de Kant K D G *et al* 2013 Exhaled breath profiling in diagnosing wheezy preschool children *The European respiratory journal* **41** 183–8

[36] Al-Daghri N M, Alokail M S, Abd-Alrahman S H, Draz H M, Yakout S M and Clerici M 2013 Polycyclic aromatic hydrocarbon exposure and pediatric asthma in children: a case-control study *Environmental health : a global access science source* **12** 1

[37] Lin T-J, Guo Y L, Hsu J-C and Wang I-J 2018 2-Naphthol Levels and Allergic Disorders in Children *International journal of environmental research and public health* **15**

[38] Cilluffo G *et al* 2022 Effects of Polycyclic Aromatic Hydrocarbons on Lung Function in Children with Asthma: A Mediation Analysis *International journal of environmental research and public health* **19**

[39] Kapuci M, Ulker Z, Gurkan S and Alpsoy L 2014 Determination of cytotoxic and genotoxic effects of naphthalene, 1-naphthol and 2-naphthol on human lymphocyte culture *Toxicology and industrial health* **30** 82–9

[40] Carratt S A, Morin D, Buckpitt A R, Edwards P C and van Winkle L S 2016 Naphthalene cytotoxicity in microsomal epoxide hydrolase deficient mice *Toxicology letters* **246** 35–41

[41] Holz O *et al* 2021 Breath volatile organic compounds and inflammatory markers in adult asthma patients: negative results from the ALLIANCE cohort *The European respiratory journal* **57**

[42] Kamal F *et al* 2021 Virus-induced Volatile Organic Compounds Are Detectable in Exhaled Breath during Pulmonary Infection *American journal of respiratory and critical care medicine* **204** 1075–85

[43] Wilkinson M, White I, Hamshere K, Holz O, Schuchardt S, Bellagambi F G, Lomonaco T, Biagini D, Di F F and Fowler S J 2021 The peppermint breath test: a benchmarking protocol for breath sampling and analysis using GC-MS *Journal of breath research* **15**

[44] Filipiak W *et al* 2012 Dependence of exhaled breath composition on exogenous factors, smoking habits and exposure to air pollutants *Journal of breath research* **6** 36008

Legends to figures

Figure 1. Study population. Breath VOCs were analyzed from 142 subjects. 40 subjects were sampled on two visits, which were at least one year apart. The Hannover site had a focus on wheezers and started the recruitment for VOC sampling later during the study. The one patient with asthma from Hannover was planned to be included as wheezer but the visit could not be scheduled before the 6th birthday, therefore based on the definition criteria, this subject was considered as child with asthma.

Figure 2. Heatmap of VOC modules: VOCs are first clustered together based on pairwise Pearson correlations. A reliability score (i.e., the fraction of times that a given VOC is assigned to the same cluster) is calculated over 1,000 permutations of participants and is used to assign VOCs to modules. The color bars on the left side depict the module membership for each VOC (a full list is given in supplement **table S4**), and the coloring in the heatmap represents the reliability score of each VOC.

Figure 3. Differences in the levels of naphthalene between groups. Median and interquartile ranges. Individual data is shown in grey open circles. *** $p<0.001$, ** $p<0.01$.

Table 1. Demographics

		n	asthma	n	wheeze	n	healthy
Number of subjects	M/H	51	50 / 1	55	30 / 25	36	33 / 3
Gender	female/male	51	16 / 35	55	20 / 35	36	12 / 24
Age	years	51	12.9 ± 3.5***	55	5.7 ± 1.3***	36	9.7 ± 3.3
Age at inclusion	years	51	10.8 ± 3.1	55	2.7 ± 1.5***	36	9.7 ± 3.3
Atopy§	% pos.	46	85	43	58	35	37
FeNO	ppb	44	24.7 (11.1;36.9)**	23	9.6 (5.1;19.9)	24	9.1 (5.6;20.2)
FEV1	z-score	49	-0.53 (-1.09;0.23)***	48	-0.39 (-1.09;0.55)***	36	0.24 (-0.24;1.26)
FEV1/FVC	z-score	49	- 0.57 (-1.34;0.15)***	47	- 0.28 (-1.18;0.87)	36	0.16 (-0.67;1.12)
Blood leucocytes	cells/µL	46	7350 (6100;8300)	42	7500 (6400;9200)**	34	6675 (5500;7730)
Blood eosinophils	cells/µL	46	358 (198;570)*	42	381 (260;582)**	34	201 (139;395)
Blood neutrophils	cells/µL	46	3327 (2654;4300)	42	3430 (2603;4304)	34	3046 (2452;3696)
Total IgE	kU/L	45	232 (162;562)**	41	111 (35;302)	34	110 (41;240)
Asthma medication:							
SABA	n (%)	51	19 (37 %)	55	15 (27 %)		
Montelukast	n (%)	51	2 (4 %)	55	4 (7 %)		
ICS	n (%)	51	10 (20 %)	55	13 (24 %)		
LABA/ICS	n (%)	51	26 (51 %)	55	8 (15 %)		
wheezing episode (yes)$	n (%)	51	13 (26 %)	55	25 (45 %)		
≥ 1 exacerbation (yes)$	n (%)	51	2 (4 %)	55	5 (9 %)		

M=Munich, H=Hannover
Mean ± SD or median (25%;75% quartile) are presented. §: specific IgE >= 0,7 against at least 1/36 food- and aeroallergens.
Medication: number of children with respective treatment in the month before the VOC visit
Wheeze episode: symptoms of wheeze which required treatment with salbutamol for > 2 / 7 days
Exacerbation: wheeze episode which required treatment with systemic steroids or admission to hospital
$: refers to 12 month prior to the VOC collection visit
***p<0.001,
**p<0.01,*p<0.05

Table 2. Differences between healthy controls, asthmatic children, and children with wheeze

VOC	Module 4	Asthma vs. Healthy		Wheeze vs. Healthy		Fragrance / Flavor	Food	Personal Care	Cleaning	Disinfectant	Air pollutant
Lower in Diseased Patients		**pMWU**	**pBH**	**pMWU**	**pBH**						
unidentified C10H18 (**a**)	x	<0,001	**<0,001**	<0,001	**0,007**						
Geranyl acetate	x	<0,001	**0,003**	0,003	**0,035**	x	x	x	x		
Pentadecane		<0,001	**0,003**	0,006	**0,042**	x	x	x			
Menthyl acetate	x	<0,001	**0,005**	0,001	**0,028**	x	x	x	x	x	
Decanoic acid methyl ester	x	0,001	**0,009**	0,002	**0,033**	x			x		
Decanoic acid	x	<0,001	**0,009**	0,008	**0,054**	x	x	x	x		
unidentified VOC (**a**)		<0,001	**0,009**	0,064	0,213						
Neryl acetate	x	0,002	**0,019**	0,038	0,144	x	x	x	x		
2-Propanol		0,002	**0,018**	0,128	0,302	x	x	x	x	x	
1-Decanol		0,002	**0,020**	<0,001	**0,007**	x	x	x	x	x	
Ionone		0,002	**0,017**	<0,001	**0,007**	x	x	x	x		
Citronellol	x	0,001	**0,015**	0,002	**0,033**	x	x	x	x		
b-Linalool	x	0,001	**0,017**	0,129	0,302	x	x	x	x	x	
Decanal		0,003	**0,023**	0,025	0,120	x	x	x	x	x	
Linalyl acetate	x	0,005	**0,031**	0,158	0,328	x	x	x	x	x	
Dioctyl ether		0,003	**0,025**	0,078	0,222	x		x	x		
Camphene	x	0,002	**0,018**	0,057	0,201	x	x	x	x	x	
Undecane	x	0,003	**0,025**	0,005	**0,038**	x	x	x			
2,5-Dimethylfuran		0,251	0,425	0,004	**0,037**	x	x				
t-Butyl alcohol		0,115	0,261	0,003	**0,034**	x	x	x	x		
unidentified Terpene	x	0,030	0,107	0,004	**0,038**						
3,7 - Dimethyloctan-1-ol	x	<0,001	**0,005**	0,003	**0,034**	x		x	x		
Octanoic acid		0,007	**0,043**	0,385	0,570	x	x	x	x		
Nonane		0,007	**0,043**	0,033	0,141						
Hexanoic acid		0,401	0,554			x	x	x	x		
Myrcene	x	0,029	0,107			x	x	x	x		

VOC	Module 6	Asthma vs. Healthy		Wheeze vs. Healthy		Fragrance / Flavor	Food	Personal Care	Cleaning	Disinfectant	Air pollutant
Increased in Diseased Patients		**pMWU**	**pBH**	**pMWU**	**pBH**						
Isoprene (**b**)		<0,001	**<0,001**	0,318	0,507						
1-Phenoxypropan-2-ol	x	0,005	**0,034**	0,039	0,144			x	x	x	
Cycloheptasiloxane	x	0,021	0,093	0,004	**0,037**			x	x		
Benzoic acid	x	0,022	0,093	<0,001	**0,008**	x	x	x	x		
1-Nonene	x	0,068	0,179	0,004	**0,037**						
1,2-Propanediol	x	0,23	0,40	0,001	**0,028**	x	x	x	x	x	x
2-Ethylhexanol		0,008	**0,047**	0,005	**0,038**	x	x				x
Limonene		0,051	0,146	0,001	**0,028**	x	x	x	x		
Naphthalene	x	<0,001	**0,005**	<0,001	**0,007**						x
Ethylbenzene	x	0,070	0,179	0,159	0,328						x
Benzene	x	0,046	0,138	0,344	0,537						x
p-/m-Xylene	x	0,144	0,297	0,151	0,328						x
o-Xylene	x	0,238	0,411	0,114	0,280						x
Toluene	x	0,113	0,260	0,036	0,144						x
Hexanoic acid				0,002	**0,034**	x	x	x	x		
Myrcen				0,005	**0,038**	x	x	x	x		

The table contains all compounds which were different between healthy and asthma or healthy and wheeze

Additional information was provided for other airpollutants irrespective of the level of significance

(**a**) sign. correlation with e.g. Geranyacetate and other lifestyle VOCs suggests common source

(**b**) endogenous VOC, difference between groups mainly due to age dependency

Occurence and main use information based mainly on European Chemical Agency (ECHA)

pMWU: p-value uncorrected using non-parametric testing

pBH: p-value corrected for multiple testing (Benjamini-Hochberg)

For more details on clusters refer to supplement table S4

Table 3. Multivariable regression model investigating the association between module 6 and asthma or pre-school wheeze while controlling for known risk factors for asthma development

	Estimate	SE	Odds ratio	98% CI	p-value
Module 6	0.83	0.29	2.28	1.34 - 4.17	0.004
Atopy	1.96	0.53	7.08	2.62 - 21.0	<0.001
Age	-0.02	0.06	0.98	0.87 - 1.10	0.7
Gender (male)	-0.11	0.51	0.9	0.32 - 2.42	0.8
parental asthma	0.1	0.56	1.10	0.37 - 3.42	0.9
Eosinophils ≥470/µl	-0.87	0.59	2.39	0.79 - 8.05	0.14
Cigarette smoke exposure	-0.001	0.59	1.00	0.32 - 3.34	>0.9